Ultrafast Lasers and Optics For Experimentalists

J D Pickering

Second Edition

Contents

Dedication

For Chioma.

About the Author

James David Pickering is a Lecturer in Physical Chemistry at the University of Leicester.

Originally from Braintree, Essex, he attended Notley High School and Braintree Sixth Form, and obtained his MChem in Chemistry from Jesus College, University of Oxford, and his Ph.D from Aarhus University. Before arriving at Leicester he undertook postdoctoral research at both Oxford and Aarhus, where he taught both physical chemistry and mathematics. His research interests lie in ultrafast and nonlinear spectroscopy.

Acknowledgements

Whilst the cover shows a single author, in reality this book could not have been written without the knowledge and experience gained over many years of coffee-fuelled technical discussions and thousands of hours spent in the lab with countless colleagues, supervisors, students, and friends. However, there are specific individuals that have made a significant contributions to the production of this text, whom I would like to mention specifically.

The first thanks must go to Claire Vallance (University of Oxford), as without her encouraging words and putting me in touch with IOP, I would not have considered turning the initial lecture handout into this book. In a similar vein, Michael Burt (University of Oxford) was always encouraging, and many of the pitfalls we encountered together are recounted here so they can be avoided by future generations. Both Claire and Michael also helped to proof read parts of this manuscript. Ensuring a manuscript like this is as error-free as possible is a challenge, and I am especially grateful for the expert proof reading and advice provided by Adam Chatterley (Aarhus University), Constant Schouder (CEA Saclay), and Jan Thøgersen (Aarhus University). Useful discussions were also had with Marco Arrigoni and Jonas Berzinš from Light Conversion about Yb laser media – and they kindly agreed to provide a figure showing OPA tuning curves. Finally, some of my more diligent students from Merton College gave invaluable advice on explanations that were unclear or confusing, and the advice of Beth, Nana, Jia Jie, and Frank was particularly valued. Thanks must also go to those who emailed me pointing out errors they had spotted in the first edition – any errors that still remain are my own.

Finally, any teacher is aware that the only way to truly learn and understand a topic is to teach it to others. I have been, and still am, fortunate enough to work with and teach many outstanding students, and much of the time I feel I learn more from them than they do from me. I am indebted to them all.

Common Terminology

You will encounter the following terms throughout the book, they are briefly defined here for quick reference. Familiarity with basic concepts of wavelength, frequency, etc... is assumed.

- **Pulse Duration**, $\Delta\tau$: The FWHM width of our laser pulse in the time domain. Tells us how short our pulse is.
- **Bandwidth**, $\Delta\omega$: The FWHM width of our laser pulse in the frequency domain. Tells us the range of colours present in our pulse.
- **Spectral Phase**, $\phi(\omega)$: The phase ϕ of a particular frequency ω in a laser pulse.
- **Speed of Light**, c_0: The speed of light in a vacuum. $2.99 \times 10^8\,\mathrm{m\,s^{-1}}$.
- **Refractive Index**, n: A factor which determines the speed of light in a medium other than vacuum (in a vacuum $n = 1$).
- **Phase Velocity**, v_p: The speed of a light wave in a given medium. Related to the speed of light in a vacuum by $\frac{c_0}{n}$.
- **Chromatic Dispersion**: The spreading out of frequency components in time as a pulse propagates due to their different phase velocities. Also called 'temporal dispersion' or simply 'dispersion'.
- **Chirp**: If the frequency of the pulse changes with time, it is 'chirped'.
- **Transform Limit**: If the frequency of the pulse does not change with time, it is 'transform limited'. A transform limited pulse has the shortest pulse duration for a given bandwidth.
- **Propagation direction**: The direction of travel of a laser beam, defined by the direction of the wave vector, $\mathbf{k}$. The *wavenumber*, k, is the magnitude of the wave vector.
- **Polarisation direction**: The direction of the oscillations of the electric field of the laser beam. The polarisation plane is orthogonal to the propagation direction.

- **Beam waist**, ρ_0: Defined as the radius from the centre of the beam at which the intensity is reduced to a factor of $1/e^2$ of the peak intensity.
- **Rayleigh Length**, z_R: Defined as the distance along the propagation direction at which the beam waist is $\sqrt{2}$ times larger than at it's smallest point.
- **Repetition rate**, T_{rep}: How many pulses per second a laser produces.
- **Pulse energy**: The total energy contained within each pulse.
- **Average Power**: The average power (energy per unit time) transmitted in the laser beam.
- **Peak Power**: The maximum power transmitted in a single laser pulse.
- **Fluence**: The energy per unit area of a given laser pulse.
- **Intensity**: The power per unit area of a given laser pulse.

Foreword

In the 1960s, the same decade as the invention of the laser, the writer Frank Herbert lamented that the 'scientists were wrong, that the most persistent principles of the universe were accident and error'. Anyone labouring in an academic laboratory, strained by the fast rate of staff and student turnover, will eventually encounter the consequences of this adage. This will be truer still when the research in question involves the intersection of several scientific fields, where the traditional divisions used to demarcate undergraduate degrees break down. As such, any postgraduate students conducting research in one of these areas, as the author James Pickering did when he undertook a PhD in chemical physics, will often find that resources are aimed at the expert rather than the student. Researchers involved with laser applications are notably vulnerable. Lasers are used in a wide variety of techniques in modern experimental chemistry and physics, particularly due to their intersection with analytical methods such as mass spectrometry, but they form only a small part of undergraduate curricula and younger students often only have limited access to them due to the risks involved. This is a real detriment that only serves to compound the potential for 'accident and error' raised by Herbert.

This text counters the above problem by focusing on the practical aspects of ultrafast laser use, and is written as a handbook that students can refer to as needed. The concepts are clearly explained at a level that will be comfortable for upper-year undergraduates, and are written in a way that will help provide a common language that students can use to accurately and confidently present their work to a broader scientific audience. Any scientist delving into ultrafast laser applications will find something to learn from this book. The first section covers the nature and generation of ultrafast laser pulses, their dispersion through media, and techniques to measure them. The second describes how to prepare an instrument beamline for ultrafast pulses, including pragmatic descriptions of day-to-day considerations such as purchasing optics and optomechanics, overlapping and focusing beams, and outlining a laser table arrangement. The latter aspect is a major achievement of this book: case studies of laboratories are rare in literature, but are invaluable in helping young scientists to make a start while avoiding mistakes.

James Pickering is a talented scientist with many years of experience in research and instrument development. He refined his teaching ability while working as a teaching fellow at the University of Leicester and as a tutorial lecturer at the University of Oxford. Throughout this, he has always demonstrated an ability to straightforwardly explain complex concepts

to students in an approachable way. I have been lucky to work with James at various stages of his career, including his master's degree at the University of Oxford, PhD at Aarhus University, and as his postdoctoral supervisor during his return to Oxford in 2019. He is always eager to talk about the practical aspects of science, and this aspect of his nature is evident in this book. James was responsible for developing the beamline in my own laboratory. The 'accident and error' we shared during those days motivated this text, and I believe the reader will benefit from that experience.

- Michael Burt, Oxford, 2020.

Preface

This book grew in 2020 as I was preparing to give a short talk to my research group about the basic principles of ultrafast optics and lasers that are relevant for non-specialists using 'turn-key' systems. The project rather ran away with me and I ended up producing an extended handout as a supplement to the slides, which eventually turned into this book. My intention is that this will be useful as an introductory text and basic reference for anyone new to the exciting world of ultrafast optics and lasers - and is specifically aimed towards those entering the field without an extensive formal background in optics or laser physics (such as myself). As somebody who initially trained as a chemist, but that moved into a field on the chemistry-physics boundary for a PhD, there was a serious learning curve to becoming competent in use of the ultrafast laser systems that the research work demanded. To my knowledge, there are minimal written resources codifying the basic principles of practical use of ultrafast laser and optics - with most knowledge being passed on orally from postdocs to students. This book aims to fill that gap, and is written largely to be 'the book I wish I had had when starting my PhD'.

The aim therefore is to present the key information about practical use of ultrafast lasers optics in a way that is accessible to non-specialists who don't have the formal background in optics, electromagnetism, or laser physics that would be gained in a typical undergraduate physics course. Minimal prior knowledge is assumed (only that which may have been gained in typical undergraduate spectroscopy or quantum mechanics courses). Emphasis throughout is placed on qualitative, intuitive, understanding rather than dense mathematics. Hopefully it will allow students and researchers using these systems to not see lasers as a mysterious 'black box' that they use in their experiments, but to start to appreciate the underlying physics that drives them. As more and more research groups across the natural sciences are starting to use ultrafast laser systems in their research, I hope that this book can prove to be a useful companion for people moving into the field, both young and old.

The structure of the book is divided into two parts. The first part, '*Fundamentals*', is written with the aim that it can bring a final year undergraduate student unfamiliar with ultrafast lasers up to speed quickly with accessible descriptions of the most important phenomena surrounding ultrafast laser physics. The basic principles of laser action will be discussed, and the properties of laser light and laser beams – with a special focus on those properties that especially differentiate an *ultrafast* laser from a non-ultrafast laser. Useful phenomena within nonlinear optics will then be covered, focussing on those phenomena which are critical

to a working understanding of ultrafast pulse generation and characterisation.

The second part of the book, '*Practical Ultrafast Optics*', essentially aims to take the background presented in the first part and walk the student through ultrashort pulse generation, characterisation, and manipulation. Things that perhaps seem mundane to a more experienced campaigner, but can often seem bewildering to a new student (such as what to look for when buying optics) are discussed in detail. Techniques for practical alignment of optical beamlines are written down, and finally some examples of beamlines for an ultrafast experiments are presented.

It is my hope that the book can be used not only by new students, but also by more experienced scientists who are moving into this field. For this reason, it has been written (as far as possible) to be easily dipped in and out of - so that an experienced postdoc does not need to read lots of information on basic laser action in order to follow the narrative. Hopefully this book can form a useful reference text for people in many laboratories. Throughout suggestions for further reading are given, as this book is only intended to provide an accessible gateway into this vast topic.

Update for 2nd edition in July 2024. Since writing the original version of this book around 5 years ago, much has changed. My own career has moved away from a 'molecular beams, small molecules, and titanium sapphire' world towards a 'solutions, proteins, and ytterbium' one. Encountering new lasers and new techniques over the last five years has been a joy and I realised that the initial edition of this book was missing a lot of information that would render the book useful to a wider range of readers. Expanded introductory chapters dig a bit more deeply into certain important concepts, and throughout the 'new' Ytterbium based laser has been used in examples alongside the venerable Ti:Sa. Further additions of worked examples throughout, updated figures, and expanded chapters on practical optics I hope will be found useful for a wider range of users, and many errors have been caught and corrected. I was touched by the number of people contacting me following publication of the first edition saying that the book was found genuinely useful – I hope that this expanded and improved new edition makes it even more so.

The book was typeset using LaTeXand all figures (except photographs and ones taken from other sources with permission) were made using Python and Matplotlib. Further information and some resources can be found at https://jamesdpickering.com/.

Part I

Fundamentals

Chapter 1

Lasers and Laser Light

In this chapter fundamental aspects of laser physics are presented, as well as properties of laser light and laser beams that are not specific to ultrafast lasers, but essential for someone new to the field to understand (such as polarisation states and Gaussian beam propagation). On the ultrafast side, key concepts such as the time-bandwidth product and the transform limit are introduced to set the stage for further exposition later.

LASER is an acronym, standing for **L**ight **A**mplification by **S**timulated **E**mission of **R**adiation. Lasers are one of the most widely used experimental tools in the physical natural sciences. Fundamentally, a laser is a device that produces highly directional electromagnetic (EM) radiation. In the visible region, this radiation is visible light, hence the name, but lasers can emit radiation across the EM spectrum – but all of this radiation is called 'light' for short. In this chapter we will learn the basic principles of laser operation and discuss key points around the light produced by lasers.

1.1 Why Lasers?

The utility of lasers for experimental scientists lies in several areas. Firstly, the directionality of the emitted light means that all of the light can be efficiently directed onto the sample of interest, which is typically not the case when a sample is irradiated using a source like a lamp or LED. Secondly, the intensity of the light can be many times higher than is easily obtained using one of these sources. Thirdly, the ease with which a wide variety of different energies of light can be produced means that a wide variety of different interactions of light with matter can be studied. Finally, there are many qualities of the light aside from the frequency that can be tuned. Some of these will be discussed in more detail later, but most critically for the topic of this book is the ability for lasers to produce light in very short pulses. These pulses can last less than a femtosecond (1×10^{-15} s) in time. These short pulses can be beneficial in a number of areas:

- They allow dynamics of events occurring on very short timescales to be measured –

electronic control of timing is (at best) several orders of magnitude slower than that which can be achieved using an ultrafast laser.

- Ultrashort laser pulses produce light with a very high intensity, allowing nonlinear effects to be exploited, and enabling the study of matter under 'extreme' conditions.
- More indirectly, we will see that having an ultrashort laser pulse requires that there are many frequencies of light in the laser pulse. This can be useful in many kinds of spectroscopy.

Lasers that produce these ultrashort pulses are termed **ultrafast lasers**, which we will define as 'lasers producing pulses with a duration of less than one picosecond'. While this book focusses on ultrafast lasers, large parts of the book are applicable to all kinds of laser. I will assume that the reader is working in an area of science where ultrafast lasers are common tools but are not the main focus of the research - such as in multiphoton microscopy, ultrafast spectroscopy, or nonlinear spectroscopy. The assumed knowledge is that which would be gained in a typical undergraduate science course, and many key physical concepts are recapped in the appendix to help readers from less physical backgrounds.

1.2 Laser Action

As the acronym suggests, lasers work via stimulated emission of radiation. Figure 1.1 shows a basic schematic of how a four-level laser functions. Other energy level structures are possible (such as a 'three-level laser'), depending on the type of laser, but the basic principles of laser operation are common across all laser types. Here we choose to focus on the four-level laser as this is how most commonly encountered lasers function, including the Titanium Sapphire and Yb-doped laser[1], which are the two most common types of ultrafast laser at the time of writing.

There are four distinct energy levels in the system shown in Figure 1.1. A ground level (1), is excited into level 4, often via optical excitation with photons of energy E_{pump} provided by a flashlamp or another laser (known as a **pump laser**). Level 4 is shown as a broad band of levels, rather than a single level, and this is often the case in real systems. Precise frequency matching of the pumping radiation to a specific upper level is then not necessary, allowing the use of broadband excitation sources.

There is then rapid population transfer from the excitation level(s) 4, into the upper lasing level, level 3. Lasing occurs on the 3→2 transition, and requires a **population inversion** to exist between levels 2 and 3. The idea of a population inversion may be familiar from undergraduate spectroscopy courses, and is the situation in which there is more population in the upper level (3) than the lower level (2) – sometimes referred to as 'non-Boltzmann behaviour'. Achieving the population inversion requires that the population of level 3 is kept high, while the population of level 2 is kept low. This is achieved by excitation from level 1

[1] Yb-doped laser operation can, in some cases, be considered 'quasi three-level', but for our purposes considering the four-level system here is sufficient.

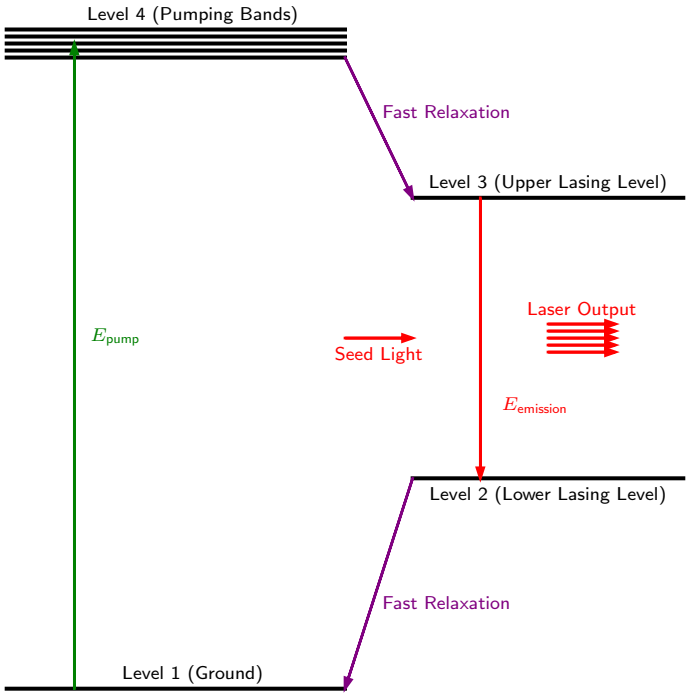


Figure 1.1: A schematic of basic laser action. Lasing is achieved between levels 2 and 3, via excitation from level 1 to level 4.

to level 4, followed by the rapid relaxation from level 4 into into level 3. Stimulated emission from level 3 into level 2 then produces the laser radiation, with the energy gap between these levels, E_{emission}, determining the frequency (colour) of the laser radiation. This emission is either stimulated using a **seed laser** with photon energy E_{emission} as shown, or relies on spontaneous emission from level 3 of a nearby system to induce the stimulated emission. Another rapid population transfer from level 2 back to the ground state has the effect of ensuring that the population of the lower lasing level remains low, such that a large population inversion can be maintained for efficient lasing.

Clearly, there must be some material in which all these levels exist, and this is known as the **gain medium** or **laser medium**. There are a huge number of different possible laser media, spanning the gas, solution, and solid phases. Here we will limit our discussion to solid-state laser media, as these are the type most commonly found in modern ultrafast lasers[2]. Different laser media have different emission energies (the gap between levels 2 and 3), and different pump energies (the gap between levels 1 and 4).

Solid-state laser media are centered around an **active ion**. This ion is doped into a host glass or crystal, and the energy levels of the ion *within this crystal field* are what determine the pump and emission energies. Common active ions are Nd^{3+}, Ti^{3+}, and Yb^{3+}, and they

[2]The earliest femtosecond lasers were 'dye lasers', where the gain medium was a solution of organic dye. These are now relatively rare.

can be doped into a variety of hosts to give a range of different laser actions. Examples of lasing schemes that can be produced by these active ions doped into common hosts[3] are shown in Table 1.1, note the standard nomenclature of '*ion:host*'. The laser gain media for two of the most commonly encountered ultrafast lasers in research are:

- Sapphire (Aluminium Oxide - Al_2O_3) doped with Ti^{3+} ions to make the **Titanium Sapphire** (Ti:Sa) laser.
- KGW (Potassium Gadolinium Tungstate) doped with Yb^{3+} ions to make the **Ytterbium KGW** (Yb:KGW) laser.

These are the lasers we will use as example cases throughout this book. Other ultrafast lasers may be encountered using Yb active ions with different hosts. Besides these, two very commonly encountered laser gain media are YAG (Yttrium Aluminium Garnet) and YLF (Yttrium Aluminium Fluoride). Both of these are common hosts for Nd^{3+} ions to make the Nd:YAG and Nd:YLF laser, which are used as pump lasers in many ultrafast laser systems, and will likely be encountered in a lab where Ti:Sa lasers are in use.

Table 1.1: Common solid-state laser media and their characteristic lasing wavelengths.

Laser Medium	Pump Wavelength (nm)	Emission Wavelength (nm)
Ti:Sapphire	400-600	650-1100
Yb:KGW	981	1020-1060
Nd:YAG	730-820	1064
Nd:YLF	730-820	1053

Note that the emission energies in Table 1.1 do not need to correspond to a single well-defined energy. Levels 2 and 3 in Figure 1.1 can just as easily be broader bands that lead to a wider lasing bandwidth – this is infact necessary for ultrafast laser pulse generation. The pump and emission energies shown are merely selected for illustrative reasons. Titanium Sapphire has the broadest emission bandwidth of any solid-state laser medium, which makes it exceptionally well suited for production of ultrashort pulses. Ti:Sa lasers are historically the most common type of ultrafast laser. Ti:Sa is generally pumped at around 527 nm to produce emission between 700 nm and 900 nm. More recently (over the last decade), Yb-doped lasers are becoming increasingly common in ultrafast laboratories worldwide, as they allow higher average powers and higher repetition rates to be achieved than Ti:Sa lasers. Throughout this book we will use the Ti:Sa and Yb-doped lasers as illustrative cases. We will now turn to a discussion of the light produced by lasers, and cover some of the fundamental physics that is needed for later discussion.

[3] Many laser media can be pumped to lase at a variety of different wavelengths. For example, Nd:YAG can be pumped to lase at 1440 nm, as an alternative to it's 'usual' emission at 1064 nm.

1.3 Laser Light

Looking back at the diagram in Figure 1.1, our laser will produce light with a photon energy E_{emission}. It will be most useful for us to think of this light as an electric field moving in space and time (see Appendix A for a refresher). There are many properties of this light field that are tunable and interesting to us as laser scientists, but the a key property to discuss here is the **temporal shape** of the field. There are two main options here:

- The light field can form a **continuous wave**, which is a perfect sinusoidal wave of a single frequency. This is called a **cw laser**, or a laser operating in **cw mode**.
- The light field can take the form of a train of **pulses**. A laser like this is called a **pulsed laser**, and if the the temporal width of a pulse is less than around one picosecond, we say it is an **ultrafast pulsed laser**, or that the pulses are **ultrashort**.

The latter mode of operation is what will concern us for the rest of this book. An example of a pulse train is shown in Figure 1.2.

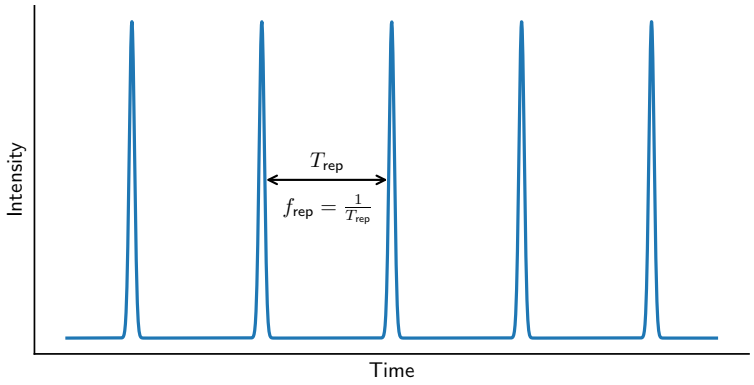


Figure 1.2: A pulse train from a laser. The *repetition rate* of the laser f_{rep} is defined as the number of pulses produced per second, which is the inverse of the time between pulses T_{rep}. Note that in an ultrafast laser, the pulse duration (width of the pulse in time) is normally 6-10 orders of magnitude shorter than T_{rep} - much narrower than depicted here.

1.3.1 Pulsed Lasers and Time-Bandwidth Products

It is now useful to remind ourselves of the general time-energy uncertainty relation, a recasting of the familiar quantum-mechanical position-momentum uncertainty relation:

$$\Delta E \Delta t \geq \frac{\hbar}{2} \tag{1.1}$$

This relation tells us that a state that has a very well defined energy (small ΔE) must exist for a very long time (large Δt), and vice versa. This has profound implications for our pulsed laser, as if the laser output is **pulsed**, then by definition it exists for a finite length of time - so Δt is very small for an ultrafast laser. A small Δt means that ΔE is relatively

large (Equation 1.1) for our laser pulse, and means that there is an uncertainty associated with the photon energy (colour) of the light in the pulse. This can be simplistically thought of as the emission coming from a broad band of levels (as drawn in Figure 1.1), rather than a single level with a well-defined energy. Thus, for our laser output to be pulsed, there must be more than one single photon energy, or more than one single wavelength of light, making up the light in the pulse. The principle you learnt at school about lasers being monochromatic is, at best, an oversimplification!

Equation 1.1 is true in a general sense, but we need to be a bit more precise about how exactly we define ΔE and Δt for our laser pulses, since these depend on the temporal and spectral shape of the pulse. Generally ultrashort laser pulses can be well described as **Gaussian pulses**, with a Gaussian profile[4] in the time domain. We can then identify the width of our Gaussian pulse in the time domain as $\Delta\tau$; this is defined as the full-width-at-half-maximum (FWHM) of the pulse, as illustrated in Figure 1.3 (the blue curve). The quantity $\Delta\tau$ is called the **pulse duration**.

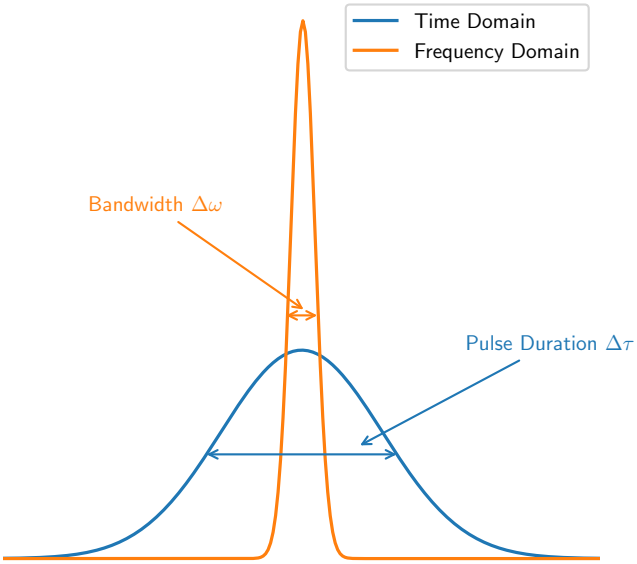


Figure 1.3: A typical Gaussian shaped laser pulse represented in both the time domain (blue) and frequency domain (orange). The definitions of the temporal width and bandwidth are annotated.

This is the orange curve in Figure 1.3 shows the same pulse in the frequency domain. Frequency is directly proportional to energy ($E = \hbar\omega$), so looking in the frequency domain is equivalent to looking in the energy domain. We find the FWHM of this frequency domain pulse[5] which is called the **pulse bandwidth** $\Delta\omega$. This bandwidth can also be called the

[4] Different types of laser oscillators produce different pulse shapes. Another common pulse shape is the sech^2 pulse shape. Generally you will find that laser manufacturers fit their pulse durations to whichever shape allows them to quote the lowest number for the pulse duration.

[5] The Fourier transform of a Gaussian is also a Gaussian, so we can define the FWHM in the frequency

'spectral width', as it is the width of the spectrum that our pulse spans. If there are many colours present in our pulse, then it has a broad bandwidth, and vice versa.

Application 1.1: Energy Units

At this early stage it is worth clarifying a common source of confusion for people new to this topic. In the discussion above, we seamlessly moved from energy to frequency units on the basis that they are proportional to each other. It is similarly common in this field and related ones to mix up energy, frequency, and wavelength units in discussions – so you need to get used to it!

Furthermore, we actually defined our frequency as the **angular frequency**, ω. The *angular frequency*, ω is related to the *temporal frequency* ν by:

$$\omega = 2\pi\nu \tag{1.2}$$

The temporal frequency, ν has units of Hz (s^{-1}), whereas the angular frequency has units of $\mathrm{rad\,s^{-1}}$. Again, these quantities are clearly proportional to one another, so it doesn't *really* matter which we use. Angular frequencies are often neater to use in equations involving sinusoidal waves as they avoid lots of factors of 2π floating around. We will generally use angular frequency, ω in this book, except for in some cases where it is more natural to use the temporal frequency.

An important point to note here is that the spread of frequencies is not, in reality, continuous as Figure 1.3 suggests. In a very well resolved spectrum we would see many individual frequencies in our total spectrum. This arises because generally the light has to form a *standing wave* within the oscillator cavity, so we can **only** have frequency components present where the wavelength is such that an integer number of waves will exactly fit into our cavity length. Mathematically this means that:

$$\text{Wavelength of an allowed component} = \frac{\text{Cavity Length}}{\text{Integer}}$$

So in reality we would see wavelengths at integer fractions of our cavity length, if we had a powerful spectrometer capable of resolving them. Overall, the gain medium produces a range of different photon energies as determined by the emission bandwidth. Only a subset of these are allowed by the cavity length. The interplay between these effects is illustrated in Figure 1.4.

Turning back to discussion of bandwidth, we can now use our new definitions for bandwidth and pulse duration to define a new kind of uncertainty relation. It is most normal to do this using the temporal frequency bandwidth $\Delta\nu$ rather than $\Delta\omega$ - but it can be described using either definition equivalently. We define a product $\Delta\nu\Delta\tau$. This is referred to as the

domain exactly as we did in the time domain

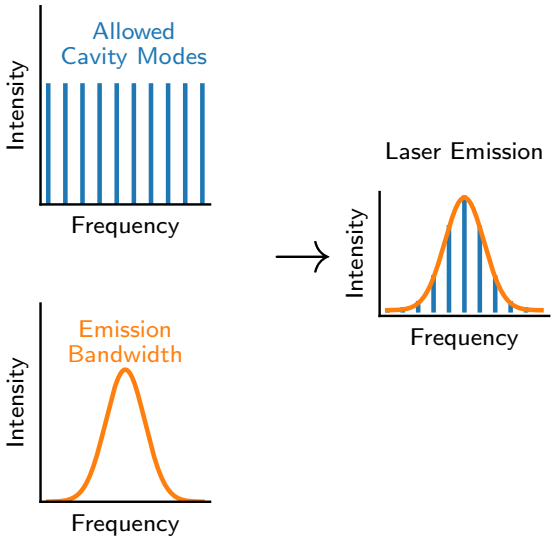


Figure 1.4: Illustration of how the interplay between the modes allowed by the cavity (top left, blue) and the emission bandwidth of the gain medium (bottom left, orange), give rise to the overall laser emission spectrum (right). Note that in most real cases the individual cavity frequencies shown in blue would be much more closely spaced.

time-bandwidth product. For a Gaussian pulse, it is found that:

$$\Delta\nu\Delta\tau \geq 0.441 \tag{1.3}$$

This means that for a Gaussian pulse with a given bandwidth $\Delta\nu$, the shortest possible pulse is found when this inequality is an equality, and then $\Delta\tau = {}^{0.441}/_{\Delta\nu}$. An important point to note here is that Equation 1.3 originates ultimately from the properties of the Fourier transform, even though the physical origin of emission broadening comes from the Uncertainty Principle. Fourier theory predates quantum mechanics, and arguably Fourier was describing the same situation as Heisenberg, but was unable to provide a physical explanation. Heisenberg's Uncertainty Principle provides a physical basis for these properties of the Fourier transform.

Application 1.2: Transform Limits

Taking a typical bandwidth of a Ti:Sa laser of 60 nm at a central wavelength of 800 nm, to use Equation 1.3 we first need to convert the bandwidth into Hz. Using the relationship:

$$\Delta\nu \approx \frac{c\Delta\lambda}{\lambda^2} \tag{1.4}$$

We find that $\Delta\nu$ is 28 THz. Using Equation 1.3 then shows that then the shortest possible pulse this bandwidth could support is theoretically around 15 fs. When the temporal width of the pulse is as short as the bandwidth will allow, we say that the pulse is **transform-limited**, **Fourier-limited**, or that it is **at the transform limit**.

Envelopes and Carriers

Before we continue, at this point it is useful to understand the distinction between the **carrier wave** and the **pulse envelope**. Figure 1.5 shows a hypothetical laser pulse with the pulse envelope (blue) and carrier wave (orange) highlighted. The **pulse envelope** is the

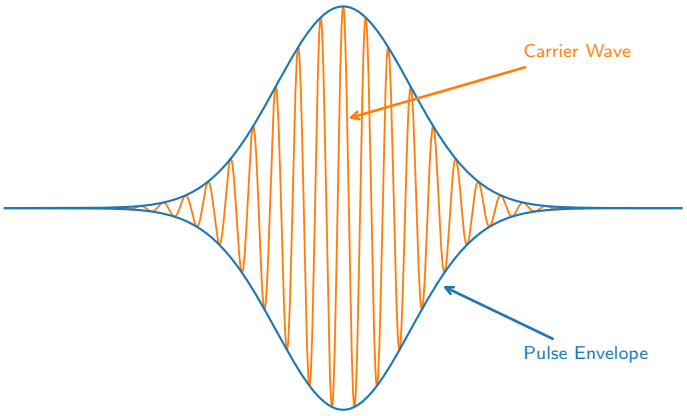


Figure 1.5: A hypothetical transform limited pulse shown with the carrier wave (orange) and pulse envelope (blue).

(in our case) Gaussian distribution that dictates the temporal shape of the pulse - shown in blue on the figure. When we talk about the pulse duration $\Delta\tau$, we talk about the width of this Gaussian distribution in time. The **carrier wave** is the superposition of all the different frequency components in the pulse bandwidth $\Delta\omega$. The product of the carrier wave and pulse envelope represents the electric field of the laser pulse.

In most practical cases, the carrier oscillates at the frequency of the light, which is extremely fast (typically with a frequency on the order of 100 THz for visible light). The envelope generally varies much more slowly than this - so treating them separately in any mathematical analysis is generally a reasonable assumption. This is called the *slowly-varying envelope approximation*.

1.3.2 The Transform Limit

Our time-bandwidth product (Equation 1.3) shows that a laser pulse with a broader bandwidth can have a shorter temporal width than one with a narrower bandwidth. We normally talk about the shortest pulse possible with a given bandwidth as being the shortest pulse that that bandwidth can **support**, and this shortest pulse occurs when the pulse is **transform limited**[6]. A clear question then, is: *'are pulses with a broad bandwidth necessarily short?'*. The answer is **no**, and we will spend a lot of the first part of this book exploring why not.

[6] i.e. the pulse duration is as short as it can be due to fundamental properties of the Fourier transform.

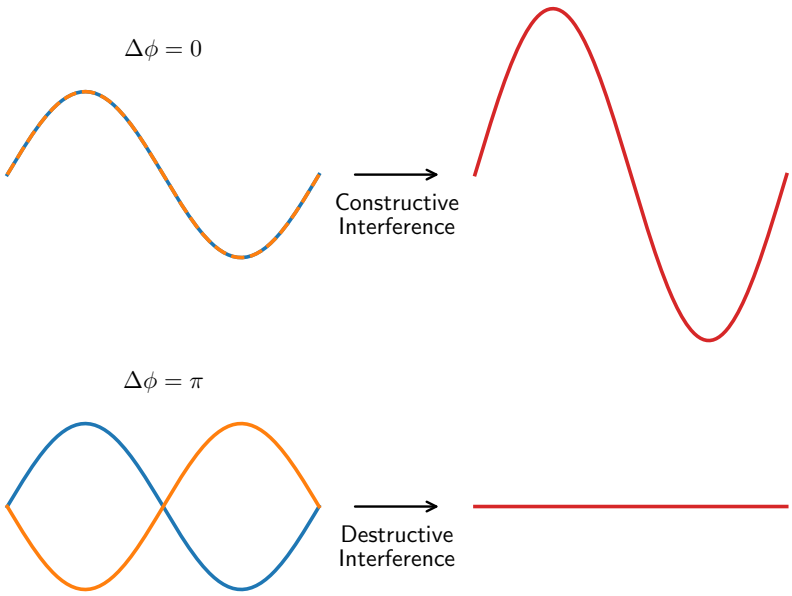


Figure 1.6: Effects of different phase relationships on the interference of two waves (orange and blue), with zero phase difference (top) and with a phase difference π radians (bottom). The resultant wave is shown in red in both cases.

To understand the concept of transform-limited pulses in more detail, it is useful to think about the pulse in the frequency domain. The pulse has a bandwidth $\Delta\omega$ which is spread around a central frequency ω_0. Within this bandwidth there are many different **spectral components** – individual frequencies within the pulse bandwidth. Each of these components can be considered as a sinusoidal wave with electric field $E(t)$:

$$E(t) = E_0 \sin(\omega t + \phi(\omega))$$

Where the frequencies ω of the waves are distributed around ω_0 in accordance with the bandwidth $\Delta\omega$ (larger bandwidth $\rightarrow$ larger spread of frequencies). Here $\phi(\omega)$ is the **phase** of the wave with frequency ω, and E_0 is the amplitude of the electric field. The quantity $\phi(\omega)$ is the **spectral phase**, as it is the phase of the spectral component with frequency ω. The phase of a wave tells you what part of the cycle of the wave you are in, and is discussed further in Appendix A if this is unfamiliar.

It is the *phase relationship* between different frequency components (that is, the difference in their phases $\Delta\phi$) that is critical to understanding ultrafast pulse generation. A simple example that illustrates why phase is important is shown in Figure 1.6 - in which two waves are interfered. In the top frame, both waves are perfectly in phase and have a phase difference $\Delta\phi$ of zero, leading to constructive interference when the amplitudes are summed. In the lower frame, both waves are out of phase by π radians, which leads to destructive interference.

In a typical Ti:Sa laser oscillator, we may have around 250000 colours (waves of different

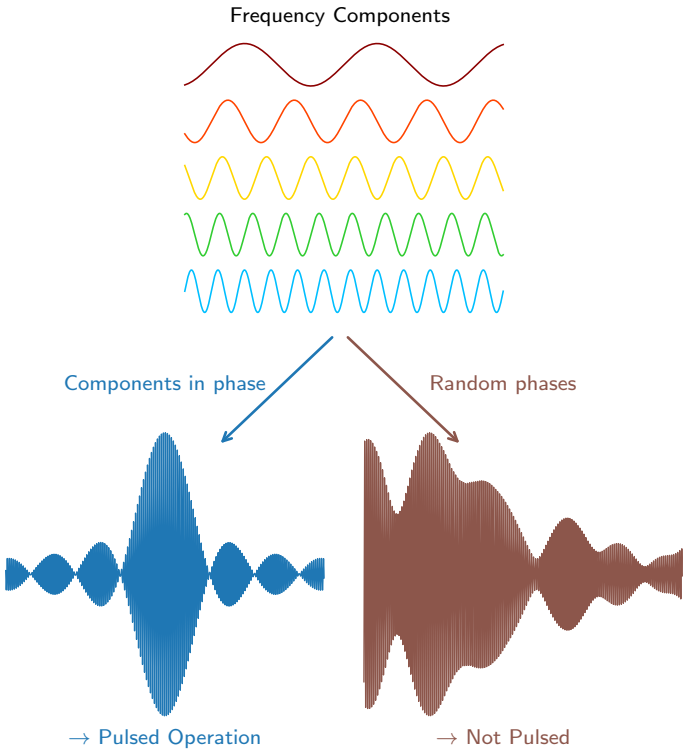


Figure 1.7: Adding waves of different frequencies with a well defined phase relationship leads to generation of a pulse, whereas adding waves with random phases does not. The output waveforms are calculated by superimposing ten waves from within a typical 800 nm Ti:Sa bandwidth, whereas the initial plane waves are chosen just for illustrative reasons.

frequencies) all resonating around the cavity. If each wave has a well defined phase relationship to all the other waves, then the waves are said to be **phase-locked**, and the cavity is said to be **mode-locked**. This results in constructive interference and creation of ultrashort laser pulses. Conversely, if all the phases are random, and there is no well defined phase relationship between one wave and the others, then on average there is destructive interference - and no pulse is seen. This is illustrated in Figure 1.7.

Creating the mode-locked cavity is key to generating ultrashort pulses, and we will discuss how to do this in chapter 4, but the broader point to take from this is that the more modes we have that are phase-locked, the larger our bandwidth $\Delta\omega$, and the shorter our output pulse can be. When all of the frequency components have a well defined phase relationship[7], then they will temporally coincide at a certain point; and this situation is the transform limit. In short: **if all of the colours in the pulse arrive at the same time, the pulse is transform limited.** This concept will be explored more mathematically in the next chapter, but it is illustrative to consider the effect of adding a greater number of frequencies together on the output pulse. This is shown in Figure 1.8, and it is clear

[7] This is equivalent to saying that the $\phi(\omega) = c_n\omega$, where c_n is a constant specific to each frequency. Then $\frac{d\phi}{d\omega} = c_n$, and the phase difference between different components does not depend on the frequency.

that adding more frequencies together in phase creates a shorter output pulse. In fact, the only reason that we see any pulsed output at all is because we add together many different frequencies in phase.

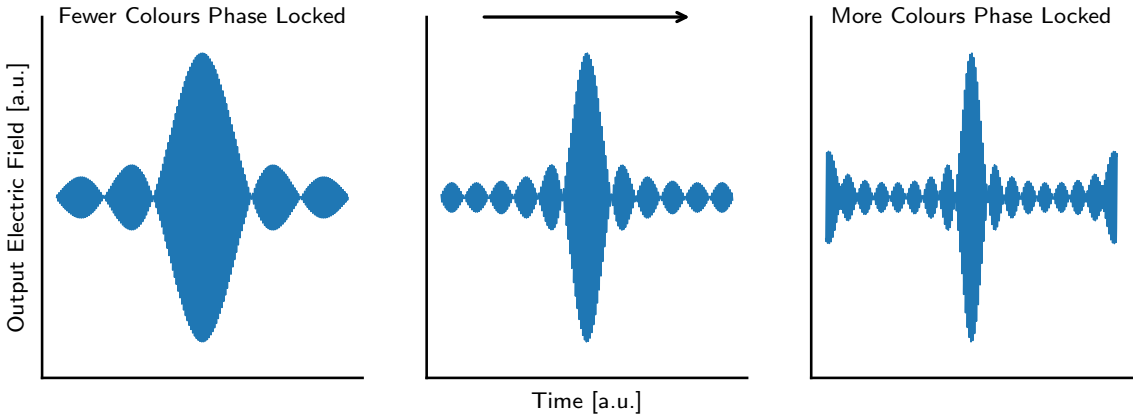


Figure 1.8: Adding together fewer frequencies in phase (left) gives a relatively broad output pulse. Adding more frequencies in phase (middle and right) produces a narrower output pulse. This lies at the heart of ultrashort pulse generation.

For the remainder of this chapter, we will discuss properties of laser light and laser beams that are crucial to understand for effective lab working, but are not specific or unique to ultrafast lasers.

1.3.3 Repetition Rates and Laser Power

We have seen that in an ultrafast laser the light exists as short pulses of EM radiation. Any pulsed laser will produce a certain number of pulses every second, and this number is known as the **repetition rate**, or **rep rate**, of the laser system (measured in Hz). Thus, from the laser output we get a train of the output pulses known as a **pulse train** - this was shown in Figure 1.2.

The energy (in Joules) contained within a single pulse of laser radiation is called the **pulse energy** of the laser. The pulse energy, pulse duration, and repetition rate of a laser lead to two important quantities: the **average power** and the **peak power** if the laser. The average power is defined as:

$$\text{Average Power (W)} = \text{Pulse Energy (J)} \times \text{Rep Rate (Hz)}$$

and the peak power is defined as:

$$\text{Peak Power (W)} \approx \frac{\text{Pulse Energy (J)}}{\text{Pulse Duration (s)}}$$

The average power can be thought of as how much energy the laser will deposit into a sample over a long period of time, whereas the peak power is the amount of energy the

Laser (typical)	E_{pulse}	$\Delta\tau$	f_{rep}	P_{av}	P_{peak}
Amplified Ti:Sa	5 mJ	35 fs	1 kHz	5 W	142 GW
Ti:Sa Oscillator	37.5 nJ	100 fs	80 MHz	3 W	375 kW
Yb:KGW (for pulse energy)	2 mJ	200 fs	10 kHz	20 W	10 GW
Yb:KGW (for rep rate)	20 µJ	200 fs	1 MHz	20 W	100 MW
Yb:KGW (high power)	1 mJ	250 fs	120 kHz	120 W	4 GW

Table 1.2: Some typical representative ultrafast lasers and their output characteristics.

laser will deposit into a sample at the peak of a pulse[8]. As is clear from the equations, shorter pulses will increase the peak power, and higher repetition rates will increase the average power. However, in practice all of these things lie in a balance – most commonly encountered ultrafast lasers do not have average powers of above 80 W, and an increased repetition rate comes at the cost of a decreased pulse energy and vice versa. Most tabletop femtosecond systems have a repetition rate between 1 kHz and 1 MHz. Some examples of how these factors interplay for commonly encountered lasers are given in Table 1.2.

It is clear from Table 1.2 that there is generally a trade off between pulse energy and repetition rate, although recent developments in high-power laser architecture are pushing the boundaries of this. In addition, some lasers allow you to manually tune the repetition rate, either with a fixed or variable pulse energy. Table 1.2 also shows the amplified Ti:Sa laser producing shorter pulses, but operating at a much lower repetition rate than, the Yb-doped lasers – this is generally the case at the time of writing. Key takeaways for someone new to this field, however, are:

- High peak power (intensity) needs short pulses and high pulse energy.
- High pulse energy tends to demand lower repetition rates.
- High repetition rates tend to demand lower pulse energy and/or longer pulses.
- Ti:Sa lasers tend to produce shorter pulses 'out of the box' when compared to Yb-doped lasers.

1.3.4 Polarisation

Having discussed the temporal and spectral characteristics of our ultrafast laser output, there is one more characteristic of the light to discuss before discussing the characteristics of the laser *beam* itself. Recall that we think of our laser light is simply an oscillating electric field travelling in a beam with a well defined direction (see Appendix A for a refresher). The direction of travel of the beam is referred to as the **propagation direction**, and the direction of oscillation of the electric field must be orthogonal to the propagation direction. This arises from **Maxwell's Equations**, which are beyond the scope of this work, but an

[8] The exact equation for peak power depends on the pulse shape – in reality it will be slightly lower than predicted by this simple equation.

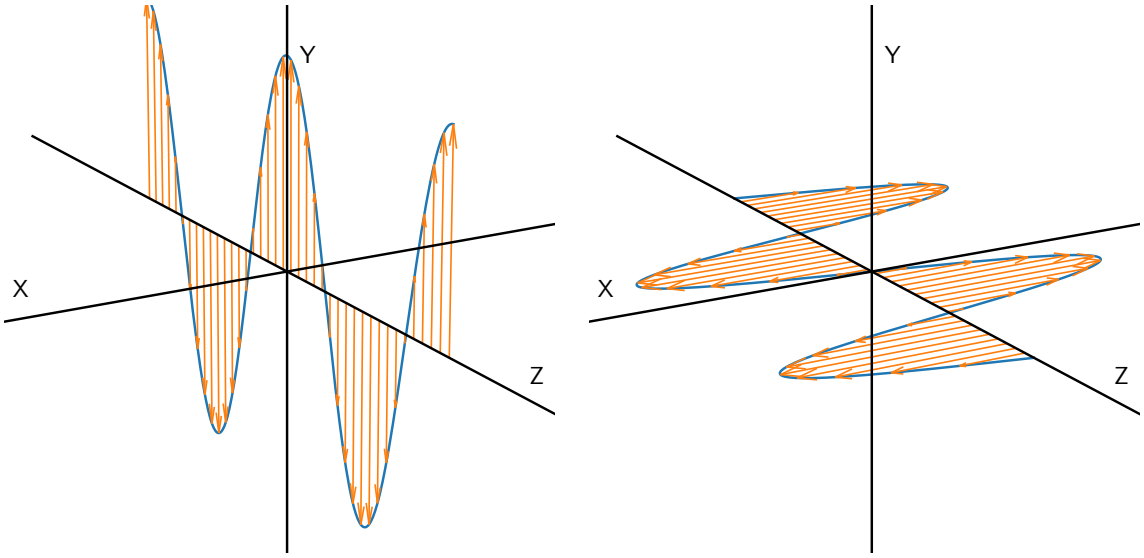
Application 1.3: Reading Power Meters

Laboratory power meters typically measure the **average power** of a laser beam. To convert this number into a useful pulse energy requires that you know the repetition rate of the laser being measured. A laser beam with a power of 50 mW at 1 kHz has a pulse energy of:

$$\frac{50 \times 10^{-3}\,\mathrm{W}}{1 \times 10^{3}\,\mathrm{Hz}} = \frac{50 \times 10^{-3}\,\mathrm{J\,s^{-1}}}{1 \times 10^{3}\,\mathrm{s^{-1}}} = 50 \times 10^{-6}\,\mathrm{J} = 50\,\mu\mathrm{J} \tag{1.5}$$

Some power meters allow the rep rate to be programmed in so that the display is measured in Joules - be aware of this if you have lasers running at different rep rates in your lab!

excellent accessible introduction to Maxwell's equations and vector calculus can be found in [1, 2]. The direction that the electric field oscillates in is known as the *polarisation* of the laser light. Polarisation is a vector that points in the direction of the field oscillation. The plane that is orthogonal to the direction of travel contains the polarisation vector(s) is therefore known as the **polarisation plane**. This is illustrated in Figure 1.9 for a linearly polarised wave.



(a) Linear Polarisation along Y. (b) Linear Polarisation along X

Figure 1.9: Coordinate frames with an illustrative example of polarisation of a light wave (blue). The wave is propagating along the Z axis, and the polarisation plane is the XY plane. The electric field vectors as the wave travels are shown in orange.

Light can be linearly, elliptically, or circularly polarised within the polarisation plane, depending on the exact way in which the field oscillates. More detail on how each of these polarisation states form can be found in any introductory text on optics [3, 4], and is not

included here in detail. Polarisation in ultrafast optics is essentially no different to polarisation in non-ultrafast optics. There are, however, some aspects of polarisation nomenclature which we will use going forward, so are documented below for clarity.

Linear Polarisation

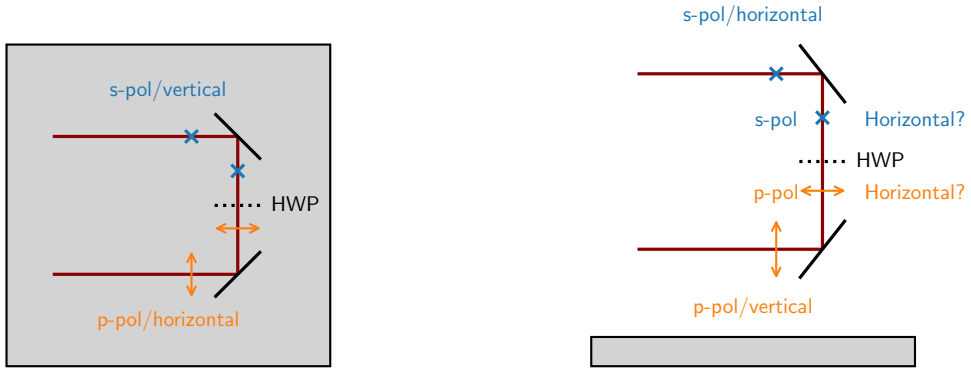
Linearly polarised light is most commonly used, and there are a couple of different systems that are used to refer to the polarisation state of the light. You will hear people talk about 'vertical', 'horizontal', 's-polarised', or 'p-polarised' light.

The vertical/horizontal nomenclature seems at first most intuitive. The direction of the field oscillation is specified with respect to its orientation to the surface of the earth (and optical tables are normally parallel to the surface of the earth!). Then, if the light is polarised such that the direction of oscillation is parallel to the plane of the earth/table, it is said to be **horizontally polarised**. If it is polarised such that the direction of oscillation is perpendicular to the plane of the earth/table, it is said to be **vertically polarised**. These designations are normally very convenient, as when standing in a lab, 'vertical' means 'vertical' with respect to your laser table.

However, this designation becomes more challenging if a beam is not propagating parallel to a table (for example, if you are sending a beam upwards in a periscope). Then the entire polarisation plane is parallel to the table, and so all light would be 'horizontally polarised' by this definition. To remedy this issue, the **s and p polarisation** designations are used.

The s and p polarisation designation defines the polarisation an an incoming wave with respect to a plane known as the *plane of incidence*, which is the plane spanned by the surface normal of an optical element (such as a mirror) and the propagation direction of the incoming wave. More simply, it can be thought of as the plane that the beam travels in both before and after reflection or refraction. If the polarisation direction is parallel to this plane, the light is said to be **p-polarised** ('p' for parallel). Conversely, if the polarisation direction is perpendicular to this plane, the light is said to be **s-polarised** ('s' from the German 'senkrecht', meaning perpendicular). You can see that when a beam is propagating parallel to a laser table, then s-polarised light is vertically polarised and p-polarised light is horizontally polarised. The advantage of the s/p designations is that it makes defining the polarisation when the beam is *not* propagating parallel to the table much less ambiguous. This is illustrated in Figure 1.10.

Figure 1.10 (a) shows the specific case where the beam is travelling parallel to the table. In this case, the plane of incidence is parallel to the laser table, and p-polarised light is always 'horizontally polarised', and s-polarised light is always 'vertically polarised'. Both systems of naming work equally well in this case. However, Figure 1.10 (b) shows the specific case where a beam is propagating perpendicular to the table. At this point where the beam is travelling perpendicular to the table, the plane of incidence is perpendicular to the table, and both s and p polarisation states would be considered 'horizontal' to the table. The horizontal/vertical designation becomes ambiguous in this case. Using the s/p designation removes this ambiguity, so you will normally find that when (for example) optics



(a) Beam parallel to the table. (b) Beam perpendicular to the table.

Figure 1.10: Illustration of different ways to describe polarisation of a given laser beam (red), which is reflected off two mirrors and has its polarisation rotated by a waveplate between the mirrors. The polarisation state is shown as an orange arrow or blue cross, the cross denoting polarisation into the plane of the page.

manufacturers quote polarisation dependent metrics for their optics, they will use the s/p designation.

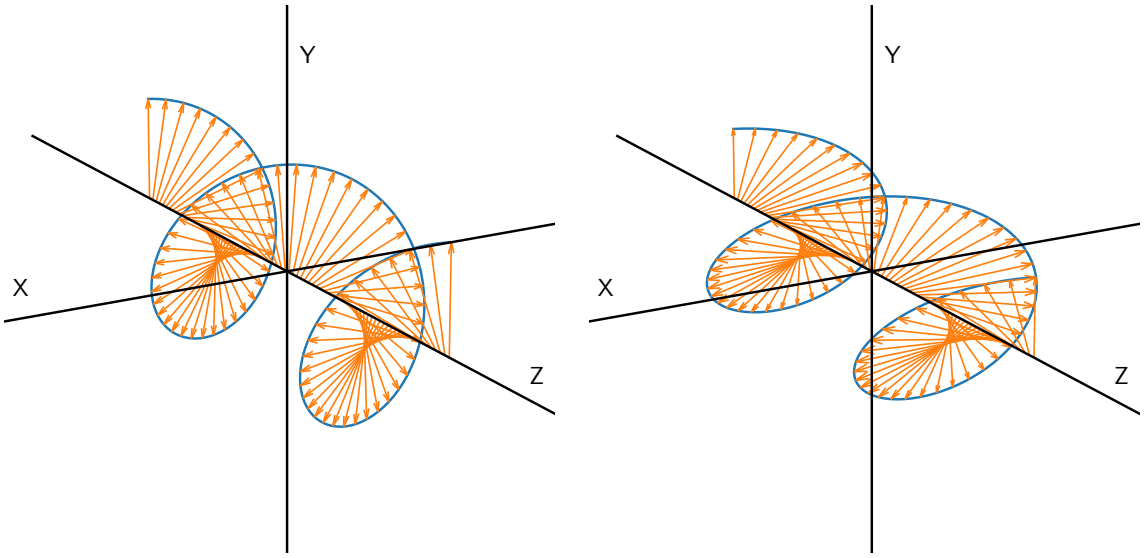
1.3.5 Beyond Linear Polarisation

Most lasers produce linearly polarised light from the output, which can happen because the physical processes that govern the behaviour of the laser oscillator and/or amplifier favour linear polarisation[9]. However, certain applications of ultrafast lasers, such as high-harmonic generation and circular dichroism spectroscopy, make use of circular and elliptical polarisation states.

These polarisation states are where the electric field vector precesses *around* the direction of propagation of the beam rather than oscillating perpendicular to it, as shown in Figure 1.11. Irritatingly, there are multiple conventions for how to define these states (depending on whether you are looking at the beam going away or coming towards you), but they can be either **left-handed** or **right-handed** depending on the direction of the precession – both examples shown in Figure 1.11 are left-handed if you imagine the beam travelling along Z.

The circular polarisation shown in panel (a) of Figure 1.11 is circular because the polarisation state draws a perfect circle if we were to imagine the projection on the polarisation plane. Elliptical polarisation will drawn an ellipse, and linear polarisation will draw a line. Circular and linear polarisations can be thought of as the two limiting cases of elliptical polarisation.

[9] For example, the cavity loss for one polarisation state may be higher than another – see later chapters for details.



(a) Left-Handed Circular Polarisation. (b) Elliptical Polarisation

Figure 1.11: Coordinate frames showing circular and elliptical polarisation. The light wave is shown in blue and the electric field vectors in orange. In both cases the beam is propagating along Z (towards the top of the page) and the polarisation plane is the XY plane.

1.4 Gaussian Beams

We now turn to discuss the nature of the *beam* that the laser light travels in, rather than the nature of the light itself. Almost all laser beams you will come across are well described as **Gaussian beams**. A Gaussian beam is a beam where the amplitude of the beam in the *transverse plane* is described by a Gaussian function. The **transverse plane** here refers to the plane perpendicular to the propagation direction of the beam, i.e. along the 'face' of the beam[10].

This discussion is, like that of polarisation, not specific to ultrafast lasers. However, it is of critical importance for an experimentalist to have some knowledge of Gaussian beams, especially when it comes to characterising laser beams. We will focus on knowledge that helps us use and manipulate Gaussian beams; many standard texts on laser physics will give more information as to *why* and *how* Gaussian beams are produced by laser oscillators [5, 6].

1.4.1 Ideal Beam Parameters

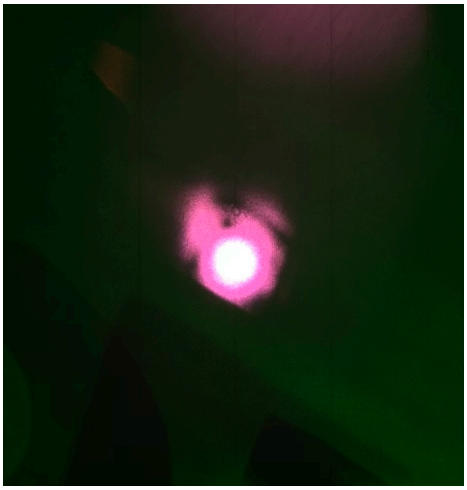
In an ideal world, when we look at the output of our laser beam on a card, we will see a circular spot that trails off towards the edges. Ideally, and in the simplest case, the variation in **intensity** I of this spot as a function of radius r from the center of the spot is given by

[10] The same plane as the polarisation plane.

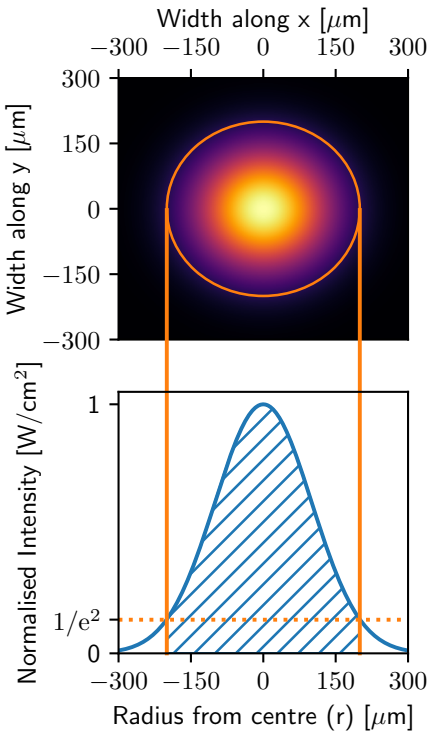
Equation 1.6.

$$I(r) = I_0 \exp\left(-\frac{2r^2}{\rho^2}\right) \tag{1.6}$$

Where I_0 is the **peak intensity**, and ρ is known as the **beam radius**. **Intensity** is a physical quantity that simply refers to the **power per unit area** transferred by something, in this case a laser beam. The beam radius is defined as the distance away from the center of the beam (distance from the axis of propagation of the beam) where the intensity of the beam has fallen to $1/e^2$ of I_0. This is illustrated in Figure 1.12. As Figure 1.12 shows, the vast majority of the beam intensity is enclosed within the $1/e^2$ radius; about 86% of the power in the beam will pass through a circle with a radius ρ.



(a) Typical Ti:Sa laser output.



(b) Simulated laser spot (top) and radial projection (bottom).

Figure 1.12: (a) A photograph of a Ti:Sa laser beam on some paper. (b) Simulation of a circular 2D Gaussian laser spot with a beam radius of 200 µm (top), and the corresponding radial projection (bottom). The orange circle and lines to the projection denote the area enclosed by the $1/e^2$ beam radius. This is also shown by the blue hatched area on the projection.

However, what we have measured by looking at the beam on our card in this way is only the beam radius *at the point where we measured it*. Laser beams will almost always **diverge** or **converge** as they propagate, meaning that the beam radius changes as the beam propagates. The extent to which the beam diverges or converges is called the **divergence** of the beam, and this means the beam radius will *vary*, depending on where in the beam

path we measure it. The beam divergence θ is defined as:

$$\theta = \arctan\left(\frac{\rho_2 - \rho_1}{L_{21}}\right) \tag{1.7}$$

Strictly, θ is the *radial* beam divergence (the 'divergence half-angle'), and sometimes you will see it defined with a capital $\Theta = 2\theta$ (the 'divergence full-angle'). Qualitatively, we can explain Equation 1.7 simply by saying that we measure the beam radius ρ at two positions 1 and 2, and the distance between these positions L_{21}. Straightforward trigonometry then leads to the definition of θ given. These concepts are illustrated in Figure 1.13 in green. The lowest possible divergence for a Gaussian beam (i.e. the divergence of an ideal Gaussian beam) is given by:

$$\theta = \frac{\lambda}{\pi\rho} \tag{1.8}$$

For us, this means that:

- Longer wavelength beams diverge more than shorter wavelength beams.
- Larger beams diverge less than smaller beams.
- All Gaussian beams will diverge to some extent.

Normally a beam taken straight from the output of a good-quality laser (i.e. not a laser pointer or bare laser diode) will have very low divergence, and so the beam radius doesn't vary appreciably except over very large distances. As an example, an ideal Gaussian beam from an typical Ti:Sa or Yb:KGW laser with an initial beam diameter of 10 mm would have to propagate for over 30 m before the beam diameter gets bigger than 11 mm.

However, if we **focus** our beam, then the divergence changes rapidly, and the size of the beam radius will decrease drastically as we measure it nearer and nearer to the focus. This is shown in Figure 1.13 as the orange lines, which are the outline of a Gaussian beam which is focussed at the center of the plot (and diverges away from the center). In the case that the beam diverges or converges, it's not clear how we should define our beam according to Equation 1.6. To account for this variation of the beam radius as the beam propagates, we have to refine Equation 1.6 such that the beam radius $\rho \to \rho(z)$, where z is the distance away from the point in the beam path where the radius is smallest.

$$I(r, z) = I_0 \exp\left(-\frac{2r^2}{\rho(z)^2}\right) \tag{1.9}$$

When the beam radius is at its smallest, it is known as the **beam waist**, and is given the symbol ρ_0. It can also be shown that the variation in beam radius $\rho(z)$ can be expressed as:

$$\rho(z) = \rho_0\sqrt{1 + \left(\frac{z}{z_R}\right)^2} \tag{1.10}$$

Where z_R is known as the **Rayleigh range** or **Rayleigh length**. These concepts are illustrated in Figure 1.13 in blue. At a distance of z_R from the position of the beam waist,

the beam radius has increased by a factor of $\sqrt{2}$; the Rayleigh length therefore gives some measure of the range over which the beam remains small - i.e. if we have focussed the beam, how long is the region where the beam is still effectively 'focussed'?

The beam waist ρ_0, divergence θ, and Rayleigh length z_R are three of the most useful parameters when characterising Gaussian beams, and are readily measured in the lab. In fact, the quality of a Gaussian beam is defined using the **beam parameter product (BPP)**. This is defined as the product of the divergence θ and beam waist ω_0:

$$\text{BPP} = \text{Beam Waist}\,(\rho_0) \times \text{Divergence half-angle}\,(\theta)$$

The units of θ and ρ are normally given in mrad and mm respectively; so the units of the BPP are mm mrad. Comparison with Equation 1.8 readily reveals that the BPP of an ideal Gaussian beam is simply λ/π. A beam with this BPP will have the smallest possible beam waist for a given divergence, and vice versa.

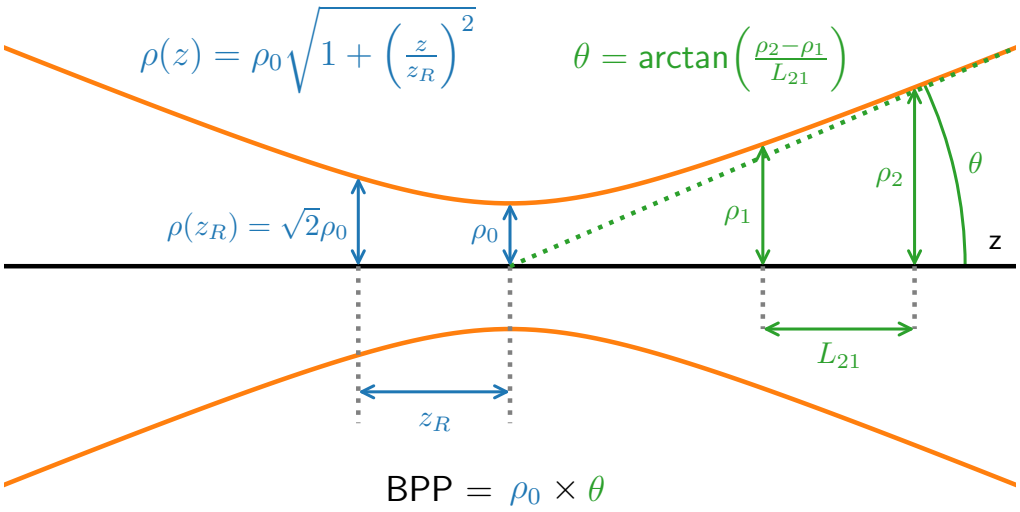


Figure 1.13: Illustration summarising the various parameters used to describe a generic Gaussian beam. The beam profile (orange) is shown side-on. Parameters relating to the beam waist ρ_0 and Rayleigh length z_R are shown in blue. Parameters relating to the divergence angle θ are shown in green.

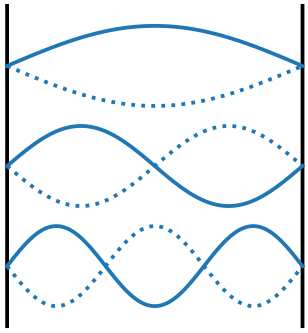
The more practical ramifications of beam waists, divergences, Rayleigh lengths, and BPPs, such as how you can measure them, and how you can use them, in a later section. However, we made reference to the preceding discussions only being valid in 'ideal' conditions - what these conditions are is the subject of the next section.

1.4.2 Deviations from Ideality

Beam Quality Factors - M^2

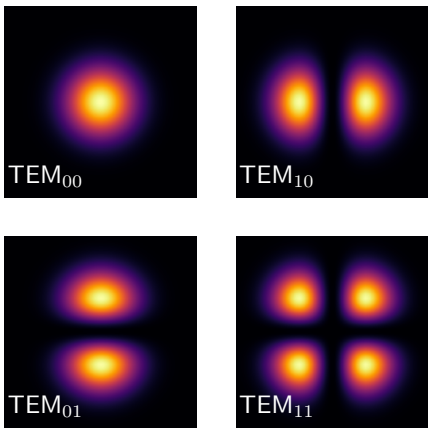
Laser beams are rarely perfectly ideal Gaussian beams, and one way in which they can be non-ideal is in the presence of different **transverse modes** in the beam. A transverse mode

is essentially 'what the front of the laser beam looks like', that is, the shape of the 'face' of the beam. The idea of a *longitudinal mode* in a standing wave may be familiar from school physics, such as when a guitar string oscillates. The difference between longitudinal and transverse modes is shown in Figure 1.14.



Longitudinal Modes

(a)



Transverse Modes

(b)

Figure 1.14: (a) Lowest three *longitudinal modes* of a standing wave in a cavity. (b) Lowest four *transverse modes* of a Gaussian beam.

Figure 1.14 (a) shows longitudinal modes of a standing wave in a cavity. Here the direction of travel of the wave is in the plane of the page (we are looking at the wave 'side-on'). This situation is analogous to when guitar strings vibrate. Figure 1.14 (b) shows transverse modes of a Gaussian beam. Here the direction of travel of the wave is out of the plane of the paper (we are looking at the wave 'head-on'), and a tangible analogy would be that it is like the skin of a drum as it resonates after being hit. The modes are labelled with the 'TEM$_{m,n}$' nomenclature; TEM stands for ´Transverse ElectroMagnetic mode', and the subscripted numbers refer to the degree of the *Hermite polynomial* in the mathematical expression of the mode. The precise mathematical form of the modes need not detain us here, and we will focus on the important point, which is that **ideally, our laser will output light only in the lowest, TEM$_{00}$, transverse mode.**

If our laser is not outputting light in only this lowest mode, and there are contributions from higher order modes present[11], then our beam is no longer an ideal Gaussian beam. This means that our BPP will be greater than the ideal value of λ/π. Intuitively, we could get a qualitative feel for *how* non-ideal our beam is by taking the ratio of the BPP for our beam to the BPP for an ideal Gaussian beam of the same wavelength. This ratio has a name, the **beam quality factor**, or **M^2 ('M-squared')** factor.

$$\text{Beam Quality Factor } (M^2) = \frac{\text{BPP of real beam}}{\text{BPP of ideal Gaussian beam}}$$

[11] This can occur due to poor oscillator alignment, for example.

As we are only considering Gaussian beams, we know that the lowest possible BPP is for that of an ideal TEM_{00} Gaussian beam. This means that $M^2 \geq 1$. If $M^2 = 1$, then the beam is ideal. You may also hear people talk about a beam being **diffraction-limited**[12] this is equivalent to saying that $M^2 = 1$. In practice, provided that $M^2 < 1.3$, the beam can be considered as behaving reasonably ideally.

We will discuss the practical implications of this, and how you can measure the M^2 value, in chapter 5. For now, just being aware of the terminology will be useful as you start to work in the lab and think about your beams.

Elliptical Beams

Another area where our assumptions of ideal behaviour in subsection 1.4.1 break down is if our beam spot is **elliptical**, that is, is not circular. In this instance, we cannot define the beam radius as we had previously - as the radius will vary depending on where on the ellipse we look at the beam! In this instance, the normal thing to do is to quote the radius at the widest and narrowest parts of the ellipse[13]. Then, rather than quoting the beam radius/waist as a single number, it would be quoted as two numbers, for example '25 µm by 35 µm'.

Sometimes elliptical beams are formed deliberately using cylindrical lenses, as they can be advantageous in some spectroscopy experiments. Often they are undesirable, but form anyway due to imperfections in optics in a beam path, or imperfections in the laser output. Dealing with elliptical beams can be difficult, and often they form when an optic is over-tightened in a mount (and so deforms slightly producing a lensing effect), or when a focussing optic isn't placed exactly square in the beam path. Checking that all your telescopes are properly aligned, and that no thin optics are over-tightened can remedy the situation. Sometimes a beam is only very slightly elliptical, such that just taking the mean of the two beam radii gives a reasonable number for an 'effective' circular beam radius that can be used in subsequent calculations.

We have now summarised all of the useful properties of laser light and laser beams that we will need going forward. In the next chapter, we will briefly explore some fundamental points about ultrafast laser *architecture* that will be useful as we continue our discussions.

Bibliography

[1] Fleisch D. *A Student's Guide to Maxwell's Equations*. Cambridge University Press, Cambridge, 1st edition, 2008.

[12] This terminology arises because at the diffraction limit, the beam waist cannot be lowered any further. Attempting to reduce the beam waist below the diffraction limit will lead to broadening of the beam waist due to diffraction.

[13] The semi-major and semi-minor axes of the ellipse

[2] Schey H M. *Div, Grad, Curl, and All that: An Informal Text on Vector Calculus*. W. W. Norton, New York, 4th edition, 2005.

[3] Pedrotti F L, Pedrotti L M, and Pedrotti L S. *Introduction to Optics*. Cambridge University Press, Cambridge, 3rd edition, 2017.

[4] Goldstein D H. *Polarized Light*. CRC Press, Boca Raton, FL, 3rd edition, 2017.

[5] Milonni P W and Eberly J H. *Laser Physics*. Wiley, New York, 1st edition, 2010.

[6] Hooker S and Webb C. *Laser Physics*. Oxford University Press, Oxford, 1st edition, 2010.

Chapter 2

Bandwidth and Dispersion

In this chapter a more detailed description of the interplay of bandwidth and pulse duration is given, together with accessible descriptions of how dispersion can lead to pulse broadening. Some initial discussion of pulse compression is given, together with implications that a broad pulse bandwidth has on other properties of the laser light.

In the previous chapter we described some of the important properties of laser light, but didn't really fully answer the question we set out to answer, which was *'are pulses with a broad bandwidth necessarily short?'.* We said that to have a transform limited pulse (the shortest possible pulse for a given bandwidth), there had to be a fixed phase relationship between all the frequency components (colours) in the pulse, such that at some point all of the different components temporally coincide to produce a short pulse.

Dispersion is the phenomenon which stops this happening. Dispersion is a phenomenon which is critical in ultrafast optics. If you are familiar with non-ultrafast optics then you may have never really thought about it, and it is the single biggest difference that we have to be mindful of when moving from non-ultrafast optics to ultrafast optics. Dispersion, unless properly taken care of, can cause ultrafast pulses to no longer resemble anything close to ultrafast pulses - so it is a very important concept to understand.

An initial disclaimer is that when we refer to 'dispersion' we are strictly referring to **chromatic dispersion** or **temporal dispersion** - the spreading out of different colours *in time*. There are other kinds of dispersion too, such as **spatial dispersion**, which will we mention later. Whenever we refer to an unqualified 'dispersion', we are referring to chromatic dispersion.

2.1 Origins of Dispersion

The **refractive index**, n, of a material is defined as:

$$n = \frac{c_0}{c} \tag{2.1}$$

Where c_0 is the speed of light in vacuum, and c is the speed of light in the material. In all normal materials, $c < c_0$ and therefore $n > 1$. Moreover, the refractive index of a material is frequency dependent:

$$n = n(\omega) \tag{2.2}$$

With the result that different colours of light travel through the material at different speeds. **Equation 2.2 is the physical origin of dispersion.** If different colours of light experience different refractive indices, then (for example) the higher energy 'bluer' colours could travel more slowly through the material than the lower energy 'redder' colours. This situation in which higher energy light travels more slowly is called **normal dispersion**. The opposite situation, when higher energy light travels faster, is called **anomalous dispersion**. These two cases are summarised in Table 2.1.

Dispersion	Refractive Index	Blue Light	Red Light
Normal	Increases with ω	Slows More	Slows Less
Anomalous	Decreases with ω	Slows Less	Slows More

Table 2.1: Summary of different types of dispersion

The wavelength (energy) dependence of the refractive index for a material can be calculated easily using the *Sellmeier Equation* for that material, but for practical use it is much more convenient to use a website like https://refractiveindex.info, which contains wavelength-dependent refractive index data for a wide range of common optical materials. Figure 2.1 shows the refractive indices of some common optical materials, plotted against *wavelength* (λ – refractive index data is almost always given in wavelength units, so you need to be comfortable converting these plots to photon energy or frequency in your head).

Most materials exhibit normal dispersion (refractive index increases with ω, or decreases with λ), but some show anomalous dispersion in narrow regions[1]. The left hand panel of Figure 2.1 shows the variation in refractive index over the whole transmission spectrum of the material (truncations are where the material is not transparent). The right hand panel shows the variation in refractive index over the most commonly used region in ultrafast optics (from the UV to the near IR). You can see that as we approach the UV region (400 nm and lower) the refractive indices tend to change more sharply.

Rearranging Equation 2.1, we find that in general the speed at which a given frequency component travels through a medium is given by:

$$c(\omega) = \frac{c_0}{n(\omega)} \tag{2.3}$$

The speed of a particular frequency component in a pulse is known as the **phase velocity** of the component. This is distinct from the **group velocity**, which is the velocity of the

[1] Note that some sources will talk about 'anomalous dispersion' to mean 'negative group delay dispersion (GDD)', so it is important to be aware of the context. Many materials have large regions where they exhibit negative GDD, but this does not necessarily mean that the refractive index of the material decreases with frequency.

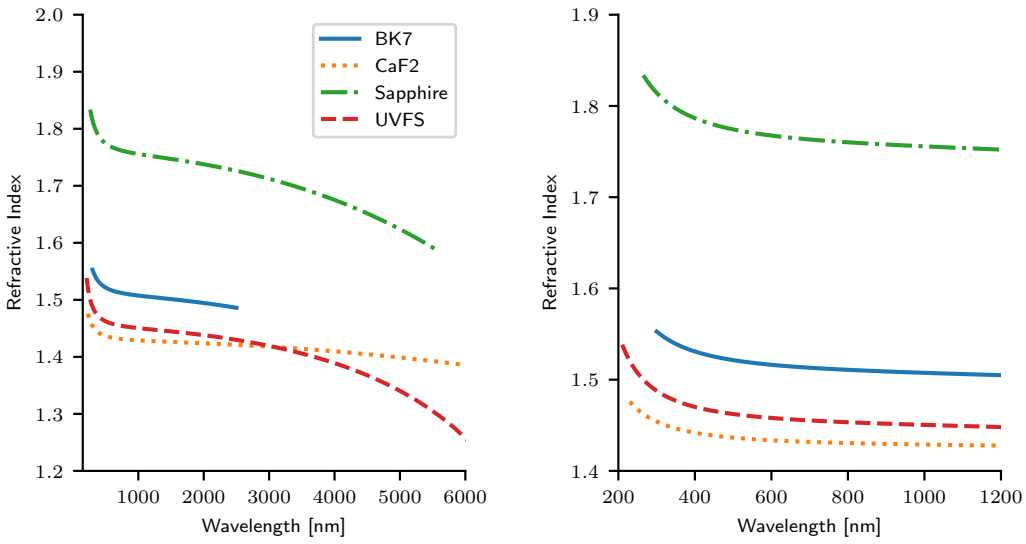


Figure 2.1: Refractive indices of some common optical materials, as a function of wavelength. The right hand plot is truncated to cover only the UV-Vis and near-IR regions, to illustrate how refractive index changes in this commonly used region. Materials chosen are BK7 (blue); Calcium Fluoride (orange); UV-fused silica (red); and Sapphire (green). Note that not all materials have data available for all wavelengths of light due to different transparencies, hence some curves are longer/shorter than others.

entire pulse, as the different frequencies that make up the pulse travel at different speeds through a given medium, according to Equation 2.3. It is important to note that because $n > 1$, all colours of light will slow down on entering a medium – it is the relative amounts by which different colours slow down that is important.

2.2 Dispersion, Ultrafast Pulses, and Chirp

We have seen that the refractive index is frequency dependent, and that this can cause dispersion as different colours of light travel at different speeds through a given medium. For non-ultrafast laser pulses, this is not usually problematic - as the bandwidth is sufficiently narrow that the effects of dispersion simply cause the entire laser pulse to slow down by uniform amount. When working with ultrashort pulses though, dispersion becomes a critical consideration because an ultrashort pulse **inherently has a broad bandwidth.** This is an important subtlety to understand. *It is not directly the short temporal width that causes ultrashort pulses to be affected more by dispersion.* Rather, it is that *broadband* pulses are affected more by dispersion, and ultrashort pulses are necessarily broadband.

To qualitatively understand the effect that dispersion can have on an ultrafast pulse, imagine we have a hypothetical transform limited pulse with a very broad bandwidth, containing frequency components from dark red to deep blue. The pulse is transform limited, so all of the different frequency components arrive at the same time. If, however, we pass this pulse through a piece of glass with normal dispersion, the following will happen:

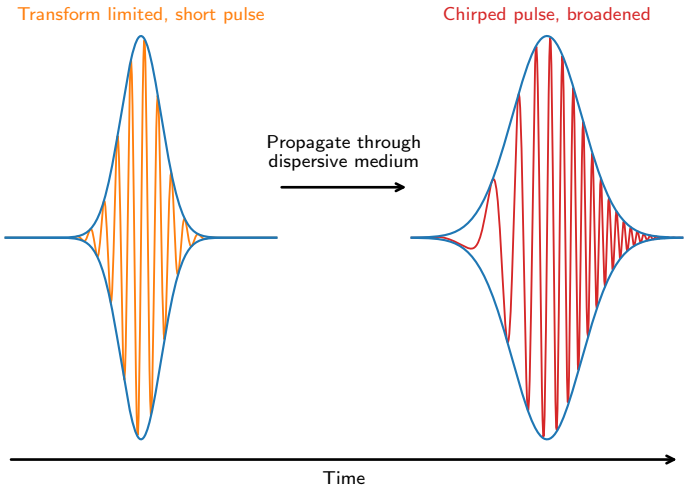


Figure 2.2: Schematic of dispersion of a TL pulse (orange) passing through a dispersive medium with normal dispersion, so it becomes chirped and broadened (red). The extent of the chirp has been exaggerated for illustrative reasons. Note that longer wavelengths occur at earlier times in the broadened pulse, as they are not slowed by the material as much.

- Both the red and blue frequencies slow down, but the red frequency components are slowed down less than the blue (as they experience a lower refractive index: normal dispersion).
- So, the red frequencies travel faster through the material than the blue frequencies. After propagation, there would be more red components at the front of the pulse (earlier times) than at the back (later times).
- Therefore, the colours are no longer arriving at the same time. **The pulse is no longer transform limited.**

The different frequency components have been **dispersed** in time by the propagation through the material, and therefore the pulse **must** have broadened. This effect is more dramatic for pulses with wider bandwidths, and for materials with a higher dispersion.

Following the passage through a medium with positive dispersion, the low-frequency red components arrive before the high-frequency blue components - this is the case plotted in Figure 2.2[2]. At this point, we say that the pulse is **chirped**. The reason for this name becomes clear if we look at a spectrogram (plot of frequency against time) for the pulse, as in Figure 2.3.

Clearly for the two chirped pulses show in Figure 2.3 (orange dotted and green dashed lines), the frequency of the pulse changes during the course of the pulse. For a **positively chirped**

[2]I appreciate that these time axes can be confusing at first. The figure looks like the high-frequency waves are coming 'first' at the end of the chirped pulse, but actually as time increases from left to right, they are coming *at later times*, so are coming later.

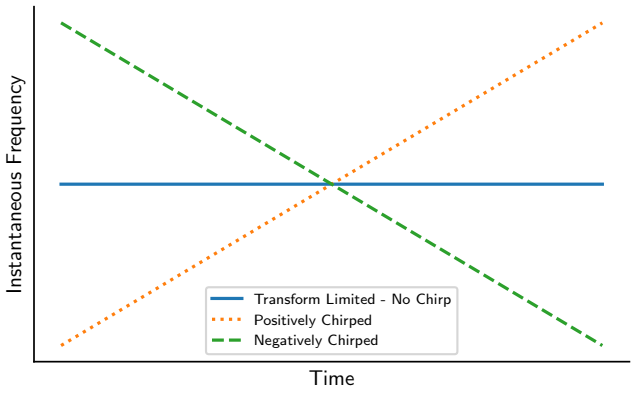


Figure 2.3: Spectrograms for an arbitrary unchirped pulse (blue solid), positively chirped pulse (orange dotted), and negatively chirped pulse (green dashed).

pulse (orange dotted line), the frequency **increases** during the course of the pulse, because the lower frequencies arrive earlier. For a **negatively chirped** pulse (green dashed line), the frequency **decreases** during the course of the pulse, because the higher frequencies arrive earlier. The name 'chirp' is given, as if you were to downconvert the optical frequencies into audio frequencies, you would hear a rising or falling frequency - rather like a bird chirping. Finally, for the **unchirped** pulse (blue solid line), the frequency does not change during the course of the pulse - all of the frequencies arrive at the same time.

The next obvious question is *'how much broader will my pulses get if I propagate them through a dispersive medium?'*. The answer to this requires that we calculate how *much* dispersion the pulse experiences. This, in turn, requires that we think more mathematically about what happens when light propagates through a dispersive medium.

2.3 Propagation Through a Dispersive Medium

We will now think more mathematically about the propagation of light through a dispersive medium, to put our prior (and later) discussion on a more formal footing. To avoid dense mathematics getting in the way of an intuitive understanding, this is kept rather brief and some mathematical complexity is omitted. More rigorous treatments of this topic can be found in references [1, 2, 3].

As our initial discussions have suggested, an ultrafast laser pulse can be described as a superposition of many different frequency components, which all interfere in a way that produces pulsed output (as shown in Figure 1.7). We can therefore write the electric field of our pulse in the time domain $E(t)$ as a sum of different frequencies ω:

$$E(t) \propto \sum \mathcal{E}(\omega) e^{i\omega t} \tag{2.4}$$

Where the wave with frequency ω is $e^{i\omega t}$, and the amplitude factor $\mathcal{E}(\omega)$ determines how much of that frequency is present in the pulse. Given that the range of frequencies spanned

by our pulse is continuous[3], we can write the sum as an integral:

$$E(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathcal{E}(\omega) e^{i\omega t} \mathrm{d}\omega \tag{2.5}$$

In Equation 2.5, $\mathcal{E}(\omega)$ is really just the laser pulse $E(t)$ in the frequency domain. The mathematical machinery that connects these two descriptions is the **Fourier transform**, which we can leverage by writing the sum as an integral[4]. There are many excellent resources available if you are interested in the mathematical background of this, such as those in references [4, 5]. We choose to write our pulse like this, because as the pulse propagates through a medium, each frequency in the pulse will be affected differently due to the frequency dependent refractive index $n(\omega)$ - so it is much simpler to look at our pulse decomposed into individual frequency components as in Equation 2.5. Herein lies the power of the Fourier transform. Problems that are seemingly intractable in the time domain can be rendered trivial in the frequency domain and vice versa. You will encounter this methodology frequently in the physical sciences, so being comfortable thinking in both the frequency and time domain is a crucial skill.

As the pulse travels through a material, each individual frequency will accumulate a phase, $\phi(\omega)$ (see Appendix A if this is unfamiliar). The quantity $\phi(\omega)$ is known as the **spectral phase** - that is, the **phase** of each part of the **spectrum** of colours in our pulse. Mathematically, adding the accumulated phase to our wave $e^{i\omega t}$ amounts to simply multiplying it by a phase factor, $e^{i\phi(\omega)}$, such that:

$$e^{i\omega t} \xrightarrow{\text{propagation}} e^{i\omega t} e^{i\phi(\omega)} \tag{2.6}$$

The output pulse after propagation, $E_{\text{out}}(t)$, can then be written as:

$$E_{\text{out}}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathcal{E}(\omega) e^{i(\omega t + \phi(\omega))} \mathrm{d}\omega \tag{2.7}$$

Thus, if we can find the accumulated spectral phase $\phi(\omega)$, then we can calculate the effect that propagation through the dispersive medium has on our pulse. One way to illustrate this effect is to expand $\phi(\omega)$ as a Taylor series around the central frequency ω_0.

$$\phi(\omega) = \phi(\omega_0) + \frac{\partial \phi}{\partial \omega}\bigg|_{\omega=\omega_0} (\omega - \omega_0) + \frac{1}{2} \frac{\partial^2 \phi}{\partial \omega^2}\bigg|_{\omega=\omega_0} (\omega - \omega_0)^2 + \ldots \tag{2.8}$$

Higher order terms than this can normally be neglected, and we will now discuss each of these terms in turn and gain an understanding of their physical meaning, following notation convention used in Hooker and Webb [1].

Zeroth order term:

$\phi(\omega_0) \equiv \phi^{(0)}$. This is the total phase that is accumulated at the central wavelength. This is responsible for the **carrier-envelope phase** (CEP) which is the phase difference between

[3] At least, the spacing between two components is very small.

[4] The additional factor of $1/\sqrt{2\pi}$ comes from the relationship between the temporal and angular frequency.

the carrier wave and the pulse envelope. The CEP is generally only significant for few-cycle pulses, where each pulse only consists of very few optical cycles. This will not be an important factor for the laser pulses we will consider.

First order term:

$\frac{\partial\phi}{\partial\omega} \equiv \phi^{(1)}$. This term is called the **group delay**, with units of time, and gives the time taken for the pulse to propagate through the dispersive medium. If the dispersive medium has a length L, we can also intuitively define the **group velocity**, $v_g = {}^{L}/_{\phi^{(1)}}$. The group velocity defines the *velocity at which the pulse envelope (i.e. the whole pulse) travels*. This is distinct from the **phase velocity**, which is specific to each frequency (each carrier wave) and defines the velocity of the individual carrier waves. Changing the group delay has the effect of changing the temporal position of the pulse, but does not change the shape of the pulse envelope, so does not lead to broadening. This can be seen clearly via a mathematical argument. If we take the group delay to be non-zero, it implies:

$$\frac{\partial\phi}{\partial\omega} = a \tag{2.9}$$

where a is a non-zero constant. The equation above implies that a change in the frequency $\partial\omega$ results in a change in the phase $\partial\phi$ that is equal to a, and thus is not dependent on frequency. Thus, the phase relationships between all colours in the pulse are affected equally by this term, and the overall phase relationship between each pair of colours is preserved.

Second order term:

$\frac{\partial^2\phi}{\partial\omega} \equiv \phi^{(2)}$. This term is called the **group delay dispersion (GDD)**, and has units of $(\text{time})^2$. It is the lowest-order term responsible for broadening of pulses via dispersion[5]. It is also generally the main source of dispersion that will broaden our pulses. When this GDD term is non-zero it implies that the phase difference between components depends on frequency, i.e:

$$\frac{\partial^2\phi}{\partial\omega^2} = a \rightarrow \frac{\partial\phi}{\partial\omega} = a\omega + a_0$$

where a and a_0 are arbitrary constants. Thus, for a given change in frequency $\partial\omega$ there is a change in the phase $\partial\phi$ that depends on frequency, ω. In this case it is no longer possible to achieve the fixed phase relation between all colours in our pulse that we would need in order to have all of our frequencies arrive simultaneously and make a transform-limited pulse. **GDD is the main phenomenon that will cause your pulses to broaden.**. The effect of different orders of the spectral phase on a pulse is shown in Figure 2.4.

Higher order terms can also play a role, especially **third order dispersion (TOD)**, but outside of certain specialist areas these tend to not be significant (and they're much harder to compensate for!). The above discussion was rather qualitative and a more quantitative description is provided in subsection 2.3.1 for more mathematically-inclined readers. However, if you are happy enough with the qualitative description then the following section can be skipped.

[5] One can also calculate the **group velocity dispersion** as ${}^{\text{GDD}}/_{L}$ - the GDD per unit length.

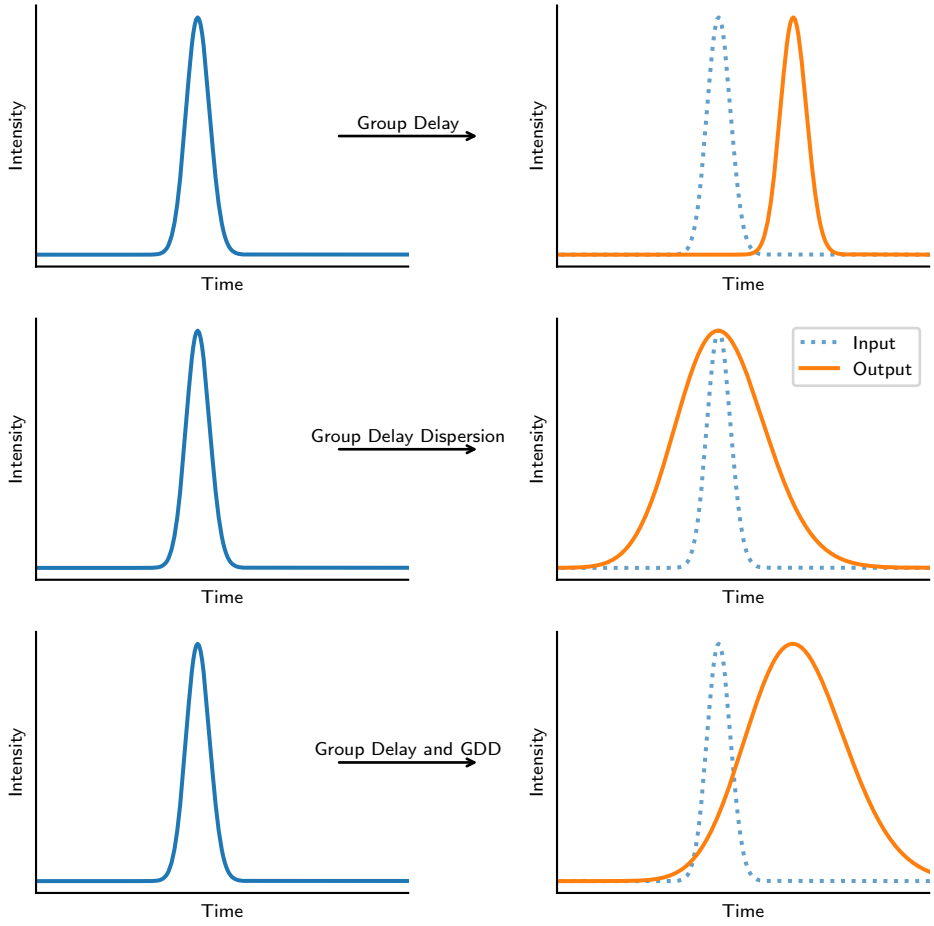


Figure 2.4: Effect of different orders of the spectral phase on ultrashort pulses. Pulses before propagation are shown on the LHS (blue solid), pulses after on the RHS (orange solid, with the original pulse in dotted blue). Propagation was numerically modelled using the equations in subsection 2.3.1 and a Gaussian envelope in the frequency domain. From top to bottom, propagation with: group delay only, GDD only, and combined group delay and GDD. The effect of the different terms on the output pulse is clearly visible.

2.3.1 The Role of the Spectral Phase

It is perhaps not immediately clear why the GDD causes pulse broadening but the group delay does not. We can understand this mathematically by simply plugging Equation 2.8 into the exponential term in Equation 2.7. Before we do this we will rewrite our equations in terms of a factor $\Delta\omega = \omega - \omega_0$, as this will facilitate the later discussion. We write Equation 2.8 as:

$$\phi(\omega) = \phi^{(0)} + \phi^{(1)}\Delta\omega + \frac{1}{2}\phi^{(2)}\Delta\omega^2 \tag{2.10}$$

Neglecting higher order terms and using the contracted notation for the partial derivatives introduced above. The factor $\Delta\omega$ can be simply thought of as a kind of 'detuning parameter' - the larger it is, the further we are from ω_0, and the larger the bandwidth. We can then substitute this into the exponential term of Equation 2.7 to illustrate the effect of each of

the phase terms.

$$\exp[i(\omega t + \phi(\omega))] = \exp[i((\omega_0 + \Delta\omega)t + \phi^{(0)} + \phi^{(1)}\Delta\omega + \frac{1}{2}\phi^{(2)}\Delta\omega^2)] \tag{2.11}$$

Where we have also made the substitution $\omega = \omega_0 + \Delta\omega$. Equation 2.11 can be further separated by factoring out the term containing ω_0 as the phase terms are all in $\Delta\omega$. We are then left with the following expression for the propagated waves:

$$\exp[i(\omega t + \phi(\omega))] = \exp[i\omega_0 t]\exp[i(\Delta\omega t + \phi^{(0)} + \phi^{(1)}\Delta\omega + \frac{1}{2}\phi^{(2)}\Delta\omega^2)] \tag{2.12}$$

To illustrate the effect that each of the phase terms has on the pulse, we will consider them in turn and let the other two phase terms be equal to zero to isolate the effect of each individual term, as in the previous section.

Zeroth order term:

Letting $\phi^{(1)} = \phi^{(2)} = 0$ leads to the following expression for the waves after propagation:

$$\exp[i(\omega t + \phi(\omega))] = \exp[i\omega_0 t]\exp[i\Delta\omega t]\exp[i\phi^{(0)}] \tag{2.13}$$

We can clearly see from Equation 2.13 that the effect of $\phi^{(0)}$ is to produce a **phase shift** of $\phi^{(0)}$ to the entire pulse. This phase shift is the CEP discussed earlier. This term does not lead to broadening or movement in time.

First order term:

Now letting $\phi^{(0)} = \phi^{(2)} = 0$ leads to the following expression for the waves after propagation:

$$\exp[i(\omega t + \phi(\omega))] = \exp[i\omega_0 t]\exp[i\Delta\omega(t + \phi^{(1)})] \tag{2.14}$$

We can see from Equation 2.14 that in essence the effect of $\phi^{(1)}$ (the group delay) is to produce a shift to the time axis of our pulse bandwidth, from t to $t + \phi^{(1)}$. The shift is uniform across the whole pulse bandwidth, and so the whole pulse moves in time by the group delay, as expected.

Second order term:

Finally letting $\phi^{(0)} = \phi^{(1)} = 0$ leads to the following expression for the waves after propagation:

$$\exp[i(\omega t + \phi(\omega))] = \exp[i\omega_0 t]\exp[i\Delta\omega(t + \frac{1}{2}\phi^{(2)}\Delta\omega)] \tag{2.15}$$

We can see from Equation 2.15 that the effect of $\phi^{(2)}$ (the group delay dispersion) is to produce a shift to the time axis of our pulse bandwidth, from t to $t + \phi^{(2)}\Delta\omega$. However, this shift is not uniform across the whole pulse bandwidth, as the shift for each individual frequency depends on $\Delta\omega$ - how far that frequency is from ω_0. The result of this is that each frequency in the pulse is shifted in time differently, with frequencies further from the central frequency (larger $\Delta\omega$) being shifted further. As each frequency is shifted by a

different amount, the pulse broadens, as expected. Moreover, the larger $\Delta\omega$ is, the more severe the broadening - in line with what we expect: broadband pulses undergo more severe broadening than narrowband pulses. This term can also be understood as the rate of change of the group delay with frequency; if the the group delay depends frequency then each part of the pulse experiences a different group delay and the pulse broadens.

2.3.2 Spectral Phase and Refractive Index

We have spoken about the spectral phase $\phi(\omega)$, but it would be nice to have a way to link our quantities $\phi^{(1)}$ and $\phi^{(2)}$ to the refractive index of a medium so we can get a feel for how our pulses will broaden and move with time as they pass through it. One way to do this is to first note that the wavenumber, $k(\omega)$, can be written as follows [3, 1]:

$$k(\omega) = \frac{\omega}{c} = \frac{n(\omega)\omega}{c_0} \tag{2.16}$$

Where $n(\omega)$ is the frequency dependent refractive index, ω is the frequency of the wave,c is the phase velocity of the wave with frequency ω and c_0 is the speed of light in vacuum. The wavenumber $k(\omega)$ is the magnitude of the **wave vector**, and has units of radians per unit length, and points in the direction that the wave propagates. The momentum of the wave is directly proportional to the wave vector (see Appendix A for further detail).

A good way to think of the wave vector is as a vector which propagates in the direction of the travel of the wave, and has a magnitude of the wavenumber, which is proportional to the energy of the wave. A wave with a higher frequency has a higher energy, and the units of $k(\omega)$ are essentially inverse wavelength - shorter wavelengths leads to larger $k(\omega)$ and a higher energy wave.

The accumulated phase can be written in terms of the wavenumber, because at any given point on the wave's travel r, the phase can be written as $k(\omega)r$. The *accumulated phase* as the wave travels between two points r_1 and r_2 is therefore given by

$$\phi(\omega) = k(\omega)r_2 - k(\omega)r_1 = k(\omega)(r_2 - r_1) = k(\omega)L \tag{2.17}$$

Where $L = r_2 - r_1$ and represents the distance that the wave has travelled.

We can now learn something about the form of the spectral phase $\phi(\omega)$ by expanding it as a Taylor series around ω_0 again, but now in terms of k rather than directly as ϕ:

$$\phi(\omega) = k(\omega)L = k(\omega_0)L + \frac{\partial k}{\partial \omega}\bigg|_{\omega=\omega_0}(\omega - \omega_0)L + \frac{1}{2}\frac{\partial^2 k}{\partial \omega^2}\bigg|_{\omega=\omega_0}(\omega - \omega_0)^2 L\,\mathrm{d}\omega \tag{2.18}$$

We know from comparison with Equation 2.8 that we can identify these three terms as relating to the CEP, group delay, and group delay dispersion. The question is now whether we can gain some insight into the physical origins of these quantities by considering the derivatives of the wave vectors. From Equation 2.16, we can write the derivatives in Equation 2.18 in terms of $n(\omega)$ and find the following by using the product rule on Equation 2.16:

$$\text{Group Delay} = L\frac{\partial k}{\partial \omega}\bigg|_{\omega=\omega_0} = \frac{L}{c_0}(n(\omega_0) + \omega_0 n'(\omega_0)) \tag{2.19}$$

$$\text{GDD} = L\frac{\partial^2 k}{\partial \omega^2}\bigg|_{\omega=\omega_0} = \frac{L}{c_0}(2n'(\omega_0) + \omega_0 n''(\omega_0)) \tag{2.20}$$

Where the notation $n'(\omega_0)$ and $n''(\omega_0)$ refers to the first and second derivatives of n with respect to ω respectively. We can also relate the derivatives given here with our previous derivatives from Equation 2.8 to unpick the meaning of the wavevector derivatives explicitly by thinking about the units of the related quantities:

$$\frac{\partial \phi}{\partial \omega} = L\frac{\partial k}{\partial \omega} \rightarrow \text{Group Delay (s)} = L(\text{m}) \times (\text{Group Velocity})^{-1}(\text{s}\,\text{m}^{-1})$$

So $\frac{\partial k}{\partial \omega}$ is the reciprocal of the group velocity. Similarly:

$$\frac{\partial^2 \phi}{\partial \omega^2} = L\frac{\partial^2 k}{\partial \omega^2} \rightarrow \text{GDD (s}^2) = L(\text{m}) \times \text{GVD}(\text{s}^2\,\text{m}^{-1})$$

So $\frac{\partial^2 k}{\partial \omega^2}$ is the GVD (Group Velocity Dispersion - GDD per unit length).

Now, looking into equations Equation 2.19 and Equation 2.20 can give us some physical insight into the origins of the group delay and the GDD with respect to the refractive index. Initially considering the group delay (Equation 2.19), we can see that the group delay will be larger if the refractive index of the material at ω_0 is large, and also if the rate of change of the refractive index with ω ($n'(\omega)$) is large. If the refractive index of the material does not change with frequency (so that $n'(\omega_0) = 0$), we would still have a group delay – the pulse would still be delayed in time as it propagates through the medium.

Now considering the GDD (Equation 2.20), we see that there are no longer terms that depend only on $n(\omega_0)$. Instead, the GDD depends on both the first and second derivatives of the refractive index with frequency. This means that unless our refractive index is independent of frequency ($n \neq n(\omega)$), then our pulse will **always broaden** as it passes through a material. This is exactly what we described qualitatively earlier in this chapter, where we stated that it was the frequency dependence of n which causes our dispersion. Moreover, we can now go further and say that we will have more GDD (and more broadening) if our refractive index changes more drastically with frequency (so the derivatives are larger), and a situation where $n(\omega)$ is a higher order function of ω will generally cause even more broadening (as then $n''(\omega_0)$ is higher). Finally, it is also clear from the expressions above that the group delay and GDD are larger if the pulse propagates for a longer distance through the medium (larger L) – so sending a short pulse through a thicker piece of glass will incur more broadening than sending it through a thin piece of glass would.

2.4 Group Delay Dispersion

We now return to a more qualitative discussion of the GDD, as needed to enable effective lab working, and to summarise the conclusions of the previous two sections. If you are

buying optics and want to assess how much dispersion they will add, you will typically find the GDD quoted in units of fs^2. The effect of the GDD on a pulse can be summarised by writing the time taken $T(\omega)$ for a given pulse to propagate through a medium as:

$$T(\omega) = \text{Group Delay} + \text{GDD} \times \Delta\omega \tag{2.21}$$

Equation 2.21 should make clear four things:

- If the GDD is zero, then the propagation time is just the group delay, and the pulse is transform limited.
- If the GDD is positive, the propagation time for frequencies greater than ω_0 is longer than that for frequencies lower than ω_0. Thus, the 'blue' end of the pulse takes longer to travel through the medium and lags behind the 'red' end. A **positive chirp** has been induced in the pulse.
- If the GDD is negative, then the inverse is true. The 'red' end of the pulse lags behind the 'blue' end, and a **negative chirp** has been induced in the pulse.
- With non-zero GDD, a larger bandwidth $\Delta\omega$ will lead to a longer travel travel time, more pronounced chirp, and a broader pulse.

Hopefully by now you understand they key message: dispersion (and specifically GDD) causes pulses to broaden. As pulses are dispersed away from the transform limit, they become either positively or negatively chirped, leading to temporal broadening of the pulse as the different colours in the pulse spread out in time.

2.5 Predicting Broadening from Dispersion

Next we consider how we can predict the broadened pulse length of a Gaussian pulse, which will be of considerable practical utility. Assuming that there is only second-order dispersion[6], then the broadened pulse duration τ of an initially transform-limited Gaussian pulse of duration τ_0, having passed through a medium with a given GDD is:

$$\tau = \tau_0 \sqrt{1 + \left(\frac{4 \ln 2 \, GDD}{\tau_0^2} \right)^2} \tag{2.22}$$

Or, equivalently, for a material with a known GVD of length L:

$$\tau = \tau_0 \sqrt{1 + \left(\frac{4 \ln 2 \, GVD \times L}{\tau_0^2} \right)^2} \tag{2.23}$$

Equation 2.22 is plotted for for a transform-limited 35 fs input pulse as a function of GDD in Figure 2.5. Clearly, as the initial pulse was transform limited, then any amount of positive or

[6] This is usually reasonable - and most often if there is third-order dispersion we can't do a lot about it anyway.

negative GDD added will cause it to broaden. It can also be seen from the above equations that a shorter input pulse (smaller τ_0) will lead to a broader output pulse. This is easily explained by the fact that a short pulse has a larger bandwidth, and will experience a wider range of refractive indices, with the result that the different colours in the pulse will spread out more in time.

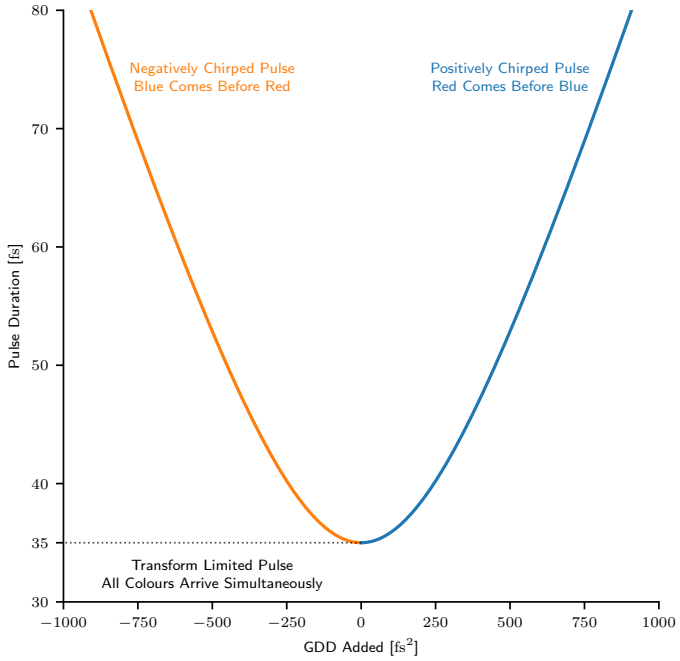


Figure 2.5: Pulse duration of an initially transform limited 35 fs 800 nm Gaussian pulse as a function of the added GDD. Negative GDD added leads to negative chirp - shown in orange. Positive GDD added leads to positive chirp - shown in blue.

Figure 2.5 and Equation 2.22 both ignore higher order dispersion. This is normally a reasonable approximation, except when considering the shortest pulses. A further point to consider is *how the GDD depends on wavelength*. In general, shorter wavelengths (especially in the UV), experience higher refractive indices, leading to more severe broadening as they accumulate **much** more GDD than an infrared pulse of similar duration. Table 2.2 the GVD (GDD per unit length) for some common materials at a range of wavelengths.

Clearly the amount of dispersion is higher in the UV than in the IR for all of these materials, and passing a 35 fs TL 266 nm pulse through a 5 mm UVFS window will lead to $\sim$1000 fs^2 of added GDD - broadening the pulse from 35 fs to $\sim$90 fs - more than a factor of two! It is therefore important to be aware of the effects of dispersion, especially when working with short wavelength pulses. To illustrate this further, we can consider the actual temporal broadening produced by some typical scenarios, summarised in Table 2.3. The characteristics of the input pulses have been chosen to reflect common outputs of Ti:Sa and Yb:KGW laser systems.

As should now be clear, to avoid pulse broadening, there are some good rules of thumb:

Substrate	GVD (fs^2/mm)			
	1030nm	800nm	515nm	266nm
BK7	25.1	44.6	84.3	N/A
UV Fused Silica	18.9	36.2	68.7	197.5
CaF_2	18.3	27.7	48.6	128.0
Air	0.013	0.02	0.035	0.10
Water	-0.97	24.88	58.0	207.2

Table 2.2: GVD of different materials at a selection of common wavelengths.

Input Pulse	Propagates Through	Output Pulse Duration
1030nm, 200fs	1cm UVFS	200fs
1030nm, 200fs	10m Air	200fs
515nm, 200fs	1cm UVFS	200fs
515nm, 200fs	10m Air	200fs
800nm, 35fs	1cm UVFS	45fs
800nm, 35fs	10m Air	38fs
800nm, 35fs	10 bounces on ultrafast mirrors	42fs
266nm, 70fs	1cm UVFS	105fs
266nm, 70fs	10m Air	80fs
266nm, 70fs	10 bounces on Al mirrors	$\sim$70fs

Table 2.3: Broadening of various input pulses by propagation through various media.

- Shorter pulses (broader bandwidth) are a lot more susceptible to broadening than longer pulses (narrower bandwidth). The pulses typically produced from the output of a Yb:KGW laser (usually >200 fs duration) will not significantly broaden except when large amounts of GVD are added. The converse is true for Ti:Sa lasers.

- UV pulses are a lot more susceptible to broadening than IR pulses, because refractive index changes with wavelength more sharply in the UV region.

- Minimising transmissive optics is the most important thing to do if you are using short pulses and want to keep pulses short.

- Bouncing on metallic mirrors or ultrafast-coated mirrors is generally OK, but other mirrors can induce quite significant broadening. Beware of 'broadband' mirrors which are not 'broadband *ultrafast*' mirrors.

2.5.1 Dispersion of Optical Elements

From the previous discussion in this chapter, we should surmise that predicting the amount of dispersion added by a particular optical element within a setup is generally a relatively straightforward matter of looking up the value of the GDD or GVD for the optical material

in question. Most suppliers of optics will give this information readily on their websites. It is important that you do not **only** consider transmissive optics when thinking about pulse broadening. Mirrors will also contribute some GDD, and especially so if you are not using mirrors that are specially coated for ultrafast pulses. Generally an 'ultrafast' mirror will have a GDD of $<30\,\mathrm{fs}^2$ per bounce. The physical origin of GDD from a mirror, and more detailed discussion of different optical elements is in chapter 6.

Aside from supplier catalogues/websites for specific optics, a very useful resource for finding the GDD for a large number of materials is the **Light Conversion Optics Toolbox**: toolbox.lightcon.com [6]. This contains a huge number of helpful calculators. Under 'Dispersion Calculators' there are applets to calculate arbitrary dispersion, pulse broadening, and much more. These calculators can also include third-order dispersion, which can be useful if you need to establish whether or not this can be neglected for your application.

2.6 Higher Order Dispersion

So far we have neglected higher order dispersion in our discussion. By 'higher order dispersion' we mean terms above the GDD term in Equation 2.8. The next term to consider in this expansion is the term that gives rise to **third-order dispersion**, $\phi^{(3)}$:

$$\phi^{(3)} = \frac{\partial^3 \phi}{\partial \omega^3} \tag{2.24}$$

In general the effects of third-order dispersion (TOD) are normally negligible except for in two important circumstances:

- Extremely short pulses ($<20\,\mathrm{fs}$) are being used (so large bandwidths).
- Extremely dispersive media at a given wavelength are being used (or we travel long path lengths through normal dispersive media, such as when using optical fibres).

We could repeat the analysis we performed for $\phi^{(1)}$ and $\phi^{(2)}$ to try and get an intuitive feel for the effect that TOD will have on a laser pulse. However, for our purposes, doing this leads to a situation where we realise that really the effect of TOD is just to introduce 'more complicated broadening'. For example, earlier on we assessed the effect of GDD by neglecting all other terms in our phase expansion to find that on propagation through a medium with a known GDD, we got a shift in the time axis of our pulse from t to $t + \phi^{(2)}\Delta\omega$, and so the shift in the time axis depends on frequency, and every colour in our pulse is affected different – resulting in broadening. An equivalent analysis for TOD would lead to finding a shift from t to $t + \phi^{(3)}\Delta\omega^2$. We still shift every frequency in our pulse differently, but in a way that now depends quadratically on the detuning of that frequency from the central frequency. This still gives broadening, but more complicated broadening that is more severe for frequencies at the edges of our pulse bandwidth. Similarly, we can try to find out the dependence of the TOD on refractive index and we find that:

$$\mathrm{TOD} = L\frac{\partial^3 k}{\partial \omega^3}\bigg|_{\omega=\omega_0} = \frac{L}{c_0}(3n''(\omega_0) + \omega_0 n'''(\omega_0)) \tag{2.25}$$

Clearly now the amount of dispersion depends on the third derivative of the refractive index with respect to frequency! Such expressions are not immediately helpful to us. We could also repeat the simple numerical model shown in Figure 2.4, and we'd find something that looks like Figure 2.6. Immediately clear from this is the distortion of the pulse envelope and the creation of some sideband type 'wings'. Really, the model being used to simulate this is too simple to capture all the effects of higher order dispersion, but it serves to illustrate the point that it gets a complicated!

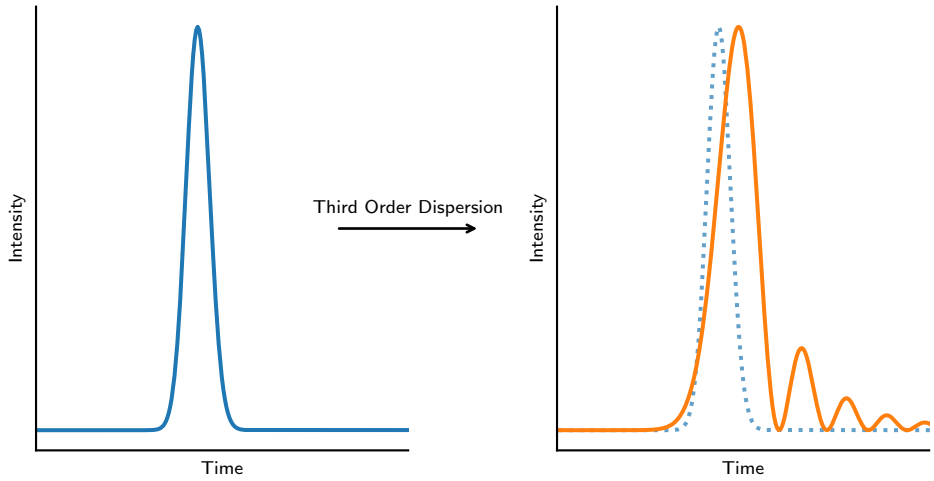


Figure 2.6: Crude numerical simulation of the effect of third-order dispersion on an ultra-short pulse. Weird 'wings' are visible – more sophisticated pulse modelling techniques are really needed to understand this behaviour when required.

So, what do you do if you are working with extremely short pulses and need to characterise the TOD? Helpfully, most manufacturers (and some online calculators) will specify and allow you to calculate TOD given by an optic, or state how a given compressor can compensate for it. It is also possible to measure the TOD, but as pulses get shorter and material interactions become more complex, serious numerical modelling of the pulse propagation becomes increasingly necessary[7]

2.7 Pulse Compression - Compensating for Dispersion

Finally, we will briefly considering how we can remove unwanted dispersion and make our pulses short again (practical discussion of this will happen later). As may be evident from Figure 2.5, we want to push our pulse back down the curve towards the transform limit, i.e. towards having zero chirp. How we achieve this depends on whether the pulse has accumulated net positive or negative GDD - whether it is positively or negatively chirped. If we can establish this, then we simply have to add GDD of the opposite sign to shift the pulse back towards the transform limit.

[7] RP Photonics produce software for this purpose (*RP ProPulse* – [7]), or if you are more mathematically inclined than me you can program your own. Reference [8] could be a useful starting point here.

While straightforward in principle, unfortunately there are not really any materials in the UV-near IR region we're normally working in which exhibit negative GDD[8]. This has two consequences:

1. If our pulses start life as transform limited pulses, they will almost inevitably end up gaining only positive GDD. The resulting positively chirped pulses will need negative GDD to recompress them back to the transform limit.

2. Adding said negative GDD is challenging due to the lack of materials that exhibit negative GDD in this spectral region.

Given the above, we usually cannot just get another piece of exotic glass and send the pulse through it to recompress it. This can be possible in the mid-IR/far IR, where some common optical materials (such as BK7) add negative GDD[9], but generally is not possible for the 200 nm-1000 nm pulses easily produced from Ti:Sa or Yb:KGW lasers.

Happily, there is a way we can add this negative GDD, by using a **compressor** made using prisms, gratings, or chirped mirrors. These devices are constructed in a geometry such that an initially positively chirped pulse enters (red comes before blue), and within the compressor the red components are delayed relative to the blue components as a result of differing path lengths. The blue components therefore have time to 'catch up' to the red components, adding negative GDD and reversing the positive chirp, compressing the pulse. We will discuss specific compressor geometries further in chapter 4.

2.8 Broad Bandwidth Miscellany

As we end our discussion of the properties of ultrafast laser light, there are a few consequences of using broadband pulses that don't inherently relate to temporal dispersion or pulse broadening, but that are worth being aware of in the lab. They are briefly discussed here for completeness, and relevant ones will be discussed in more depth in relevant sections. They pretty much all boil down to the optical properties of materials being wavelength dependent.

2.8.1 Polarisation Optics

Polarisation was briefly discussed in chapter 1 and polarisation optics will be discussed extensively in chapter 6. However, it is useful here to make some brief comments about polarisation in the context of broadband pulses.

If a laser pulse consists of many different colours, we know that we can represent the pulse by thinking of it as a superposition of different waves with different frequencies. Each of these

[8] Eagle-eyed readers will have noted water exhibits negative GDD at 1030 nm, but not a lot. Perhaps an aquarium-style compressor is possible though.

[9] As these materials have both positive and negative GDD, there is generally a wavelength where they exhibit zero GDD - **the zero dispersion wavelength**.

waves has a polarisation state and can be affected by polarisation optics (polarisers and waveplates) differently. When using ultrafast lasers it's important to ensure that whatever you do the polarisation state of the pulse, you do *to the whole pulse*. Many polarisation optics are not designed to work at broad bandwidths – so you can end up only polarising part of your pulse, which can have severe knock-on effects[10]. To avoid this, it is important to buy polarisation optics that are 'broadband', 'achromatic', or otherwise specifically specified for your use case.

2.8.2 Focussing Optics

Transmissive focussing optics such as lenses are defined by their *focal length* (see chapter 6 for further details). Like most optical properties of a material, the focal length is wavelength dependent. When using a singlet lens this can then result in **chromatic aberration** where different colours within the pulse are focussed to different points in space, as shown in Figure 2.7. If severe enough, this can have the effect of smearing out an interaction region in an experiment or creating imaging aberrations (among other things). This issue is most severe when using extremely broadband sources, such as when a *white-light continuum* (see chapter 3) is used.

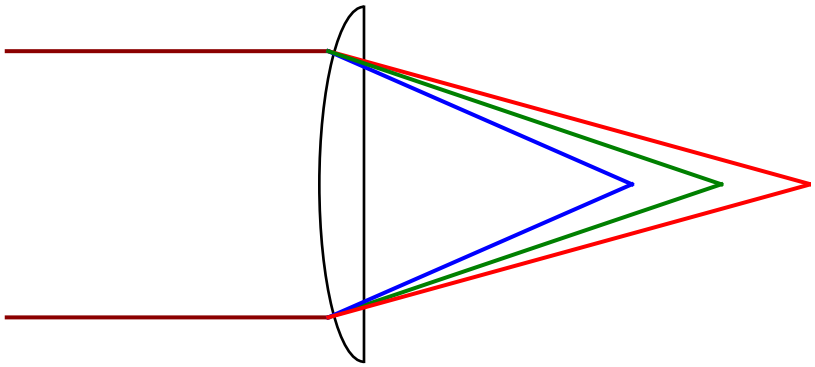


Figure 2.7: Chromatic aberration. Different colours of light experience different refractive indices of a lens and thus have different focal lengths.

To avoid this issue, if you are using very broadband pulses such as a generated white-light continuum, it is best to use **reflective** focussing optics such as spherical or parabolic mirrors[11]. However, in general this issue is not severe enough to warrant not using transmissive lenses for typical laser pulses from a laser output (they are not broadband enough to cause much of an issue). Other solutions from traditional imaging optics such as using *achromatic*

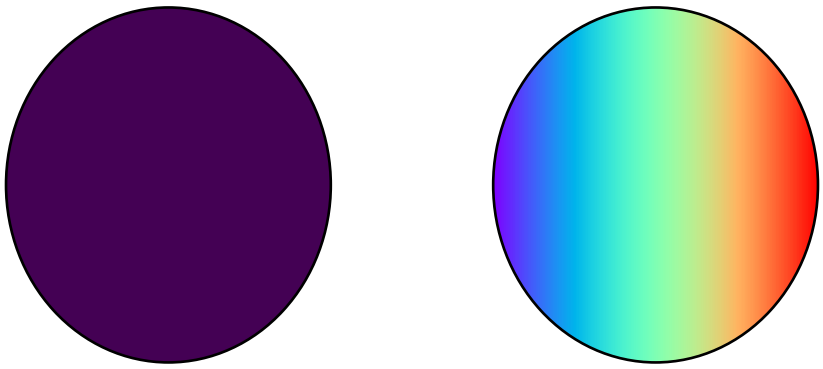
[10] For instance, you could have an ultrashort pulse which goes through a waveplate that only rotates half of the bandwidth. This half of the bandwidth would then get rejected by a polariser further downstream in the beamline.

[11] Which have the additional benefit of not adding significant dispersion to the pulses.

doublets or multiple lenses together are generally best avoided when using ultrafast lasers – the large amount of glass this number of transmissive optics will add to the beam path will generally have a negative effect on the pulse duration.

2.8.3 Spatial Chirp

We've discussed dispersion and chirp extensively in this chapter, but have always been talking about **temporal** dispersion and **temporal** chirp, where different colours of light within the laser pulse arrive at different points in time. Another kind of chirp that can arise is **spatial chirp**, which occurs when different colours within the pulse arrive at different points **in space**. Practically, this means that if you were to look at the laser beam, you'd see different colours at different points of the beam profile – as shown on the right hand panel of Figure 2.8



Colour Uniform Across Beam:
No Spatial Chirp

Colour Changes Across Beam:
Spatial Chirp

Figure 2.8: Simulated beam profiles with no spatial chirp (left) and severe spatial chirp (right), where different parts of the beam face are different colours. The extent of the spatial chirp has been exaggerated here for illustration.

Spatial chirp can be added to pulses by any dispersive optical element (prisms, gratings, tilted bits of glass). Generally when prism or grating compressors are designed, the designer takes care to not add excessive spatial chirp to pulses – however it is often difficult to eradicate completely. Spatial chirp can have severe consequences in microscopy as it is difficult to focus a beam to a tight focus when the colour of the beam is changing across the face of the beam.

Bibliography

[1] Hooker S and Webb C. *Laser Physics*. Oxford University Press, Oxford, 1st edition, 2010.

[2] Milonni P W and Eberly J H. *Laser Physics*. Wiley, New York, 1st edition, 2010.

[3] Manzoni C and Cerullo G. Design criteria for ultrafast optical parametric amplifiers. *J Opt*, 18, 2016.

[4] Kreyszig E. *Advanced Engineering Mathematics*. Wiley, New York, 8th edition, 1999.

[5] Sanderson G. But what is the fourier transform? a visual introduction. https://www.youtube.com/watch?v=spUNpyF58BY. Accessed: 23-12-2020.

[6] Light Conversion. Optics toolbox. toolbox.lightcon.com. Accessed: 23-12-2020.

[7] Paschotta R P. The rp photonics encyclopedia. https://www.rp-photonics.com/encyclopedia.html. Accessed: 22-12-2020.

[8] A Couairon, E Brambilla, T Corti, D Majus, O de J. Ramirez-Gongora, and M Kolesik. Practitioner's guide to laser pulse propagation models and simulation. *Eur. Phys. J. Special Topics*, 199, 2011.

Chapter 3

Nonlinear Optics

In this chapter we discuss nonlinear optics. Nonlinear optical effects are heavily utilised in ultrashort laser pulse generation, characterisation, and manipulation – so an understanding of them is essential. Here we will introduce key ideas around nonlinear material response, the concepts of self-focussing and self-phase modulation, frequency mixing, and finally optical parametric amplification.

Before we consider how to generate, characterise, and practically use ultrashort pulses, it will be useful to discuss **nonlinear optics**, as many of the techniques discussed later make heavy use of nonlinear phenomena. Nonlinear optics are not exclusively relevant to ultrafast lasers and optics, but ultrashort pulses necessarily have a very high *intensity*, so nonlinear effects can be very significant.

The physics and mathematics behind a lot of nonlinear optics quickly become quite dense, and would get in the way of the more elementary understanding we need as non-specialists using these pulses in our experiments. In this chapter we limit ourselves to briefer discussions of nonlinear phenomena that are important in ultrashort pulse generation and characterisation, and discuss concepts in an order which may alienate a purist reader but enable us to more quickly gain the qualitative understanding we need. For the more purist reader, most books on laser physics will offer a more rigorous treatment of nonlinear optical effects than will be presented here [1, 2], and Boyd's *Nonlinear Optics* [3] is an excellent next step.

3.1 Nonlinear Material Response

As discussed previously, a laser pulse can be effectively described as an oscillating electric field (the carrier wave), within a more slowly varying envelope. When a medium (air, glass, anything) is subjected to an electric field **E**, the response of the material is given by the **polarisation density P**, of the medium. Polarisation density has units of 'dipole moment per unit volume', so if the applied electric field produces a large dipole moment, then the

medium is said to be **polarisable**[1]. The polarisation density, **P**, can be expressed as follows:

$$\mathbf{P} = \epsilon_0 \chi^{(1)} \mathbf{E} + \epsilon_0 \chi^{(2)} \mathbf{E}^2 + \epsilon_0 \chi^{(3)} \mathbf{E}^3 + ... \tag{3.1}$$

Where $\chi^{(n)}$ is the n^{th} order electric susceptibility, and ϵ_0 is the vacuum permittivity. The electric susceptibility $\chi^{(n)}$ is best thought of as the degree to which the material is affected by the field, and intuitive mental picture is to think that the oscillating electric field applied to the material makes the charges (such as electrons) in the material oscillate too – what I like to call the **jiggling electron** picture (Figure 3.1).

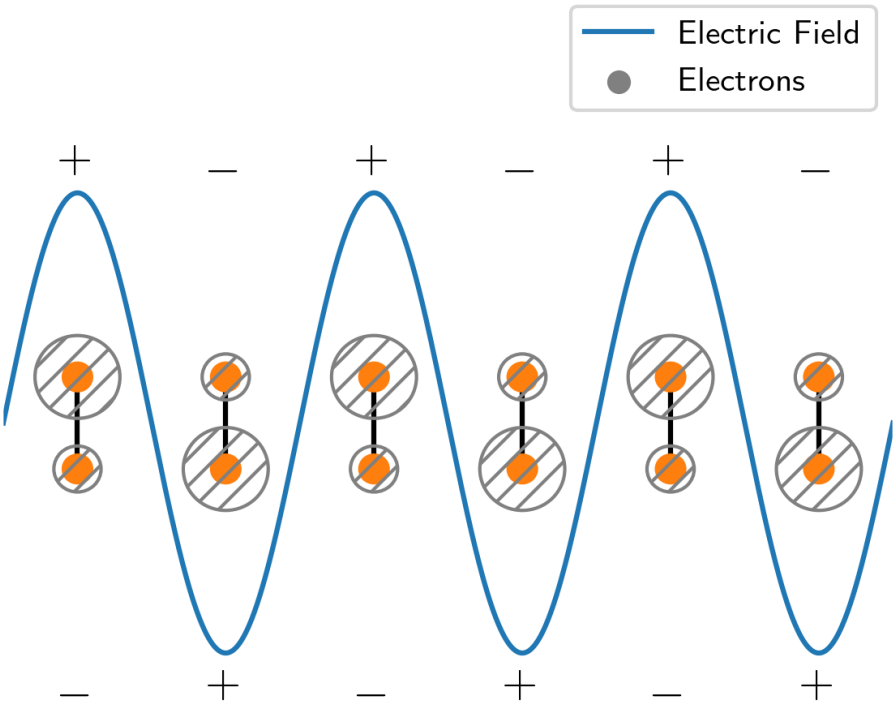


Figure 3.1: Jiggling electrons: the electrons (grey hatched) on a molecule are driven around by the presence of an oscillating electric field.

For weak fields, such as those generated by lamps, LEDs, or low power lasers all terms apart from the first term in Equation 3.1 can be considered to be zero. However, for intense fields, which we generally have when using ultrafast pulses, the higher order susceptibilities can become important. To put some rough numbers on what we mean by *weak* and *intense* fields here, a weak field is the kind of field you'd experience in your day to day life (bright LED, laser pointer). An intense field strong enough to induce significant nonlinear interactions

[1]Note that this is not the same thing as the polarisation **of** the electric field, which is the orientation of the field oscillation.

will have an intensity above around $10^{12}\,\mathrm{W\,cm^{-2}}$, but not an intensity that is high enough to rip the electrons off the material (less than around $10^{16}\,\mathrm{W\,cm^{-2}}$). In short, the field is strong enough to make the electrons in the medium jiggle in exciting ways, but not strong enough so that they are actually ionised from the medium. Such fields are readily produced by ultrafast lasers.

3.1.1 Intensity-Dependent Refractive Index

To start with a simple motivating case, it turns out that in isotropic media (where the medium behaves identically regardless of which of the three symmetry axes of it we look at - most of the media we use are isotropic, with special cases discussed later), shining an intense field on the material gives rise to an **intensity dependent refractive index**, $n(I)$:

$$n(I) = n_0 + n_2 I(r,t) \tag{3.2}$$

The first term in this equation, n_0, arises from the first order susceptibility $\chi^{(1)}$, and is responsible for the normal frequency-dependent refractive index we are familiar with from previous chapters, which in turn is responsible for familiar optical effects such as reflection, refraction, and absorption. The second term in Equation 3.2 is the product of the **nonlinear refractive index**, n_2, and the position- and time-dependent intensity of the pulse $I(r,t)$. n_2 arises from the *third-order* susceptibility[2], and the effect of the $n_2 I(r,t)$ term causes the refractive index experienced by a particular frequency within our laser pulse to be different at the most intense parts of the pulse than at the less intense parts.

In physical terms, this means that the peak of the pulse in time, and the center of the beam in space (see Figure 1.3 and Figure 1.12), will experience higher refractive indices than the at the edges[3].

This dependence of n on intensity is known as the **Optical Kerr Effect**, and leads to two phenomena that it is useful to be aware of. These phenomena happen when the intensity-dependent refractive index varies both in *space* – leading to **self-focussing** – and in *time* – leading to **self-phase modulation**.

3.1.2 Self-Focussing

We first consider the phenomenon of **self-focussing**. Self-focussing occurs when an intense laser pulse passes through a medium and experiences an intensity dependent refractive index **in space** via the Kerr effect (Equation 3.2). As discussed earlier, most laser beams we consider can be well described as **Gaussian beams**, which means that the transverse

[2] The second-order susceptibility is zero unless the medium is anisotropic - see later.

[3] The extent of the refractive index difference caused by the Kerr Effect depends on both the pulse intensity and the magnitude of n_2, which is material dependent. n_2 is normally positive, so the refractive index is higher for areas of high intensity. n_2 has units of $\mathrm{cm^2\,W^{-1}}$, and for typical transparent materials is on the order of $10^{-15}\,\mathrm{cm^2\,W^{-1}}$.Thus we can get a feel for the sort of light intensities above which nonlinear effects start to become significant: n_0 is around 1 for most transparent materials, and so n_2 starts to become comparable (within a factor of 1000) to n_0 at light intensities of around $10^{12}\,\mathrm{W\,cm^{-2}}$.

intensity across the face of the beam is described by a 2D Gaussian function (Figure 1.12), such that the intensity in the center of the beam is higher than that at the edges.

Let us imagine that a pulse with a Gaussian intensity profile in space passes through a medium. If the intensity is high enough, it could induce an intensity dependent refractive index in the medium via the Kerr effect. This effectively leads to a lens being *imprinted* on the material by the laser beam, and this *'induced lens'*[4] leads to focussing of the beam just as a conventional lens would. This focussing of the beam as it passes through the material is known as **self-focussing**, and is illustrated in Figure 3.2. Self-focussing is important in the phenomenon of **Kerr-Lens Modelocking**, discussed later.

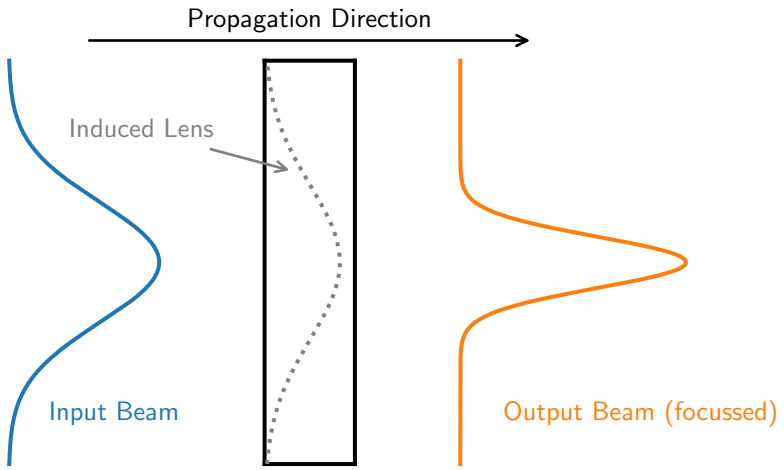


Figure 3.2: Schematic of the self-focussing process. Transverse profiles of a Gaussian beam are shown. An input beam (blue) passes through a dispersive material (black box) and imprints a transient 'lens' on the material via the Kerr effect (grey dashed line). This leads to focussing such that the output beam (orange) has a higher transverse intensity. The extent of this effect has been exaggerated here for illustrative purposes.

This process of self-focussing is important for us as laser users. It is deliberately used inside ultrafast lasers in some situations (discussed later), but in most practical optics it is a problematic effect. If we are not careful we can have a situation in which self focussing in our optics causes us to inadvertently focus the beam, which can damage or destroy optics that are further down the beamline! Aligning optics at low power and then slowly increasing the power with careful observation of the consequences should allow problems to be identified before any damage is caused. You do not want to accidentally focus the beam onto an expensive autocorrelator or inadvertently drill a hole in your vacuum chamber window[5].

[4] Sometimes known as **Kerr Lens**.
[5] Which does rather tend to spoil the vacuum.

3.1.3 Self-Phase Modulation

The second effect that arises from the Kerr effect is known as **self-phase modulation** (SPM). SPM occurs due to the intensity at the peak of the pulse **in time** being greater than at the edges. As the intensity of the pulse varies with time (with peak intensity at the peak of the pulse envelope), the pulse will 'see' a time-dependent refractive index as it travels through the medium. This causes a phase shift in different parts of the pulse in time which results in a change in the instantaneous frequency $\omega(t)$, of the pulse which is defined in Equation 3.3.

$$\omega(t) = \frac{\mathrm{d}\phi(t)}{\mathrm{d}t} \tag{3.3}$$

Where $\phi(t)$ is the temporal phase of the pulse. After undergoing SPM $\phi(t)$ contains a contribution from the SPM process, denoted ϕ_{NL}:

$$\phi(t) = \phi_0 + \phi_{NL}(t) \tag{3.4}$$

The instantaneous frequency of the pulse can then be written as:

$$\omega(t) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\phi_0 + \phi_{NL}(t)\right) = \omega_0 + \delta\omega(t) \tag{3.5}$$

Where $\delta\omega(t)$ is the change in instantaneous frequency due to SPM at a time t. The magnitude of $\delta\omega(t)$ is proportional to the rate of change of the pulse intensity with time, that is:

$$\delta\omega(t) \propto \frac{\mathrm{d}I(t)}{\mathrm{d}t} \tag{3.6}$$

From here it is clear that at early times, as the pulse intensity increases ($\frac{\mathrm{d}I(t)}{\mathrm{d}t} > 0$), that $\delta\omega$ is positive and the instantaneous frequency rises relative to ω_0. Conversely, as the pulse intensity decreases (at later times) the opposite is true and the instantaneous frequency decreases relative to ω_0. $\delta\omega(t)$ (orange solid line) is plotted over the incident pulse intensity (blue dotted line) in Figure 3.3a for an arbitrary transform-limited Gaussian pulse undergoing SPM. It is straightforward to calculate $\delta\omega(t)$ for different pulse shapes[6].

The net effect of SPM is the broadening of the pulse spectrum due to the additional frequencies created during the SPM process, as plotted in Figure 3.3b. It is a little harder to develop an intuitive picture for SPM in the same way that we did for self-focussing, but one explanation may go as follows:

- An intense enough pulse passing through a medium experiences a refractive index that varies with the pulse intensity.
- The change in refractive index causes a change to the instantaneous frequency of the pulse.
- Because the sign of the refractive index change is different on each side of the peak of the pulse in time, the sign of the instantaneous frequency change is different too.

[6] See for example Boyd [3] section 7.5.1 - for the process that was followed to make Figure 3.3

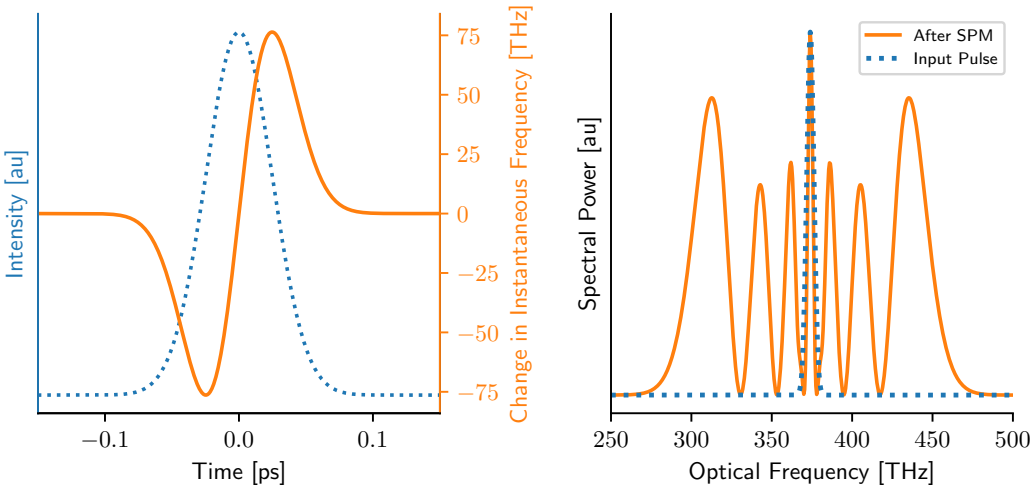


Figure 3.3: (a) An input pulse (blue dotted, left axis) experiences a time-dependent refractive index change which results in the creation of new frequencies $\delta\omega$ via SPM (orange solid, right axis). (b) Spectral broadening of a transform-limited input pulse (blue dotted) via SPM leads to the creation of new frequencies, broadening the pulse (orange solid). Note spectra in (b) are normalised to the same maximum height for illustrative reasons.

- This leads to the creation of new frequencies at either side of the pulse central frequency – broadening the pulse.

SPM can essentially be thought of as an energy redistribution process – energy from the pulse is distributed out to more frequencies in the increased bandwidth. An example of an SPM broadened pulse is shown in Figure 3.3b (orange solid line), where it is clear that the bandwidth has increased dramatically. However, it is important to note that SPM does not *always* result in an increased pulse bandwidth – it is possible that the SPM process narrows the pulse rather than broadening it, especially if the pulse was originally negatively chirped. Furthermore, the simple example of SPM spectra shown in Figure 3.3 ignore other effects of dispersion as the pulse passes through the material, and in general accurate modelling of the shape of a pulse spectrum after SPM requires complex numerical modelling of the pulse propagation that is beyond the scope of this book.

Anyway, why do we care about SPM? Well, the spectral broadening often induced by SPM can be important when creation of very short pulses is desired, as effectively you are creating 'extra bandwidth' via the SPM. Compression of this spectrally broadened pulse will then lead to a shorter pulse than could have been obtained from compression of the initial, non-broadened, pulse. It is also commonly used in spectroscopy where the broadening is used to create 'white light' from an ultrafast laser, which enable a wide range of molecular energy levels to be excited in a single laser shot (see chapter 8 for an example of how this is used).

3.2 Nonlinear Frequency Mixing

Another of the most important nonlinear effects to be aware of is that of **nonlinear frequency mixing**. This is a technique where two or more input photons are combined to give output photons at combinations of the frequency of the input photons. Frequency mixing techniques are classified as *'N-wave mixing'* where N is the number of fields involved in the process. We will mostly consider the simplest case, three-wave mixing, here.

3.2.1 Three-Wave Mixing Techniques

If we consider mixing of only two photons with frequencies ω_1 and ω_2, then frequency mixing occurs when a third photon ω_3 is produced such that:

$$\omega_3 = \omega_1 \pm \omega_2 \tag{3.7}$$

That is, the frequency of the third photon is either the sum or the difference of the frequencies of the two initial photons. This is called **three-wave mixing**[7]. In the case where $\omega_3 = \omega_1 + \omega_2$, this is called **sum-frequency generation (SFG)**, and if $\omega_3 = \omega_1 - \omega_2$, then this is called **difference-frequency generation (DFG)**. These are the most general cases, but there are other possibilities with unique names which are summarised in Table 3.1. How

Input Photons	Output Photon	Name
ω_1, ω_2	$\omega_1 + \omega_2 = \omega_3$	Sum-Frequency Generation (SFG)
ω_1, ω_2	$\omega_1 - \omega_2 = \omega_3$	Difference-Frequency Generation (DFG)
ω_1, ω_1	$\omega_1 + \omega_1 = 2\omega_1$	Second Harmonic Generation (SHG)
ω_1, ω_2 (where $\omega_2 = 2\omega_1$)	$\omega_1 + \omega_2 = 3\omega_1$	Third Harmonic Generation (THG)

Table 3.1: Nonlinear frequency mixing schemes.

the nonlinear material response described in Equation 3.1 can result in the creation of sum or difference frequencies is easily shown with a simple example (see Appendix D for a more detailed derivation). We can write an input wave $\mathbf{E}_{\boldsymbol{\omega}}$ in terms of the frequency of ω, such as:

$$\mathbf{E}_\omega \propto \exp(i\omega t) \tag{3.8}$$

This wave in the nonlinear medium gives rise to a second-order polarisation $\mathbf{P}^{(2)}$ according to Equation 3.1:

$$\mathbf{P}^{(2)} \propto \epsilon_0 \chi^{(2)} \mathbf{E}_{\boldsymbol{\omega}}^2 \tag{3.9}$$

Where:

$$\mathbf{E}_\omega^2 \propto \exp(i\omega t)^2 = \exp(i2\omega t) \tag{3.10}$$

Hence, $\mathbf{E}_\omega$ can produce a second harmonic wave with frequency 2ω through $\chi^{(2)}$, and will do so if the intensity is high enough, **provided that $\chi^{(2)}$ is non-zero**. This is a key concept, and implies that we can only perform this kind of frequency-mixing in a medium that lacks inversion symmetry[8]. Most media are isotropic (for us, this means they have

[7] Which seems bizarre when we only mixed *two* waves to do it. The naming arises because we count all three waves (including the generated one)

[8] See Boyd [3] section 1.5.10 for explanation.

same refractive index for all three coordinate axes) and so have $\chi^{(2)} = 0$. Thus, we need to use an **anisotropic medium**, where the refractive indices of the three different coordinate axes are not all equivalent. These properties are readily found in **nonlinear crystals**, which are crystals made of a material that is optimised for a given nonlinear effect. Common nonlinear crystals you may encounter include β *Barium BOrate* (BBO), *Potassium Dihydrogen Phosphate* (KDP), and *Lithium Niobate*. We will discuss crystals further later.

So, given the right crystal, how do we efficiently perform our frequency mixing to make a lot of second harmonic light? For any kind of nonlinear frequency mixing, the bottom line is that we have to **combine the photons in a way that conserves momentum and energy**. We will illustrate this using the example of second-harmonic generation (SHG).

Second-Harmonic Generation

To combine the photons in a way that conserves momentum and energy, we require:

$$\omega_1 + \omega_2 = \omega_3 \tag{3.11}$$

For energy conservation, since the photon energy E is given by $E = \hbar\omega$, and:

$$\mathbf{k}_1 + \mathbf{k}_2 = \mathbf{k}_3 \tag{3.12}$$

Where $\mathbf{k_i}$ is the **wave vector** of wave i. Equation 3.12 comes from momentum conservation because the momentum of a wave, p_i, can be written as $p_i = \hbar k_i$ where k_i is the magnitude of the wave vector $\mathbf{k_i}$. Equation 3.12 is known as the **phase-matching condition** (we will see why shortly), and states that the vector sum of the momenta of the two input photons must be equal to the momentum of the output photon. Equation 3.12 is often rewritten in terms of the **wave vector mismatch**, $\Delta\mathbf{k}$:

$$\Delta\mathbf{k} = \mathbf{k}_3 - \mathbf{k}_2 - \mathbf{k}_1 \tag{3.13}$$

Where the wave vector mismatch $\Delta\mathbf{k}$ quantifies the degree of phase matching - in the ideal case, $\Delta\mathbf{k} = 0$. If this condition is satisfied, then we can effectively add our input photons together to produce output photons at double the frequency.

If we are using collinear beams (e.g. if ω_1 and ω_2 come from the same beam), we can cancel out the angular dependence of the k-vectors and Equation 3.12 reduces to:

$$k_1 + k_2 = k_3 \tag{3.14}$$

Where k_1 is the magnitude of $\mathbf{k_1}$ and so on. To understand how this relates to material properties, note that the magnitude of the wave vector in medium i, k_i can be written in terms of the refractive index as follows:

$$k_i = \frac{\omega_i}{v_{\mathrm{p},i}} = \frac{n_i\omega_i}{c} \tag{3.15}$$

We saw a similar construction in Equation 2.16. In Equation 3.15, n_i is the refractive index of frequency ω_i in the medium, and $v_{\mathrm{p},i}$ is the phase velocity of the wave in the medium.

The phase velocity can be expressed in terms of the refractive index and the speed of light c, giving the final form shown in Equation 3.15. We can now use the relationship between k_i and n_i to write Equation 3.12 in terms of refractive indices as:

$$n_1\omega_1 + n_2\omega_2 = n_3\omega_3 \tag{3.16}$$

Which in the case of second-harmonic generation reduces to:

$$n_\omega = n_{2\omega} \tag{3.17}$$

on substitution of $n_1 = n_2 = n_\omega$ and $n_3 = n_{2\omega}$, with $\omega_1 = \omega_2 = \omega$ and $\omega_3 = \omega_1 + \omega_2 = 2\omega$. Equation 3.17 shows that to have effective SHG, we need a material in which the refractive index for the fundamental n_ω is the same as the refractive index for the second-harmonic $n_{2\omega}$.

We can get an intuitive feel for this process (and explain why it is called *phase matching*) by considering what happens to waves when they are mixed together in a nonlinear crystals:

- When Equation 3.17 is satisfied, any produced second harmonic travels with the same *phase velocity* as the input fundamental due to them having equal refractive indices.
- Therefore, any second harmonic photons created by the fundamental at the start of the crystal will be in phase with any second harmonic created by the fundamental at the end of the crystal, and at all points in between.
- These second harmonic waves then constructively interfere to give the maximum possible second harmonic output, as shown in Figure 3.4.

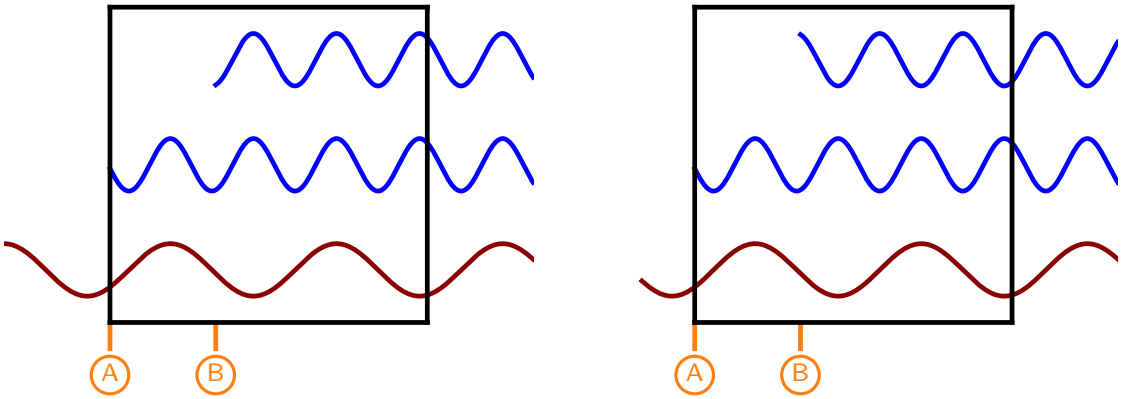
This matching of phase velocities is why the process is called *phase matching*.

Previously, we have seen that for most materials the refractive index depends on the frequency of the light (Figure 2.1) - so it seems that finding a material where $n_\omega = n_{2\omega}$ would be difficult. Luckily, there are ways to achieve this, and one way is to use a **birefringent crystal**.

Birefringent Phase Matching

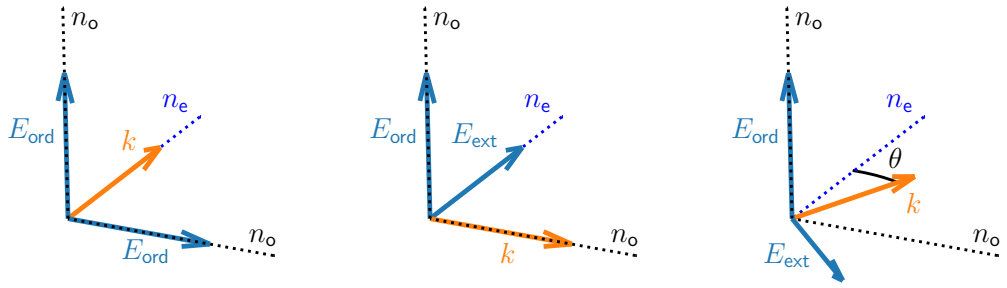
A crystal that is **birefringent** has a refractive index which depends on the polarisation and propagation direction of the light incident on it. A 3D crystal will have three refractive indices, n_x, n_y, and n_z. If these are all the same then the crystal is **isotropic**, and is not birefringent. If one of these is different from the other two, then the crystal is **uniaxial**. If they are all different, then the crystal is **biaxial**. Birefringence can be exploited to allow us to achieve phase matching, and we will now consider collinear phase matching in uniaxial crystals as an illustrative case.

In a uniaxial crystal, there is one axis which has a different refractive index from the other two, this axis is called the **optical axis** of the crystal, and is shown in dashed blue on Figure 3.5. A wave polarised along this axis would experience a refractive index n_e, whereas



(a) (b)

Figure 3.4: Simplified illustration of the need for phase matching. In the phase matched case (a) the second harmonic waves (blue) generated at points A and B by the fundamental (dark red) are in phase and constructively interfere, producing intense second harmonic output. In the non-phase matched case (b), the second harmonic wave generated at B is not in phase with that created at A, leading to destructive interference and reduced output. Note that the phase evolution of the fundamental is not presented here.



(a) Propagation along the optical axis. (b) Propagation orthogonal to the optical axis. (c) Propagation at an angle θ to the optical axis.

Figure 3.5: Propagation of light in a uniaxial crystal with the optical axis (OA - n_e) in dashed blue and the other two axes (n_o) in dashed black. The k-vector (orange) defines the propagation direction, and the light can either be polarised in the plane spanned by the k-vector and the OA (extraordinary polarisation - E_{ext}), or orthogonal to this plane (ordinary polarisation - E_{ord}). Panels (a)-(c) show different propagation directions within the plane spanned by the k-vector and OA.

a wave polarised along either other axis would experience a refractive index n_o. If $n_e > n_o$, then the crystal is said to be **positive uniaxial**, and if $n_o > n_e$ it is said to be **negative**

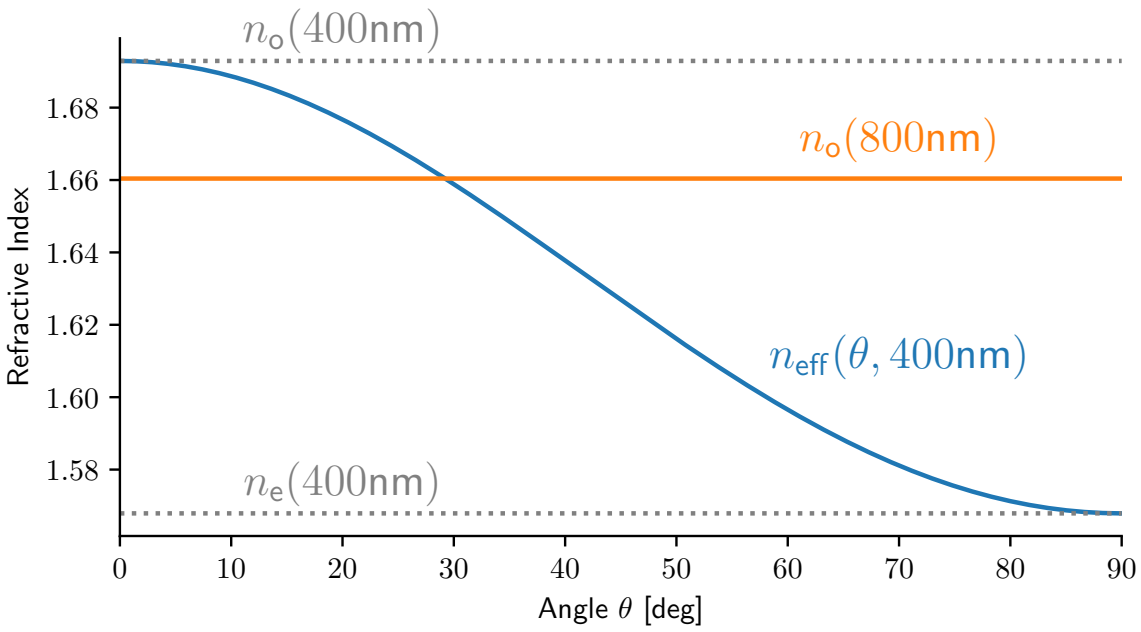


Figure 3.6: Plot of Equation 3.18 showing how by tuning the angle θ of a BBO crystal the refractive index n_{eff} (blue) experienced by 400nm light with extraordinary polarisation can be made to coincide with the refractive index n_o(orange) experienced by the 800nm fundamental with ordinary polarisation. n_{eff} varies between n_o and n_e for 400nm light, shown with dashed grey lines. Phase matching would be achieved at the angle where the curves intersect.

uniaxial. A wave propagates through the crystal in a direction defined by the direction of the wave vector, **k**, shown in orange on Figure 3.5. The polarisation plane of this wave is orthogonal to the wave vector. If the wave is linearly polarised and the polarisation direction is orthogonal to the plane containing **k** and n_e, then it is known as the **ordinary ray**, with polarisation direction E_{ord}. Conversely, if the polarisation direction lies in the plane containing **k** and n_e, then it is known as the **extraordinary ray**, with polarisation direction E_{ext}.

The ordinary ray always experiences a refractive index of n_o, regardless of the propagation direction. The extraordinary ray experiences a refractive index between n_o and n_e, depending on the angle θ between the wave vector and the OA. Panels (a)-(c) in Figure 3.5 illustrate this for the two limiting cases where **k** coincides with either n_e (a) or n_o (b), and for the case of an arbitrary angle θ (c). The effective refractive index, $n_{eff}(\theta)$ that the extraordinary ray experiences is given by:

$$\frac{1}{n_{eff}(\theta)^2} = \frac{\cos^2(\theta)}{n_o^2} + \frac{\sin^2(\theta)}{n_e^2} \tag{3.18}$$

Equation 3.18 is plotted in Figure 3.6 for the case of SHG of 800 nm light in a BBO crystal. To achieve birefringent phase matching, we need to angle our crystal in such a way that

the refractive indices experienced by both our fundamental (n_ω) and second harmonic ($n_{2\omega}$) are equal. Figure 3.6 shows that this is possible at an angle of θ =29.2° provided that the fundamental 800 nm has ordinary polarisation and the generated 400 nm has extraordinary polarisation. Turning the crystal to this angle will therefore match the refractive indices and yield efficient SHG. There are some important points to emphasise regarding this process now we have seen a basic example.

Polarisation Effects

Firstly, we have seen that in this case, to achieve phase matching, our output wave must have different polarisation to our input wave. The case where our two input waves (here both 800 nm waves from the same source) have the same polarisation and the output has different polarisation is called **Type-I** phase matching. If the two input waves have different polarisation, it is called **Type-II** phase matching. If all waves have the same polarisation, it is called **Type-0** phase matching. These schemes are summarised in Table 3.2. It is not possible to achieve Type-0 phase matching using a birefringent crystal as described here - this requires use of a technique known as *quasi-phase matching*. Further detail on this can be found in [4, 3].

Name	Wave polarisation (input/input/output)
Type-0	ooo, eee
Type-I	ooe, eeo
Type-II	oeo, oee, eoe, eoo

Table 3.2: Different polarisation schemes for achieving phase matching. o = ordinary, e = extraordinary.

Phase-Matching Bandwidth and Efficiency

Another consideration relates to the concept of the **phase-matching bandwidth**. As we are using broadband ultrashort pulses, it is important that we can achieve phase matching over the whole pulse bandwidth so that our pulses remain ultrashort after the frequency mixing. For example, if we start with an 1030 nm pulse with a bandwidth of 10 nm, we'd expect to see colours ranging from 1025-1035 nm in the spectrum of our pulse. If we frequency doubled this pulse to 515 nm, we'd expect to end up with a bandwidth of 2.5 nm (taking into account the apparent reduction in bandwidth caused by working in wavelength space[9] seen in Equation 1.4). So really we need to satisfy Equation 3.17 across all these colours.

Yet as plots of the refractive index for common materials (Figure 2.1) show, refractive indices vary with frequency and so Equation 3.17 can only strictly be realised for one specific fundamental and one specific second harmonic frequency. Every possible combination of fundamental and second harmonic frequency present in our ultrashort pulse will not

[9]Which is one of the main reasons why working in wavelength space is silly, but everyone seems to do it so you need to be used to it. For a more thorough exploration of *why* this is silly, see reference [5]

have Equation 3.17 satisfied. Furthermore, the phase matching bandwidth $\Delta\omega$ scales with interaction length L as:

$$\Delta\omega \propto \frac{1}{L} \tag{3.19}$$

So a long crystal will only phase match a narrow bandwidth. This is problematic because the amount of SHG light we make scales as the square of the interaction length (Figure 3.7 (a)), so to produce a lot of light we want a long interaction length[10] – but a long interaction length will cause our phase-matching bandwidth to drop. A reduction in phase-matching bandwidth will mean that we can't frequency mix the entire bandwidth of our ultrashort pulse – which will ultimately narrow the bandwidth of our output pulse, with a corresponding increase in the pulse duration[11]. Long crystals will also make the angle-tuning of the crystal much more sensitive.

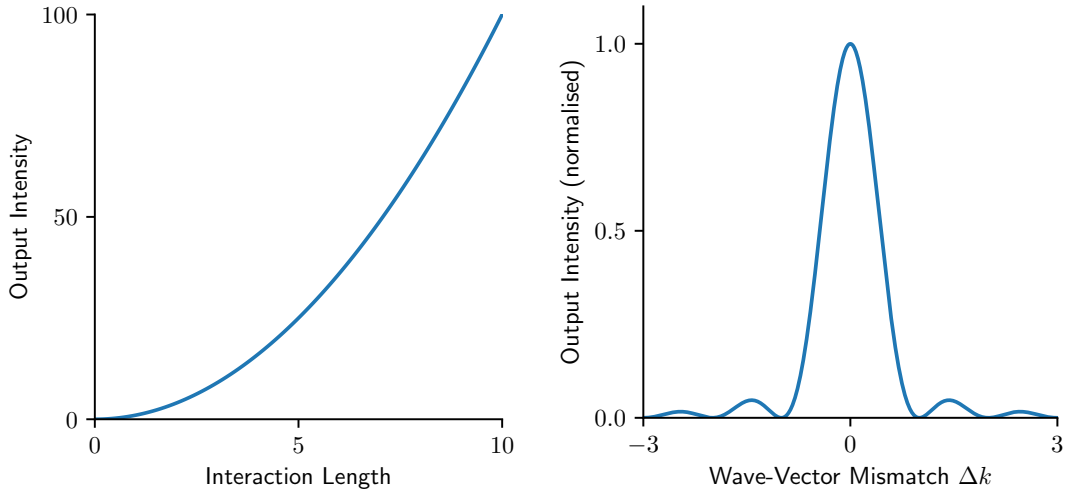


Figure 3.7: Output intensity from nonlinear frequency mixing as a function of interaction length in the crystal (a) and wave-vector mismatch (b).

Finally, it is also useful to have an awareness of how the efficiency of the frequency mixing depends on Δk – shown in Figure 3.7 (b). Clearly, the efficiency is maximised when $\Delta k = 0$ and falls away rapidly as it is tuned away from this optimum. To summarise this section, we can take away the following key messages:

- Thick crystals produce lots of light, at the expense of phase-matching bandwidth. They are more sensitive to angle and input frequency, and are more robust.

- Thin crystals will more completely frequency-convert your ultrashort pulses, retaining their whole bandwidth after conversion. They are less sensitive to angle but are more fragile.

[10] To a point – at some point a long interaction length stops being useful because the wave created gets out of phase with the polarisation that's generating it, but this is a technical detail.

[11] There are cases where this spectral narrowing can actually be helpful though - see reference [6] for a spectroscopic example.

Other ways to increase the phase-matching bandwidth involve using non-collinear beams, in which case Equation 3.12 turns back into a vector equation, and the angle between the incoming beams becomes an additional parameter that can be adjusted to optimise the characteristics of the output light. Further detail and more mathematical treatment of all these concepts can be found in [4, 3, 2, 7], and the Optics Toolbox [8] contains many useful calculators for things like phase-matching angles.

Temperature Tuning and Walk-Off

A drawback of angle-tuning (critical phase matching) is a phenomenon known as **spatial walk-off**. Spatial walk-off occurs because in an anisotropic medium (like a birefringent nonlinear crystal) the **direction of wave propagation** and the **direction of energy flow** of the laser field are not necessarily the same[12]. Walk-off limits the conversion efficiency in critical phase matching because it causes the beams with ordinary and extraordinary polarisation to "walk off" from each other and lose spatial overlap, which stops there being efficient frequency mixing.

To remedy this, we can instead do something called **non-critical phase matching**[13]. This is a technique where the refractive index matching is achieved via temperature tuning. Some crystals have a temperature dependent refractive index along certain axes, and so the crystal can be housed in an oven and the temperature tuned to match the relevant refractive indices for efficient phase-matching[7]. The advantage of this technique is that precise angle tuning is unnecessary and spatial walk-off can be avoided. The disadvantage is that the temperatures required are generally on the order of 100 °C and so require that the crystal is housed in an oven. Non-critical phase matching is commonly found inside high power pump lasers.

3.2.2 Selected Uses of Frequency Mixing

Here we briefly summarise some uses of the fundamental frequency mixing technique discussed in the preceding section, with an emphasis on those that might be encountered by the intended reader of this book.

Higher Order Mixing

All of our discussion so far has focussed on three-wave mixing, but higher-order mixing processes are also commonly encountered. The logical next extension to this is **four-wave mixing** (FWM) where there are three input waves and one output, a process that occurs through the third-order susceptibility $\chi^{(3)}$. Accordingly, FWM does not have to happen in an anisotropic medium. The range of possible interactions in FWM is much greater than in three-wave mixing, but an obvious first example would be in **third harmonic generation**[14]

[12] See Figure 8.8, p 369 in reference [2] for an explanation.

[13] The terminology originates because in temperature tuning the relative alignment of the beams and the crystal angles is less sensitive, or 'critical', than in critical phase matching.

[14] This could also be achieved in a three-wave mixing process: $\omega_1 + \omega_2$ where $\omega_2 = 2\omega_1$.

where the output frequency ω_4 is given by:

$$\omega_4 = \omega_1 + \omega_1 + \omega_1 \tag{3.20}$$

Thus producing an output at three times the input frequency. A common use of FWM is to produce laser pulses with frequencies towards the deep UV or mid IR. It is intuitive that adding together more frequencies gives the possibility of higher energy outputs (see reference [9]), and tuning the precise phase-matching can also produce lower energy output via difference frequency generation (see reference [10]). An introduction to how FWM techniques can be applied spectroscopically can be found in reference [11]

Once you extend to FWM, the sky is (almost) the limit with nonlinear frequency mixing. Techniques such as **high-harmonic generation** exist where the frequency of input laser pulses can be multiplied by a factor of 10 or 20 – these find use in the generation of attosecond laser pulses, and in the generation of tabletop extreme ultraviolet (XUV) light sources. These methods are beyond the scope of this text, but accessible places to start are in references [12] and [13]. At the other end of the frequency spectrum, you can also generate very low photon energies by a process called **optical rectification**. This finds a use in generation of *terahertz* radiation[15]. Optical rectification is essentially the case where you have DFG that creates a polarisation that varies very slowly (see Appendix D for some mathematics), so produces a very low frequency wave. More information can be found in reference [14].

Second Harmonic Bandwidth Compression

Another interesting use of second harmonic generation is in the production of narrowband picosecond pulses from broadband femtosecond ones. It may seem odd that this is something people want to do, but actually many kinds of spectroscopy benefit from the increased frequency resolution a narrowband pulse provides, and instruments for doing this bandwidth compression are relatively common in experimental setups in spectroscopy.

Bandwidth compression can be accomplished in relatively low-tech ways such as spectrally filtering the laser pulse[16] (Figure 3.8). Methods like these are conceptually simple but very inefficient. A much more elegant way to spectrally narrow a broadband femtosecond pulse is via a technique called **second-harmonic bandwidth compression** (SHBC).

SHBC works by second harmonic generation, but where the two pulses being combined are oppositely and strongly chirped. Simplistically, if we imagine that the pulses are linearly and oppositely chirped, then their frequencies ω^+ and ω^- are given by:

$$\omega^{\pm} = \omega_0 \pm \Delta\omega \tag{3.21}$$

[15] To us, 'terahertz' means really low frequency infrared light. To HAM radio geeks, it means really high frequency radio waves. It's used in spectroscopy of materials, among many other things.

[16] Literally just throwing away the frequencies in the pulse you don't want, which is as wasteful as it sounds.

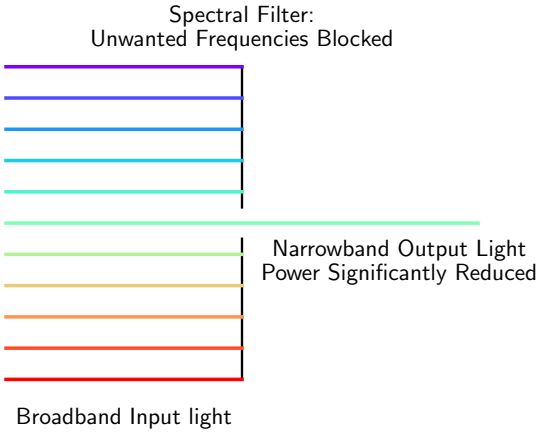


Figure 3.8: Spectral filtering: dispersing your broadband pulse into its constituent frequencies, and blocking the ones you don't want.

Where ω_0 is the central frequency and $\Delta\omega$ is the bandwidth. Sum-frequency mixing of these two pulses gives an output pulse with frequency ω_{SHBC}:

$$\omega_{SHBC} = (\omega_0 + \Delta\omega) + (\omega_0 - \Delta\omega) = 2\omega_0 \tag{3.22}$$

Hence, we have made a narrowband pulse at twice the input frequency. In reality the pulses are not perfectly linearly chirped and we have to take into account phase-matching bandwidth considerations, but conversion efficiencies in excess of 40% are possible. This technique is commonly used to generate high power, narrowband, picosecond pulses from femtosecond laser output.

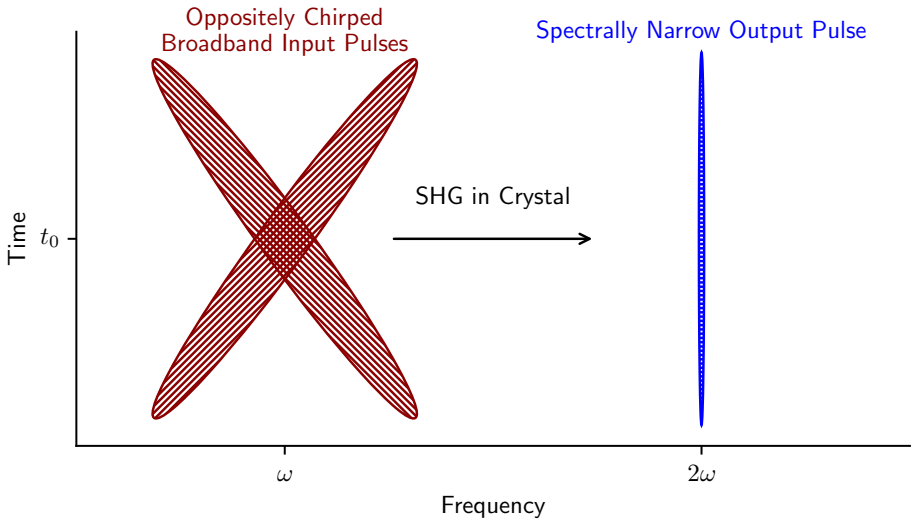


Figure 3.9: Second Harmonic Bandwidth Compression (SHBC). Two oppositely chirped broadband pulses are mixed in a nonlinear crystal. The produced SHG is both narrowband and preserves much of the power from the input pulses.

3.3 Optical Parametric Amplification

A crucial nonlinear process that almost all ultrafast laser users will encounter is **optical parametric amplification**. The development of this process has been a boon to many non-specialists (like us) who want to use lots of different colours of light in our experiments, but want to generate them all from one driving femtosecond laser[17]. At its core, an optical parametric amplifier (OPA) is just doing the frequency mixing that we have discussed at length in the previous section, but generally over a wider range of frequencies and in a way that is tunable.

Using a modern commercial OPA (which are almost invariably made by the Lithuanian[18] company Light Conversion) we can simply type the wavelength of the light we want into a computer, and motors move all the crystals and mirrors around inside the OPA so that the desired wavelengths are produced, which is very convenient. In this section, we will explore some of the key aspects of OPAs that are important to be aware of. However - on first reading for a complete novice, this section could certainly be skimmed over.

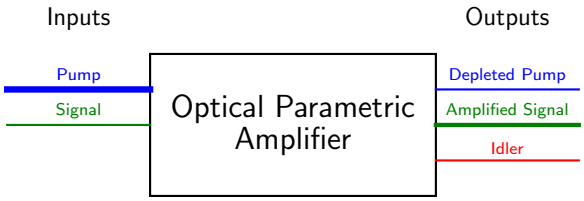


Figure 3.10: Schematic of the OPA process. A pump beam is depleted, while a signal and idler beam are amplified/generated.

3.3.1 Fundamentals

The fundamental idea of an OPA is that a strong **pump beam** with frequency ω_p amplifies a weaker **signal beam** with frequency ω_s via mixing in a nonlinear medium. Energy and momentum conservation during this process requires that a third beam, the **idler beam** is created, with frequency given by $\omega_i = \omega_p - \omega_s$. Rewritten, this equation is the central one that governs the OPA process:

$$\omega_p = \omega_s + \omega_i \tag{3.23}$$

This is illustrated schematically Figure 3.10 and as an energy level diagram in Figure 3.11. Equation 3.23 shows that we can think of the OPA process as splitting the pump photon in half, creating two new photons: the signal and idler. As the process is tunable ω_s can then

[17] So we can synchronise the pulses with ultrafast timing.

[18] Lithuanian companies dominate a lot of the laser and optical industry, for good reason – *'trust the Lithuanians'* is advice I was always given when buying optics and optomechanics.

vary from 0 up to ω_p whilst ω_i varies from ω_p down to 0. The *degeneracy point* occurs when $\omega_s = \omega_i$, which occurs at half the pump frequency.

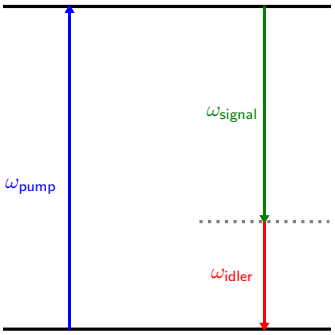


Figure 3.11: Schematic of the optical parametric amplification (OPA) process.

The beams are overlapped both temporally and spatially in a nonlinear crystal such as BBO to perform the process depicted in Figure 3.11. The OPA process needs to be phase-matched like the frequency mixing discussed previously to be efficient. It's useful to look at the OPA process in more depth to understand how commercial OPAs that you may find in a lab are constructed (and because building OPAs, or small parts of them, is a common thing to do).

Fundamentally, to make an OPA we require two things:

1. A source of a weak signal beam that we can amplify using a pump beam.

2. One or more amplification stages to amplify the weak signal beam up to a useful power level.

All of these stages are generally contained within a single box in a commercial system, and are driven by the output from a femtosecond laser system. In applications like ultrafast spectroscopy, it is important that the same driving laser drives the whole OPA (and other parts of an experimental setup) to maintain the ultrafast time resolution[19] OPAs seem like very complex devices, but the underlying physics at a simple level is quite intuitive. We will discuss key aspects here, but a much more extensive accessible introduction to how ultrafast OPAs work (and can be designed) is found in [4], and other useful discussion of various concepts can be found in [3, 1, 7].

3.3.2 Seed Generation

The most common way in which a seed pulse is generated for amplification in an OPA is via **white-light generation** (WLG). WLG is a complex process, but simplistically it amounts to focussing an ultrashort laser pulse into a material where a range of nonlinear interactions

[19] Generally one big high-power laser 'engine' will drive an OPA and other experimental apparatus. Commonly a single laser can provide the power for several experimental setups simultaneously.

happen that result in the pulse bandwidth broadening significantly. The bandwidth broadens so much that the output light looks 'white', hence the name.

A key driver of this process is self-phase modulation, as discussed earlier in this chapter, but an array other effects can also present. The range of colours present in the white light continuum (WLC) dictates the range of colours that the eventual OPA can produce (if they can all be amplified), and the range of colours in the white light is essentially dictated by two things:

- The central wavelength of the light used to drive the WLG process.
- The material in which the WLG is performed.

Driving Wavelength

At a basic level, WLG results in creation of new colours around the central wavelength of a pulse, so changing the central wavelength results in different colours being produced. Generally a small fraction of the pump beam (ω_p) is used to drive the WLG process - which will typically have a wavelength of 1030 nm or 800 nm for Yb:KGW or Ti:Sa systems respectively. These will lead to a WLC centered in on the red side of the visible spectrum. Alternatively, a common option is to generate the WLC with the second harmonic of the pump (515 nm or 400 nm), which pushes the WLC towards blue wavelengths. Generating WLC into the UV region (<350 nm) becomes incresingly challenging due to absorption of the pump pulse by the material used to generate the WLC.

Material Choice

The choice of material for generation of a WLC affects both the bandwidth and stability of the WLC[20]. Intense pump light is focussed hard into the material and so avoidance of damage is a key consideration – the widespread use of sapphire is due to its high damage threshold. More fragile materials such as CaF_2 or LiF can be used to extend the bandwidth of the WLC closer to the UV, but come at the cost of increased potential for damage and reduced stability of the resulting continuum. It is also possible to do WLG in liquid samples, but again instability is a key issue. The majority of OPAs you will encounter will use a sapphire plate to generate the WLC.

3.3.3 Amplification

Once a seed has been generated, it can be amplified[21]. The amplification process is the core of the OPA and it is worth discussing in some detail - a schematic of the process was shown in Figure 3.10.

[20] See reference [15] for a detailed review, and Fig 10 within it for a comparison of different materials.

[21] It is also possible to *not* use a specific seed and instead amplify vacuum fluctuations – this is called **optical parametric generation** or **parametric superfluorescence**. This technique is used when the driving laser produces very low pulse energy, but produces an unstable seed that generally makes the rest of the OPA difficult to work with.

Amplification Bandwidth

The first thing to consider is the range of frequencies that the OPA can effectively amplify. There are a few 'hard limits' here:

- We can't amplify anything that is at a higher frequency than the pump beam.
- We can't amplify anything that we can't generate a seed for via WLG.
- We can't amplify anything outside the transparency range of the crystal used in the OPA (for most common crystals this means any wavelength shorter than 350 nm and longer than 3 µm).

Furthermore, OPA is a phase matched process, so we also have to consider the properties of the exact nonlinear crystal being used, as this will determine the range of angles for which we can get effective phase matching and thus effective amplification. An OPA amplification stage can either use collinear or non-collinear pump and seed beams: collinear beams can give a higher conversion efficiency at the expense of amplified bandwidth, and non-collinear beams allow a much wider range of frequencies to be effectively phase-matched (as the angle between the beams can be tuned, in addition to the crystal angle), but at the expense of a slightly reduced efficiency due to the less effective spatial overlap.

Generally, OPAs are better able to amplify long wavelengths than short ones (due to strong group velocity mismatch at shorter wavelengths – see below). By 'long wavelengths' here we mean wavelengths in the near-IR range, around 1000-2000 nm. As such, if longer or shorter wavelengths than this are desired[22] then the usual strategy is to **frequency-mix the OPA output**. Almost continuous tuning in the UV-Vis region can be achieved by taking second/third harmonics of the signal or idler, or by sum- and difference-frequency mixing of the pump with the signal or idler (and possibly taking harmonics of the result).

At the time of writing, the company Light Conversion dominates the OPA market and they produce OPAs designed for both Ti:Sa and Yb lasers – a look at their website will show lots of tuning curves for various OPAs, and the interactions used to generate them[23]. Using the wavelength extension boxes means that the OPA system can cover a range from 190 nm up to 20 µm – deep in the UV into the far infrared.

Amplification Efficiency

In contrast to simple sum- or difference-frequency generation, OPA is a very efficient process. The underlying physical reasons for this are explained in depth in [4], but fundamentally the reason is that in OPA, the *gain*[24] of the OPA scales **exponentially with the length**

[22] Which they commonly are for spectroscopy

[23] These can be bewildering if you look at them as a novice. To save you a few minutes of puzzling out the acronyms: (S/T/F)H(S/I) = Second/Third/Fourth Harmonic of Signal/Idler; SF(S/I) = Sum of Fundamental and Signal/Idler; SHSF(S/I) = Second Harmonic of Sum of Fundamental and Signal/Idler; DFG1/2 = Difference Frequency Generation in Crystal 1/2. Other acronyms are all derived from these.

[24] The ratio of the output power to the input power.

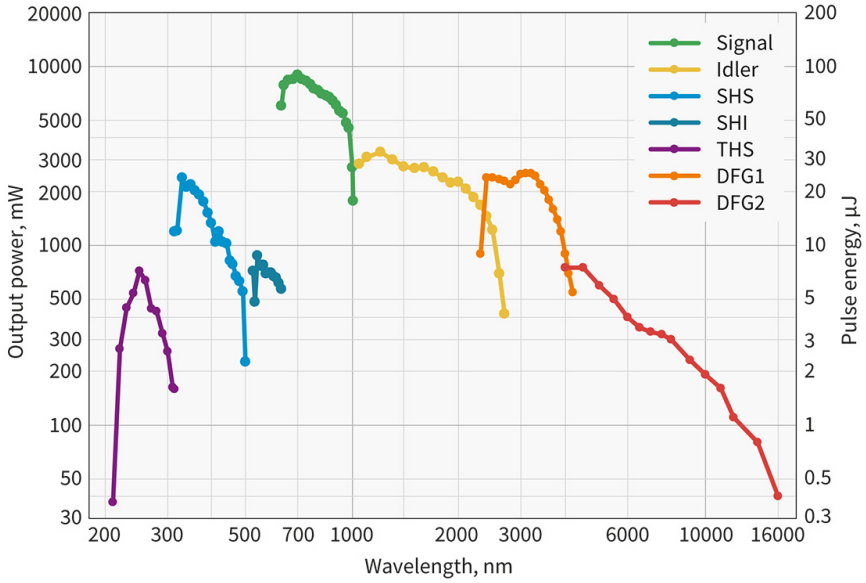


Figure 3.12: Tuning curves for an the Light Conversion ORPHEUS-NEO pumped at 1030 nm. Figure courtesy of Light Converison.

of the nonlinear crystal being used (Figure 3.13 (a)). This is in contrast to SHG, where we saw that the efficiency scales with the square of the crystal length (Figure 3.7). So, is it best to just use the longest crystal possible?

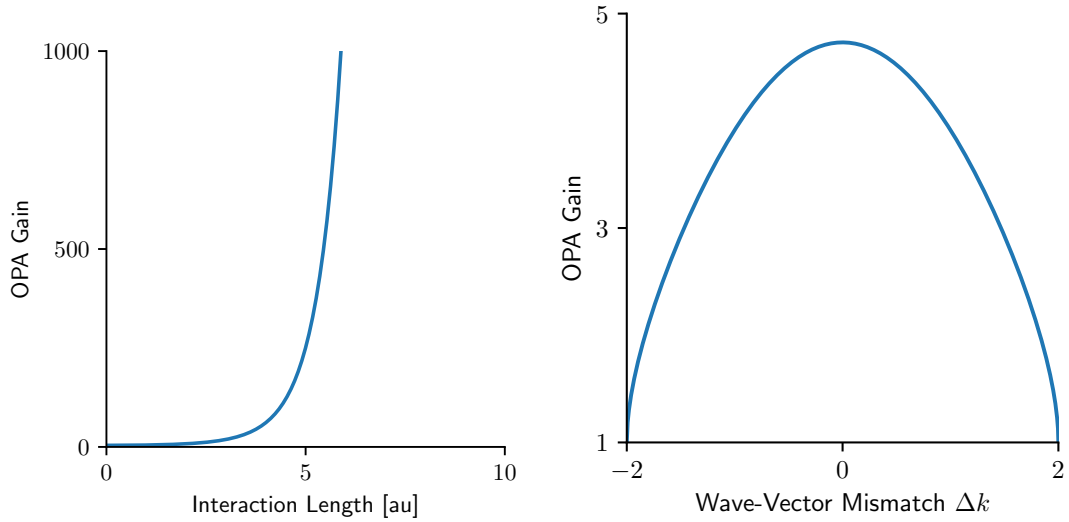


Figure 3.13: OPA gain as a function of interaction length in the crystal (a) and wave-vector mismatch (b). The gain scales exponentially with interaction length, unlike the SHG case shown in Figure 3.7. Simulated using idealised parameters from equations in reference [4].

The answer is no, because we need to also think about the nature of the pulses we are amplifying in the OPA crystal. The pulses must be overlapped in time and space, and

because the pulses are different colours they will travel at different speeds through the crystal. This **group velocity mismatch** (GVM) will cause a phenomenon called **temporal walk off** where after a certain propagation length the two pulses will lose their temporal overlap. Angle-tuning (in a non-collinear scheme) can mitigate this to some extent.

Finally, it is important to remember that OPA is still a phase-matched process, any wave-vector mismatch will cause the efficiency to drop precipitously (Figure 3.13 (b)) in a similar way that it does for harmonic generation (Figure 3.7 (b)). The wave-vector mismatch ultimately limits the phase-matching bandwidth (and therefore bandwidth that can be amplified for a given crystal angle). Again, angle-tuning in a non-collinear scheme can mitigate the effects of this and allow a very broad range of wavelengths to be amplified at the same crystal angle[25].

3.3.4 OPA Miscellany

Included here is some miscellaneous information about OPAs and the OPA process which may be instructive before we move to the next chapter – as before, this could certainly be skipped on a first reading.

OPA vs Four Level Laser

It is interesting to think about the similarities and differences between the OPA process (Figure 4.10) and the prototypical four level laser amplification process we discussed at length in earlier in this book (Figure 1.1).

The similarities are that we still *deplete* a pump beam whilst *amplifying* a weaker signal beam, but a crucial difference between an OPA and a four-level laser is that **there is no energy stored within the medium the OPA is performed in**. This is in stark contrast to the four level laser, in which we were dumping huge energy into a crystal and then extracting part of it and relying on fast relaxation between levels to maintain our population inversion. In the four-level case, we would have had to cool the crystal significantly to prevent damage. In contrast, energy conservation means no energy is actually stored in the nonlinear crystal we use for optical parametric amplification – so cooling is not necessary and high input intensities can be used without damaging the crystal. Another key difference is that an OPA can be engineered to have a much broader amplification bandwidth than a typical four-level laser amplifier, as we do not rely on the intrinsic emission spectrum of a material to create the bandwidth.

A good way to think of this difference as the OPA just providing a 'medium' that facilitates mixing of multiple beams, whereas a laser crystal in a four-level laser is more involved, as one beam puts energy into the crystal which is picked up by another. It may sound like this makes OPA a 'better' technique – but that is comparing chalk with cheese: we need to have a powerful femtosecond laser to drive the OPA with in the first place!

[25] See, for example, Fig 10 in reference [4]

NOPA

Earlier we mentioned that very broadband amplification can be achieved using an OPA in a non-collinear geometry. A Non-collinear OPA is called a **NOPA**, and these are reasonably common as homemade sources in ultrafast laser labs as they are relatively easy to align, and as such are an easy and cost-effective way to make tunable visible region pulses.

A NOPA can be pumped by the fundamental or second harmonic output of any typical femtosecond laser (Ti:Sa or Yb-based), and consists of a WLG stage and a single non-collinear amplification stage. Details on how a NOPA can be designed are found in references [4] and [16][26]. Whilst a homemade NOPA will not necessarily have the simple 'type-in-a-number-and-hit-go' tunability of a commercial OPA system, they are a lot more budget-friendly, and building a NOPA is certainly not beyond the ability of an enthusiastic non-specialist[27]

Bibliography

[1] Hooker S and Webb C. *Laser Physics*. Oxford University Press, Oxford, 1st edition, 2010.

[2] Milonni P W and Eberly J H. *Laser Physics*. Wiley, New York, 1st edition, 2010.

[3] Boyd R. *Nonlinear Optics*. Elsevier, Amsterdam, Netherlands, 4th edition, 2020.

[4] Manzoni C and Cerullo G. Design criteria for ultrafast optical parametric amplifiers. *J Opt*, 18, 2016.

[5] Pickering J D. Wavelength considered harmful. https://jamesdpickering.com/pdfs/wavelength_considered_harmful.pdf. Accessed: 23-07-2024.

[6] Weeraman C, Mitchell S A, Lausten R, Johnston L J, and Stolow A. Vibrational sum frequency generation spectroscopy using inverted visible pulses. *Opt. Express.*, 18, 2010.

[7] Paschotta R P. The rp photonics encyclopedia. https://www.rp-photonics.com/encyclopedia.html. Accessed: 22-12-2020.

[8] Light Conversion. Optics toolbox. toolbox.lightcon.com. Accessed: 23-12-2020.

[9] Beutler M, Ghotbi M, Noack F, and Hertel I V. High power red and near-ir generation using four wave mixing in all integrated fibre laser systems. *Optics Letters*, 35, 2010.

[26] A literal step-by-step laboratory build guide can also be found in reference [17], and further information from Eberhard Riedle's site at reference [18].

[27] And would be a fun project.

[10] Lavoute L, Knight J C, Dupriez P, and Wadsworth W J. High power red and near-ir generation using four wave mixing in all integrated fibre laser systems. *Optics Express*, 18, 2010.

[11] Wright J C, Carlson R J, Hurst G B, Steehler J K, Riebe M T, Price B B, Nguyen D C, and Lee S H. Molecular, multiresonant coherent four-wave mixing spectroscopy. *Int. Rev. Phys. Chem.*, 10, 1991.

[12] C M Heyl, C L Arnold, A Couairon, and L'Huillier A. Introduction to macroscopic power scaling principles for high-order harmonic generation. *J. Phys. B: At. Mol. Opt. Phys.*, 50, 2016.

[13] Krausz F and Ivanov M. Attosecond physics. *Reviews of Modern Physics*, 81, 2009.

[14] Hebling J, Yeh K-L, Hoffman M C, and Nelson K A. High-power thz generation, thz nonlinear optics, and thz nonlinear spectroscopy. *IEEE Journal of Selected Topics in Quantum Electronics*, 14, 2008.

[15] Dubietis A, Tamosauskas G, Suminas R, Jukna V, and Couairon A. Ultrafast supercontinuum generation in bulk condensed media. *Lithuanian Journal of Physics*, 57, 2017.

[16] E Riedle, M Beutter, S Lochbrunnner, J Piel, S Schenkl, S Sporlein, and W Zinth. Generation of 10 to 50fs pulses tunable through all of the visible and the nir. *Appl. Phys. B.*, 71, 2000.

[17] Wilson K and McCamant D. Guide to nopas. http://www.sas.rochester.edu/chm/groups/mccamant/assets/pdf/documents_nopa.pdf. Accessed: 23-07-2024.

[18] E Riedle. Nopa - overview and principles. http://nopa-agriedle.info/. Accessed: 06-10-2024.

Part II

Practical Ultrafast Optics

Chapter 4

Generating Ultrashort Pulses

In this chapter, we turn to a discussion of how ultrashort pulses are generated. We start with a brief discussion of general laser oscillator/amplifier concepts, and then go on to discuss the central ideas of modelocking and chirped-pulse amplification. We also discuss the optical parametric amplification process in some detail. Both Titanium Sapphire and Ytterbium gain media are discussed.

At this point, we have covered most of the relevant fundamental physics that we'll need. Before we move to the practical optics, it will be useful to have a rough idea of how an ultrafast laser actually works – which is the subject of this chapter.

4.1 Fundamentals

We know from chapter 1 that we need a **gain medium** in our laser, which is the medium that contains the energy levels used to achieve laser operation. Summarised in Figure 1.1 for a typical four-level laser, we saw that we need to **pump** this gain medium to excite it, such that lasing can be achieved. In this way, a simple building block of any laser system is a gain medium being pumped by another energy source to produce laser output.

The simplest lasers will essentially consist of this building block alone. For example, a laser pointer might just be a **laser diode** in a pen-shaped package. A laser diode is a semiconductor that will produce a laser beam when driven by an electrical current, so in this case the pumping energy is provided by an electric power supply. Other lasers may use an optical (but non-laser) energy source as the pump. The first laser was built in 1960 by Theodore Maiman[1], and consisted of a ruby crystal pumped by a **flashlamp**. A flashlamp is essentially a very bright lamp where an electric arc is produced that gives off a lot of bright white light. Flashlamps will be familiar to you from things like camera flashguns or neon signage, but they are also widely used to pump laser gain media. Ultrafast lasers all use *other lasers* as their source of pump energy.

In this way, ultrafast lasers are generally best regarded as **laser systems**. They do not

consist of simply one gain medium that is excited by an electric current or a flashlamp, but instead the combine multiple different lasers together (into the **laser system**) to produce the final ultrafast laser output. In this chapter we will focus on the key components and construction of a typical ultrafast laser system that are needed to facilitate later discussion. The most important concept to understand here is the distinction between the laser **oscillator** and laser **amplifier**.

Any individual gain medium can be set up in such a way as to function either as a laser **oscillator** or as a laser **amplifier**. A detailed mathematical discussion of these is beyond the scope of this book, and more detailed description can be found in references [2, 3, 4, 5]. Here we will focus on qualitative differences and those that are key to understanding overall ultrafast laser systems, starting with a discussion of a laser amplifier.

Amplifiers

As laser **amplifier** is a part of a laser which **amplifies** an existing laser beam up to a desired power level. There are a multitude of different geometries of laser amplifier, and some of these will be discussed in due course, but some features are common to all.

In all of the cases we will discuss in the context of ultrafast optics, amplification requires a gain medium, a pump laser, and a weaker seed laser. Energy is put into the gain medium by the pump laser beam, which is then 'picked up' by the seed laser beam. The effect of the amplifier is then that the energy of the pump beam is reduced, whereas the energy of the seed beam is increased. This is exactly the situation drawn in Figure 1.1, where E_{seed} is amplified. Figure 4.1 shows a simplified schematic of the amplification process for illustrative purposes. The extent to which the energy of the seed beam is increased is referred to as the **gain** of the amplifier, and is defined as the ratio of the output power to the input power.

Laser amplifiers can only amplify across a certain range of frequencies, as the gain medium has a finite emission bandwidth (see Figure 1.4 and Table 1.1). The bandwidth where a laser amplifier can effectively amplify is known as the **gain bandwidth**. For ultrafast laser systems, we desire a *broad* gain bandwidth, as we have to make sure we effectively amplify the full bandwidth of our ultrashort pulse. When the amplifier cannot amplify the full bandwidth of the input pulse, it leads to a situation called **gain narrowing** - the reduction in bandwidth of a pulse following amplification, which would have knock-on effects for the pulse duration in an ultrafast laser. Ti:Sa has a very broad gain bandwidth, which is the reason for the ubiquity of Ti:Sa in ultrafast laser architecture. Yb:KGW has a narrower gain bandwidth (but still broad enough for ultrashort laser pulses), but has other qualities which can lead to it having advantages over Ti:Sa – these will be discussed further later.

Often there are multiple amplification stages within a laser system. These can broadly be divided into **preamplifiers** and **power amplifiers**, depending on the gain produced[1].

[1] The preamp/power amp definition is less well defined than in electronic amplifiers, where preamplifiers typically provide high voltage gain but relatively little current gain, whereas power amplifiers produce large current gain but smaller voltage gain. For optical amplifiers, the distinction seems to depend on the power

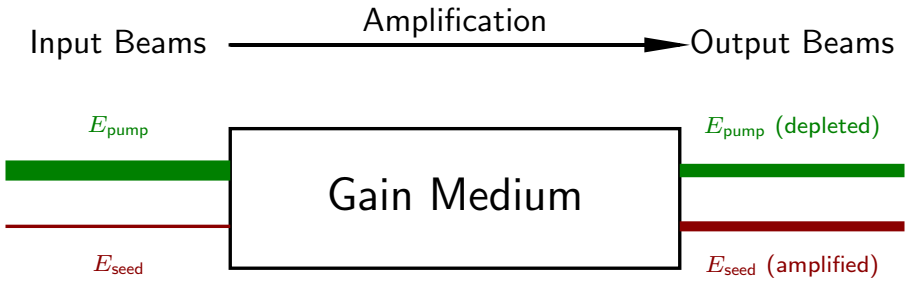


Figure 4.1: Simplified schematic of how a laser amplifier works. The incoming seed beam (red) is amplified, and the pump beam (green) is depleted, via interaction in the gain medium.

4.1.1 Oscillators

A laser **oscillator** is a part of a laser which *generates the initial, weak, laser pulses*. In a high-power laser system, these weak pulses are then used as a *seed* for an amplifier further downstream. In a simple case, taking a laser amplifier and applying feedback to it (i.e. returning some of the amplified output to the input of the amplifier) will cause to to **oscillate** at the frequency of the radiation that is fed back. Simplistically this can be thought of as taking the output light from a stimulated emission process, and re-routing it back into the input, inducing more of the same kind of stimulated emission.

Oscillation is practically achieved by building an **optical cavity** around an amplifier, so that the laser beam resonates around the cavity, passing through the gain medium on each round trip, then the emission is returned to the input by the cavity, stimulating more emission, and oscillation occurs. An optical cavity is simply a region enclosed by mirrors, such that a laser beam injected into it will bounce round and round it rather than passing straight through. A simplified schematic of how an amplifier can be turned into an oscillator is shown in Figure 4.2.

The function of the oscillator shown in Figure 4.2 can be described as follows (following the circled numbers). (1) Emission (generally stimulated) from the gain medium is directed into the cavity using mirrors. (2) This emission travels around the cavity is directed back into the gain medium, causing further stimulated emission, and oscillation. (3) The first mirror after the gain medium is an **output coupler**, a partially reflective mirror that reflects some

level of the output beam (high power beam = power amplifier), but it is relatively ill-defined.

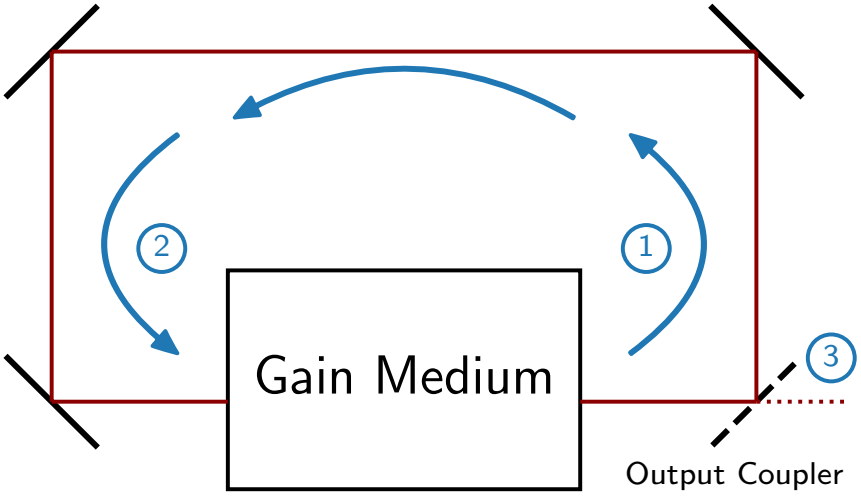


Figure 4.2: Simplified schematic of a laser oscillator. The red lines show laser beams in the cavity (formed by mirrors shown in black), and the blue arrows show the direction of travel around the cavity. Circled numbers show important areas of the cavity, see text for details. The oscillator pump beam is not shown.

light back into the oscillator (as in stage (1)), but allows some light out to be amplified further in the laser system.

Ensuring that the oscillator is oscillating in the desired way (so it is producing the pulses you want) is a key design consideration. A full mathematical description of how this is achieved is actually quite accessible and can be found in standard texts such as references [2, 3, 6], but a detailed discussion is beyond the scope of this book.

Essentially, there will often be multiple competing processes taking place within the oscillator cavity, producing different kinds of laser output. When the light makes one complete circuit around the cavity there will be a certain amount of gain, referred to as the **round-trip gain**; but also a certain amount of loss, referred to as the **cavity loss**. Intuitively, if the round-trip gain for a certain process is greater than the cavity loss, then the laser will produce output via that process. If, conversely, the cavity loss is greater than the round-trip gain for a given process, then there will be no laser output from that process. The point at

which the round-trip gain exceeds the cavity loss is referred to as the **lasing threshold**. The construction and alignment of the oscillator has to be such that the desired process has the highest round trip gain and lowest losses, so that it dominates over other undesired processes. We will discuss the workings of more specific ultrafast oscillators in due course.

4.1.2 Pump Lasers

Finally, a brief word about **pump lasers** before we continue. In ultrafast laser systems, we use pump lasers to excite our gain medium that will eventually produce ultrafast output. The pump laser itself is usually **not** an ultrafast laser, but a laser with a high average power that can effectively produce the population inversion in our gain medium. Different ultrafast lasers use different kinds of pump lasers – usually either continuous wave or nanosecond pulsed lasers, which each have different characteristics which will be touched on later. Ideally, you won't ever have to touch a pump laser. In my view they are the scariest lasers we tend to have in an ultrafast laser lab, they are obnoxiously powerful and (in the case of Ti:Sa lasers) emit blinding bright green light. They're usually locked away in hermetically sealed boxes, and tend to get adjusted by real laser engineers than by amateurs like us[2].

4.2 Ultrafast Laser Systems

Note that above we specifically referred to a **laser system** rather than a single laser. This is because ultrafast lasers tend to consist of multiple different lasers that all function together to produce the final ultrashort pulse output. As discussed, within any laser system, there will always be an oscillator and a pump laser, and potentially an amplifier and separate compressor (in a high power system).

Intuitively, the oscillator produces low energy pulses, which are amplified by the amplifier, and compressed by the compressor. Within the oscillator and amplifier, there will be pump lasers which pump the gain medium up to a high energy, stretchers to stretch pulses before amplification, and the oscillator may have a separate pump laser too. A schematic of a typical amplified ultrafast laser system, with some representative rep rates and pulse energies for the outputs of each stage is shown in Figure 4.3. We will discuss the three main components in more detail (oscillators, amplifiers, and compressors) here.

Note here that Figure 4.3 specifically shows an **amplified** ultrafast laser system. It is also possible to simply use the oscillator output directly without amplification – where you get a higher repetition rate but much lower pulse energy. To provide some focus as you read the remainder of chapter, typical applications would be:

- **Amplified Systems**: High pulse energy, low rep rate. Typically used in applications

[2]Of course, tweaking the internals of lasers to optimise them *with permission from your supervisor* is all part of the fun (why would they make adjustable mirrors if they didn't want us to adjust them?). In the case of these high-power pump lasers though, best left to experts. Sadly for those of us who like to tinker, more and more lasers are produced with no user-adjustable parts – though this is undeniably beneficial to the output stability and probably means more science gets done.

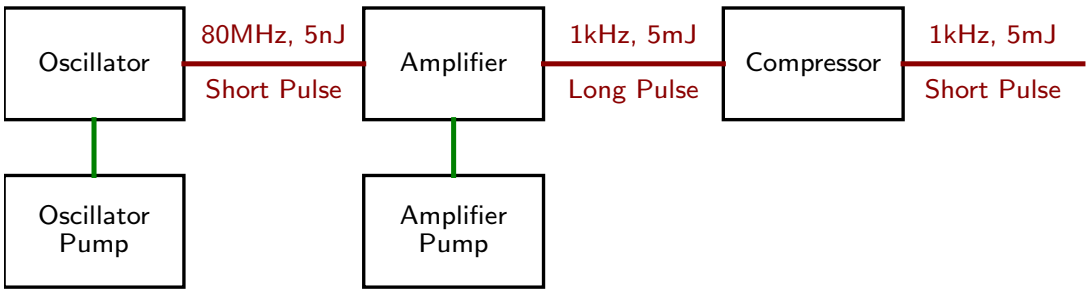


Figure 4.3: Schematic of a typical amplified ultrafast laser system. Pump laser beams are shown in green, and the final ultrashort pulse output beam is shown in red. Representative repetition rates and pulse energies after each stage are shown for illustrative reasons, and are typical of a Ti:Sa system.

where:

  - High intensities are needed, provided by the high pulse energy. (e.g. to effectively drive OPAs , or in strong-field physics experiments).

  - High time resolution ($<$50 fs) is needed, as the pulses can be made shorter by SPM, which requires high pulse energy (e.g. extremely fast spectroscopy, looking towards attosecond experiments).

- **Standalone Oscillators**: Low pulse energy, high rep rate. Typically used in applications where:

  - High pulse energy is not useful, as it would damage samples (e.g. microscopy and imaging, applications in biology and neuroscience).

  - High repetition rate is beneficial (e.g. experiments using fragile samples or that produce low signals, as more data can be collected more quickly - improving the signal-to-noise ratio).

4.2.1 Oscillators

As mentioned earlier, a **laser oscillator** is the part of the laser system where we generate very low energy ultrashort laser pulses which are amplified further down the road in the laser system. We've already discussed how an oscillator is essentially an amplifier with a cavity

built around it to provide feedback, but now we can discuss in more detail the techniques for making an oscillator produce ultrashort (femtosecond) laser pulses specifically.

Previously we gave a very simple discussion of oscillator construction in terms of the round-trip gains or losses that a particularly lasing pathway experiences in the oscillator cavity. There we said that ensuring the round-trip gain exceeds the round-trip loss was a way to make our cavity oscillate in the desired way. This idea can be put on a more formal footing by considering something called the **quality factor**, or **Q factor** of the cavity. The Q factor is defined as the ratio of the round trip gain to the round trip loss, so a cavity with a high Q factor produces a lot of gain, whereas a cavity with a low Q factor is inherently lossy.

Q Switching?

We can imagine, then, that a way to produce pulsed laser output is to find a way to *switch* the Q factor of the cavity from low to high for a very short period of time. This is called **Q switching**, and is very common way of making a laser oscillator. Q switching will be familiar if you have ever used YAG lasers, or other nanosecond lasers.

At a very basic level, Q switching amounts to switching the laser on and off very quickly to produce the pulsed output. Energy is built up in the gain medium when the Q is low, and then is quickly dumped out when the Q is high. Electronic switches can easily pulse on for a nanosecond (or so), and so Q switching can easily produce nanosecond laser output. However, electronic switches are inherently limited in how fast they can switch, and **it would not be possible** to find an electronic switch that can switch fast enough to produce a femtosecond laser pulse. The solution is to use optical means to produce our ultrashort pulses, via a phenomenon called **modelocking**.

Modelocking

We have already met the concept of modelocking in passing in chapter 1 of this book. To briefly reiterate, there we mentioned how the shortest pulses are made when the largest number of cavity modes all have a fixed phase relationship to one another. When this is the case, the cavity is said to be **modelocked**. This situation was shown in Figure 1.7 - when colours oscillate in phase, the laser produces a stable train of short pulses. When the phases are random, the laser does not produce a stable train of pulses. The final pulse duration depends on the number of modelocked colours, and the overall repetition rate of the oscillator is determined by the time it takes for a pulse to travel around the oscillator cavity (Figure 4.4).

So, to get modelocked output, there are two things to do:

1. Ensure that there is a large enough bandwidth in the cavity to support generation of ultrashort pulses.

2. Ensure that as many of the colours within this bandwidth (the cavity modes) as

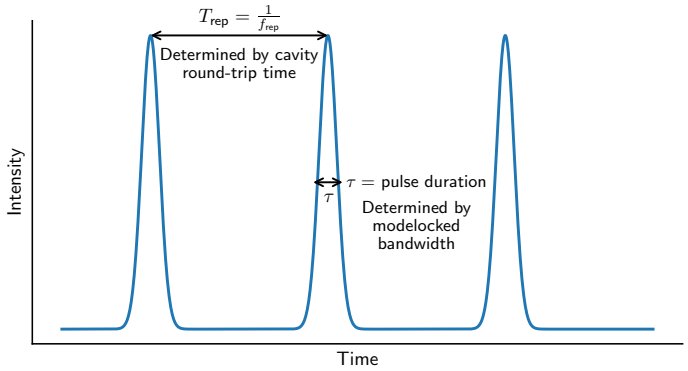


Figure 4.4: A modelocked pulse train. The rep rate is determined by the time it takes for a pulse to travel around the oscillator cavity. The pulse duration is determined by the number of modes that are locked.

possible are locked in phase.

The first of these points is easily achieved by using a gain medium like Ti:Sa or Yb:KGW and pumping it to produce broadband laser emission, as described in chapter 1. The second point (ensuring the cavity is modelocked) is harder to achieve, but ultimately it boils down to modulating the cavity gain such that the gain for the desired modelocked process exceeds the gain for the other undesired processes[3]. There are a variety of different methods for achieving this, but they can be broadly divided into **active modelocking** and **passive modelocking** techniques.

Active modelocking techniques modulate the cavity gain using some modulator driven by an external signal; whereas passive modelocking techniques rely on the light in the cavity itself to drive a modulator. Active modelocking can produce more stable pulse trains, but as this modulator is typically electronically driven, actively modelocked lasers do not produce the shortest output pulses (although can still produce pulses shorter than 100 fs). Passively modelocked lasers have a modulator driven by the light pulses themselves, so the modulator can be 'switched' much more quickly, and so they produce the shortest pulses (sub 5 fs). More detailed descriptions of modelocking can be found in references [4, 2, 7, 3, 6].

There are a very large number of different types of laser oscillator, and the purpose of this book is not to provide an exhaustive list of all the different possibilities[4]. Rather, we want to understand the basics of how an oscillator works so that understanding the specific oscillator we have in our laser system is not so mystifying. The operating principles of femtosecond laser oscillators are, I think, best shown with an illustrative example: **Kerr-lens modelocking**.

[3] As the cavity could easily oscillate in continuous wave mode, rather than being modelocked. This would produce continuous wave output (called *'cw spiking'* or *'cw breakthrough'*), which is not what we want when using ultrafast lasers and can damage amplifiers

[4] Reference [6] is an excellent place to start if you want this.

Kerr-Lens Modelocking

Within passive modelocking, there is one technique that gives an instructive and intuitive illustration of how cavity gains/losses can be modulated by the laser pulses themselves. This is **Kerr-lens modelocking**. We have discussed the creation of a **Kerr lens** by self-focussing in chapter 3; but to briefly reiterate it is when the intensity of a pulse going into a material is high enough to cause the refractive index to become intensity dependent. This then results in a transient lens being 'imprinted' on the material as the pulse goes through it, as was shown in Figure 3.2. This leads to self-focussing, and in a Kerr-lens modelocked laser we exploit this effect to enhance the cavity gain for our modelocked process, while increasing the cavity loss for undesired processes.

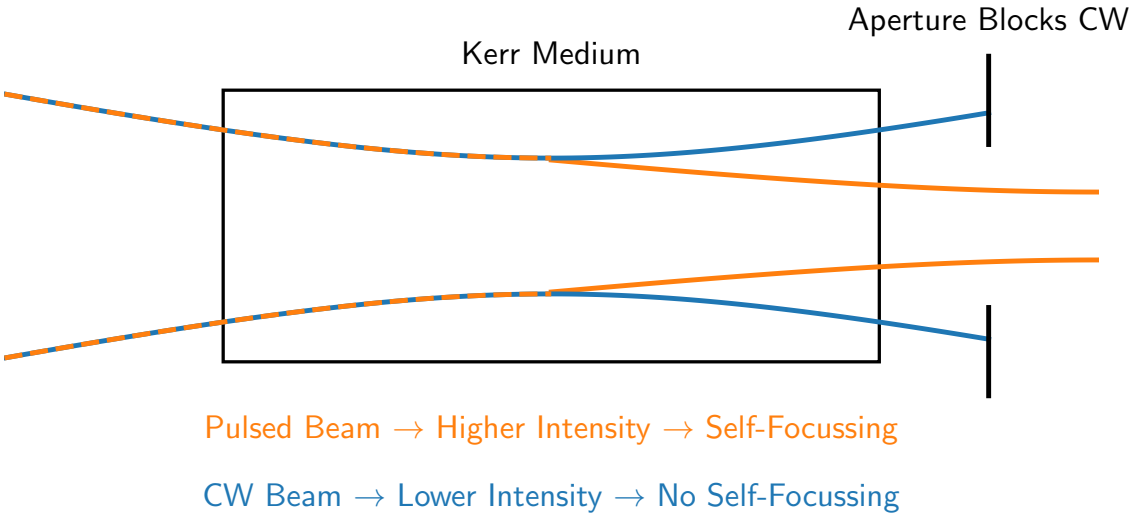


Figure 4.5: Illustration of the operating principle behind Kerr-lens modelocking. Initially co-propagating pulsed (orange) and continuous-wave (blue) beams pass through a thick glass Kerr medium. The less intense continuous-wave beam is blocked by an aperture at the exit of the medium. The more intense pulsed beam undergoes self-focussing via the Kerr effect, and can then pass through the aperture unhindered.

Figure 4.5 shows an illustration of one possible scheme for Kerr-lens modelocking[5]. Kerr-lens modelocking exploits the fact that when the oscillator is producing the shortest possible pulses, the intensity of the pulses is necessarily as high as is possible. These pulses imprint a (relatively) strong transient lens on a piece of glass in the cavity called the **Kerr medium**[6]. The pulses then undergo self focussing to such an extent that they can pass unhindered through an aperture placed after the Kerr medium. Lasing pathways which do not produce the most intense pulses (or produce cw output) do not undergo self-focussing, and are blocked by the aperture. They therefore experience much higher cavity losses, with the result is that the laser oscillator preferentially produces the ultrashort pulses.

[5] There are other geometries, but this illustrates the fundamental principle.
[6] Normally the gain medium itself fulfils this role.

In practice this is often done by using the pump beam as a pseudo-aperture. The pump beam can be focussed into the crystal such that only the more strongly Kerr lensed seed beam (the ultrashort pulses) overlap well with the pump beam, so they experience much higher gain. This is more common than use of a physical aperture, but the underlying principle is the same. At this point, the oscillator is modelocked, and a stable train of ultrashort pulses is produced. There will normally also be an intra-cavity pulse compressor to account for the GDD accumulated by the pulse on each round-trip.

We will end our discussion of ultrashort pulse oscillators by emphasising two key points. Firstly, modelocking is inherently a **multimode** phenomenon. The train of ultrashort pulses is created by the interference of a large number of different cavity modes all oscillating in phase. This is in contrast to a Q-switched laser, where the pulses are produced by a sudden reduction in cavity losses, producing a pulse. Secondly, it should be stressed that there are many other different ways to produce a mode-locked oscillator than are described here. Researching and understanding what kind of oscillator your specific laser system has is well worth the time spent.

4.3 Amplifiers

We turn now to a discussion of ultrafast laser amplifiers. We introduced the idea of a laser amplifier as a place where a seed laser beam is amplified and a pump laser beam is depleted. The seed laser in this case is the output from our oscillator, and the pump laser is a high-energy laser beam that can efficiently transfer a lot of energy to our gain medium. We will discuss pump lasers in more detail shortly.

There is one significant complication that arises when amplifying ultrashort pulses, which is that an ultrashort pulse delivers its energy in a very short time due to the short pulse duration, meaning that the **peak power** of the pulse is often on the order of 100 TW. Powers this high can easily cause unwanted effects in the gain medium, if they do not damage or destroy it. Without a solution, this problem would limit the output power of ultrafast laser amplifiers considerably. Thankfully, a solution to this problem is **chirped pulse amplification**, and the Nobel Prize in Physics in 2018 was awarded for the development of this technique [8]. Essentially all ultrafast laser amplifiers use this technique.

4.3.1 Chirped Pulse Amplification

Chirped pulse amplification (CPA) is a technique where an ultrashort pulse is **stretched** before it is amplified. Stretching makes the pulse longer, and reduces the peak power back to a level that will not damage the gain medium. We already know that to stretch an ultrashort pulse, the colours need to spread out in time, and so we have to make the pulse chirped. The stretched, chirped pulse is then amplified safely in the gain medium, and the resulting higher energy amplified pulse is subsequently re-compressed back to an ultrashort pulse. This is operating principle of CPA, and it is illustrated in Figure 4.6.

An common question regarding CPA is to ask why the pulses that are intense enough to

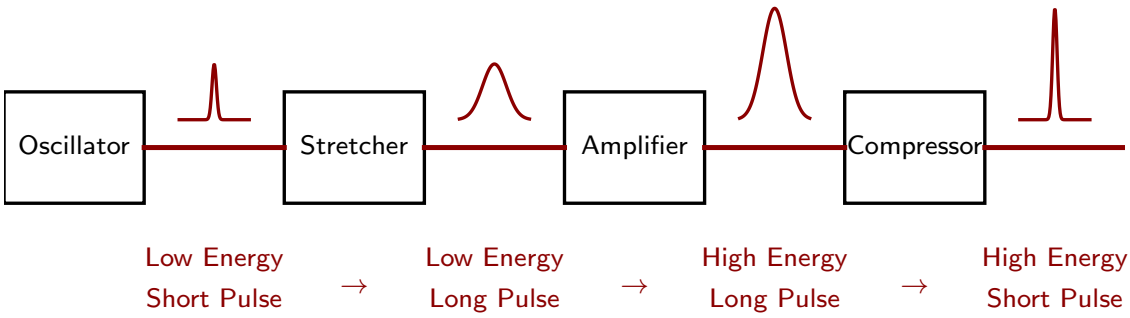


Figure 4.6: Operating principle of chirped pulse amplification. Low energy ultrashort pulses are stretched, amplified, and re-compressed to produce high energy ultrashort pulses without damaging the internal optics of the amplifier.

damage the laser gain medium without being stretched, don't damage mirrors and other optics after the laser output when they are re-compressed. The answer is two-fold. Firstly, reflective optics like mirrors are **much** less susceptible to damage than transmissive optics like a laser crystal, so external mirrors and gratings are much less likely to be damaged. Secondly, the beam inside the gain medium has a very small beam waist, and almost any optic placed in a focussed, compressed, full power beam from the laser output would definitely be damaged! In many ultrafast laser systems producing high energy pulses, the output beam waist is comparatively large (10 mm or more) to try to minimise this effect.

Having understood CPA, we now turn to look at some specifics of the amplification process.

4.3.2 Amplifier Design

There are many possible ways to make an ultrafast laser amplifier. With an aim of encouraging an appreciation of how your lab equipment works, here we will discuss two types, the **single/multipass amplifier** and the **regenerative amplifier**, which are commonly encountered. All amplifiers discussed here are chirped pulse amplifiers, so use a stretched seed pulse – the way the pulse is stretched will be discussed in due course.

4.3.3 Single and Multi-Pass Amplifiers

Conceptually the simplest kind of amplifier is one where the seed and pump beam are both overlapped in the gain medium, and amplification occurs – as illustrated in Figure 4.7. If the beams take a single pass through the gain medium, we would call this a **single pass**

amplifier. If the seed beam makes multiple passes by being steered into the gain medium multiple times by mirrors, we would call it a **multipass amplifier**. In the case that the seed makes multiple passes at the same angle via electronic switching, then we have a **regenerative amplifier**, discussed below.

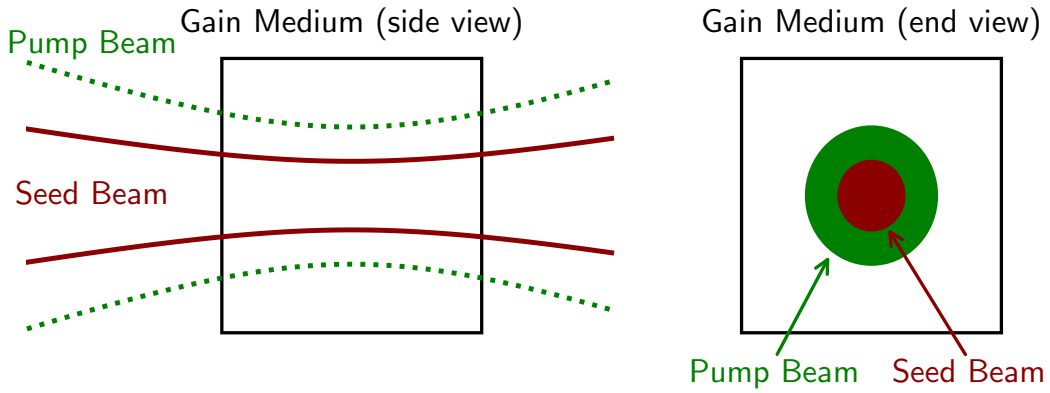


Figure 4.7: Generic laser amplifier – pump and seed beams are spatially overlapped in the gain medium. Good beam quality and good spatial overlap are essential to ensure high amplification efficiency.

Single and multi-pass amplifiers are actually relatively rare, except as additional amplification stages or when single passes through long fibre based gain media are made. However, they are conceptually simple and Figure 4.7 outlines an important point. To achieve good amplification efficiency, the pump and seed beams need to be **spatially overlapped** in the gain medium – when spatial overlap is optimised, we get the most efficient transfer of energy between the pump and seed beams, and the most efficient amplification. In general this is achieved by the means displayed in Figure 4.7: focus the seed beam tighter than the pump beam, so that all of the seed is in an area of high pump intensity, and none of it is 'unpumped'.

Furthermore, the *beam quality* plays a very important role here. Poor beam quality (not having an ideal TEM_{00} mode – Figure 1.14 (b)) reduces the effective spatial overlap and can have a dramatic effect on the output efficiency. This kind of spatial overlap is important in all kinds of laser amplifier, but also in nonlinear optics (where beams are mixed together), and in experiments where the effects of multiple beams on a sample are being investigated (such as in time-resolved spectroscopy).

Regenerative Amplification

We will now turn to look at the **regenerative amplifier**. The fundamental operating principle behind regenerative amplifiers is that the laser gain medium is pumped up to a

high energy state using the pump laser, and then the seed pulse passes through the gain medium multiple times to extract the maximum energy possible from the laser medium, achieving the maximum gain possible. So, having pumped the gain medium up to a high energy state with the pump laser, we need to find a way to make multiple passes through the gain medium. We could do this in the *multipass* way described above, but these are fiddly to align and achieving good spatial overlap can be challenging.

Instead, we will enclose the gain medium inside a resonator, much like how an oscillator is constructed. A simplified schematic of how a regenerative amplifier works is shown in Figure 4.8. However, unlike in an oscillator, you do not just let the seed pulse resonate round and round for a long time, as we only want it to complete enough round-trips to pass through the gain medium enough times to extract all the energy - in general around 10-20 round trips are required. This requires that the seed pulse is injected at a specific time relative to when the gain medium is pumped, allowed to resonate around the cavity for a fixed number of times, and then coupled out of the cavity after it has extracted all possible energy from the medium. This can be achieved by using an electro-optical gate (a *Pockels cell*) that can allow you to inject or eject a pulse into a cavity at a specific time via electronic control[7]. Adjusting the timing of the gates can then be used to adjust how much gain you add. This type of amplifier is generally what people refer to when they say 'regenerative amplifier'.

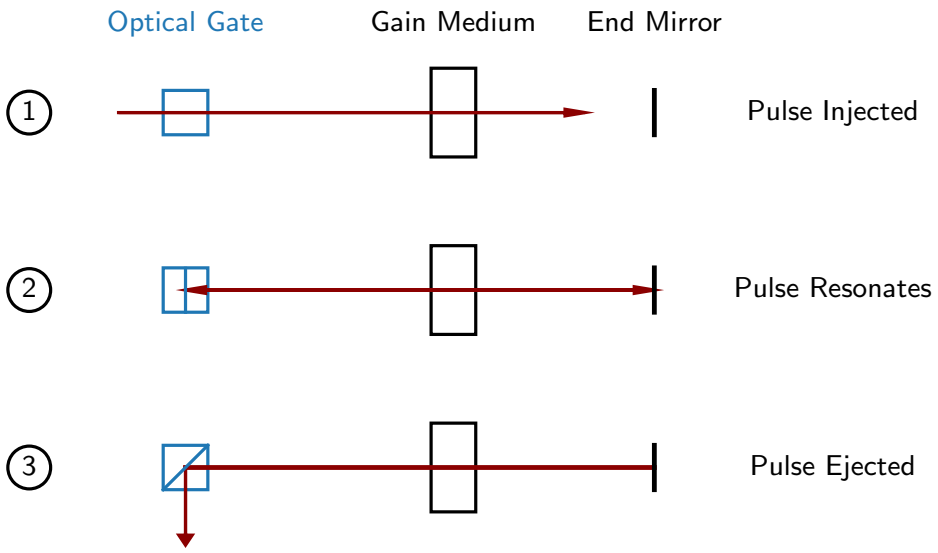


Figure 4.8: Simplified schematic of how a regenerative amplifier works. The electro-optical gate (blue) is opened to allow a pulse to enter (1), then closed such that the pulse resonates around a cavity and passes through the (pumped) gain medium multiple times (2). Finally the gate is re-opened, and the amplified pulse exits the cavity (3).

The question remains, though, as to why we would choose one type of amplifier over the

[7] As generally the cavity round trip time would be some nanoseconds - so electronically controlled gates would be able to switch fast enough.

other? Each type of amplifier has characteristics that are beneficial in certain circumstances. Single and multipass amplifiers do not need complex driving electronics; but can be difficult to align and generally the same number of passes through the gain medium cannot be achieved as in an electronically gated regenerative amplifier. Furthermore, the need to enter the gain medium from different angles reduces the effective spatial overlap, whereas the pump and seed can be almost collinear in a regenerative amplifier. This means that electro-optical gate driven regenerative amplifiers are able to provide higher overall gain. However, an electro-optical gate contains relatively thick pieces of glass and so the dispersion added by a regenerative cavity is typically much higher than that added by a multipass geometry, making multipass geometries preferable when very short output pulse durations are required[8]. In some laser systems, there is a regenerative preamplifier followed a single or multipass power amplifier.

4.3.4 Pulse Compression and Stretching

By the nature of chirped-pulse amplification, the pulses that come straight out of our amplifier were stretched before amplification and are very chirped, and therefore temporally broad. They need to be compressed again before they are useful to us as ultrashort pulses - and this is done using a **compressor**. We gave a brief overview of pulse compression at the end of chapter 2, but here we will discuss the geometries of different types of compressor. A compressor and a stretcher are identical pieces of equipment – a compressor can turn into a stretcher if you try to compress a pulse that is already at the transform limit! Hence, the geometries below are also seen as stretchers in CPA systems.

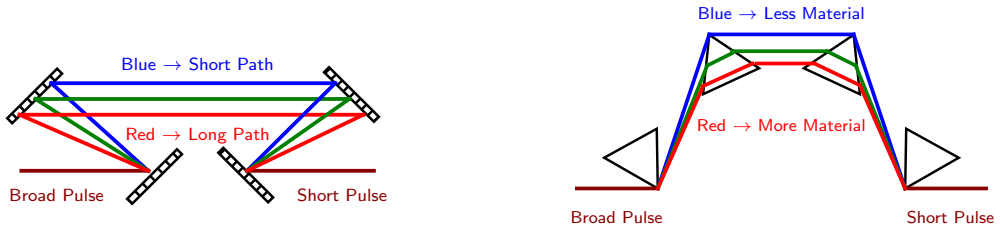
Compressor Geometries

Two very common kinds of pulse compressor are the **prism compressor** and the **grating compressor**. To discuss how each individual type of compressor works, we look at Figure 4.9. Initially, the grating compressor (a), works by making the blue components of the pulse travel over a shorter distance than the red components, with the result that negative GDD is added. The prism compressor (b) is similar in principle. However, in this case the red components do **not** have a longer path length, but they travel through more of the prism (as the blue components pass nearer the apex). This means that they spend longer travelling in a high refractive index region and slow down relative to the blue components, allowing the blue components to 'catch up'.

Grating compressors generally can produce larger amounts of negative GDD than prism compressors - but are more expensive, and generally have a lower transmission efficiency. At the time of writing, the best grating compressors are only about 70% efficient (so 10 W uncompressed becomes 7 W compressed). A prism compressor can be aligned to have a much higher transmission[9], but often cannot handle very high input powers due to self-

[8] Whilst a compressor can mitigate the dispersion added by the gate to some extent, it cannot generally compensate for higher-order dispersion. Higher order dispersion becomes a limiting factor for the shortest pulse durations.

[9] By placing the prisms at Brewster's angle.



(a) Grating compressor. (b) Prism compressor.

Figure 4.9: (a) Illustration of a grating compressor. Blue components travel over a shorter path length, so 'catch up' to the red components. (b) Illustration of a prism compressor. Blue components travel through less of the prism material than the red components, so they are slowed down less by the prism, and catch up to the red components. Note that in practical cases retroreflectors can be placed where the dispersed beams run parallel to reduce the number of gratings/prisms needed by two. Both compressors are shown as if they are compressing a positively chirped pulse (adding negative GDD).

focussing (or other non-linear effects). The compressor that you find after a chirped-pulse amplifier will almost always be a grating compressor for these reasons. If you plan to build a prism or grating compressor then a good place to start is the Optics Toolbox [9], as here there are tools that let you calculate the GDD (and TOD) for an arbitrary geometry where you can input the relevant parameters. This is ideal as a way to get a feel for how changing the geometry changes the added GDD too.

Chirped Mirrors

A third way to produce a compressor is to use a **chirped mirror**. Chirped mirrors will be discussed more in chapter 6, but essentially the mirror is designed such that each bounce on the mirror surface induces negative GDD. Typically, each bounce adds on the order of $-50\,\mathrm{fs}^2$. A fixed amount of negative GDD can then be added by simply bouncing on the mirrors multiple times. The usual strategy is then to overshoot, such that your pulses become negatively chirped, and then do fine adjustment using a pair of thin glass wedges (adding positive GDD) to optimise the compression. The advantage of chirped mirrors is that the surfaces can be engineered to effectively control higher-order dispersion over a very wide bandwidth.

Pulse Shaping

None of the compressors described above can compensate for third-order dispersion. One way to do this requires use of a 'GRISM', which is a grating etched into a prism. Another

way is to use a **pulse shaper**, which is essentially a device which allows the phase of each individual colour in the pulse to be arbitrarily controlled. Being able to directly and arbitrarily control the spectral phase makes these devices almost the 'ultimate' pulse compressor. Despite their high cost, they are becoming more and more common in university laboratories, especially with recent advances in multi-dimensional spectroscopy.

Alternative Strategies For Compression

Depending on the type of pulses you have, there may also we other ways to compress your pulses. If you have IR pulses (longer wavelength than around 1300 nm), then you may be able to use material dispersion to add negative GDD. We said that generally materials don't tend to provide negative GDD, so we could not easily compensate for the added positive GDD from transmissive optics by sending the beam through a piece of exotic material. While this is true in the visible and near-IR wavelength regions, in the mid to far IR regions, several common optical substrates actually add negative GDD. For example, BK7 adds negative GDD at wavelengths longer than around 1300 nm. Another useful technique at more conventional visible/near IR wavelengths is to utilise material dispersion to add positive GDD. If we used a compressor or chirped mirrors to give our pulse a negative chirp, we could then compress it by passing it through some thickness of glass. We could pass it through a thin glass wedge, which can be easily scanned in and out of the beam, changing the amount of material dispersion in a very fine and controllable way.

A final important idea is **pre-compensation** for dispersion. The crux of this idea is that it doesn't matter which order you add the GDD in (either positive first or negative first), it all gives the same result in the end. For example, if our ultrashort pulses exit the laser system as transform-limited pulses, then in most cases they will end up only gaining positive GDD as they propagate through our beamline, and so will become positively chirped. A very easy way to compensate for this is to utilise the compressor inside the laser system and alter the chirp of the pulses at the direct laser output such that they are negatively chirped as they exit the laser system. Then, as they accumulate positive GDD travelling through the beamline, they will arrive at our experimental target transform-limited. By looking at some kind of intensity dependent experimental signal, the compressor can be adjusted to exactly compensate for the accumulated positive GDD, ensuring that the pulses are as short as possible where it counts: at your experiment[10]

4.4 Optical Parametric Amplifiers

The final type of amplifier we will discuss is the **optical parametric amplifer** (OPA) that was discussed in chapter 3. A key difference between the amplifiers discussed previously and with an OPA is that *no energy is stored within the gain medium of an OPA*. The gain medium is simply a non-linear crystal in which a pump and signal beam with energies E_{pump} and E_{signal} are mixed in a way that amplifies the signal beam, and also produces a third

[10] Of course, if you are using the laser to drive an OPA, then you can't really do this – OPAs are very sensitive to input pulse duration.

idler beam with energy E_{idler} to conserve energy and momentum. We have already seen an overview of how the process of optical parametric amplification works, so we will now discuss some points about practical OPA construction that are instructive. Figure 4.10 shows a schematic of the internals of a typical OPA.

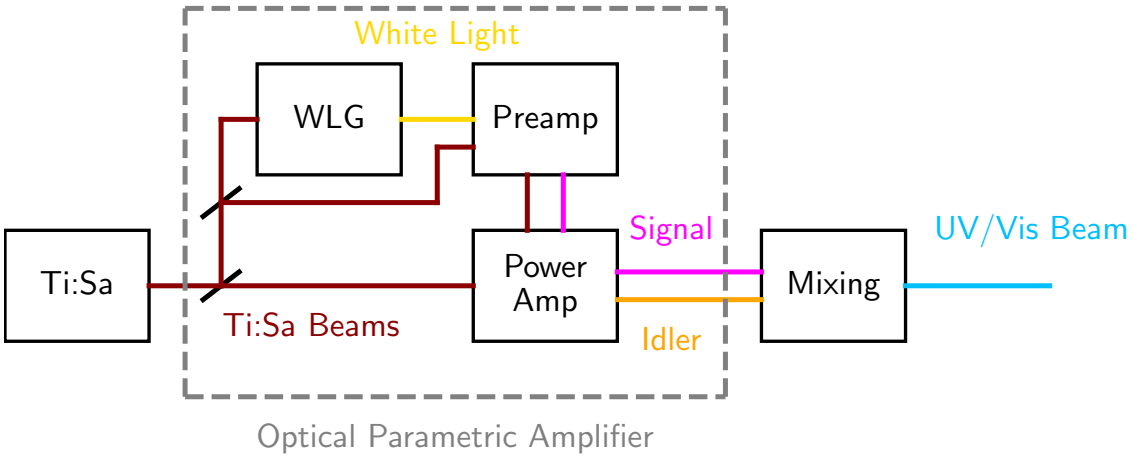


Figure 4.10: A schematic of a typical OPA system. The parts comprising the OPA are surrounded with the dashed grey box, an an external frequency mixing stage is also shown at the output. The input Ti:Sa beam is split by beamsplitters shown in black (mirrors in the beamline are not shown).

The pump beam for an OPA is typically the output from an ultrafast amplifier, and is shown in red on Figure 4.10. For users the real benefit of an OPA lies in the wide wavelength tunability, but this requires that we generate a tunable signal beam from our pump beam[11]. A way to achieve this that is commonly seen in OPAs (such as the Light Conversion TOPAS or ORPHEUS OPAs) is to split the input pump beam up, and use a small part of it to drive **white light generation** (WLG) - discussed in chapter 3 (Figure 3.3).

As discussed, WLG uses self-phase modulation and other nonlinear effects to dramatically increase the bandwidth of an input ultrafast laser pulse. In an OPA, the spectrally broad white light is then collimated and directed into a second amplification stage. In this stage, the white light beam is used as the signal beam in an OPA stage, and tuning the angle of the OPA crystal and temporal delay between the white light and pump pulses allows a specific wavelength to be amplified – this process is how the tunability of an OPA arises.

Like a laser system, an OPA may consist of multiple amplification stages, shown as a preamp and power amp in Figure 4.10. The weakly amplified signal beam passes through crystals in each of these stages, gaining energy each time (as each stage is pumped by a fraction

[11] Feeding in a signal beam from a different laser would not be possible as it would not be temporally synchronised to the main pump beam.

of the input pump beam). The phase-matching angles of all the crystals need to be tuned to amplify the desired wavelength of light, and generally an OPA has a large number of motors which are controlled in synchrony to achieve this. The output from the OPA overall is then an amplified signal beam, together with the idler beam generated by the parametric amplification process. For a commercial OPA (Light Conversion TOPAS) pumped using 800 nm light from a Ti:Sa laser, the direct output wavelength (signal or idler) is generally tunable between around 1100 nm and 2600 nm. For an equivalent commercial OPA (Light Conversion ORPHEUS) pumped by Yb lasers (fundamental at 1030 nm), the signal/idler output is actually tunable between 600 and 2600 nm - which is accomplished by using the second harmonic of the fundamental (515 nm) to drive the process too.

While OPAs do produce this broad range of output wavelengths, they are not uniformly efficient across all wavelengths bandwidth (see Figure 3.12) An OPA will normally come with **tuning curves**, which are plots of output pulse energy against wavelength, telling you how efficient it is in different spectral regions – or they can be looked up.

Finally, as was mentioned in the previous chapter, many commercial OPAs come with an option for **external frequency mixing** - this is essentially a box containing motorised non-linear crystals that can serve to further extend the tuning range. An external frequency mixer generating UV/visible light is shown as the last stage in Figure 4.10, but equivalent mixers to generate mid-IR light are available. They can easily be homemade (as they are really just a box containing a crystal and some dichroic mirrors), but having them integrated into the OPA system and computer controlled is much more user-friendly.

4.5 Amplifier Miscellany

Again, here we discuss some interesting and instructive concepts regarding ultrashort pulse generation and amplification, but which are not integral to the main text.

4.5.1 Quantum Defects

When considering the amplification process in any laser, it is clear that the photon energy of the output must be lower than that of the pump[12] (see Figure 1.1). The difference in energy between the pump photon and the output photon is called the **quantum defect** of the laser:

$$\text{Quantum Defect} = E_{\text{pump}} - E_{\text{seed}}$$

This energy has to go somewhere, and generally it is dissipated as heat in the gain medium[13]. Heating of the gain medium reduces the efficiency of the laser, as too much heating will can spoil the quality of the output beam (through processes like *thermal lensing*), and eventually can damage or burn the gain medium itself. To remedy this, the gain medium must be

[12] In the situation where we absorb one pump photon and emit one output photon, almost always the case.

[13] As discussed, the OPA is an important exception to this rule.

Example: Quantum Defects

Table 1.1 showed typical pump and emission photon wavelengths for Ti:Sa and Yb:KGW lasers. Conversion of these wavelengths to an energy allows the quantum defect to be calculated:

Laser Medium	Pump Energy (eV)	Emission Energy (eV)	Quantum Defect (eV)
Ti:Sapphire	2.48	1.55	0.93
Yb:KGW	1.26	1.19	0.07

Clearly, the quantum defect is around 10x higher for the Ti:Sa laser, and so for the same pulse energy and repetition rate, 10x more energy has to be dissipated in the gain medium for a Ti:Sa laser!
The effect of this in practice is that it puts an upper limit on the possible output power of a Ti:Sa laser – Table 1.2 showed that the average power of Yb:KGW lasers is roughly an order of magnitude greater than that of Ti:Sa lasers. The reduced quantum defect accounts for a lot of this difference.

cooled[14]. Minimising the quantum defect of a laser reduces the amount of energy that is stored in the gain medium and allows higher average powers to be achieved – comparing the quantum defect of a typical Ti:Sa laser with a typical Yb:KGW laser illustrates this point.

4.5.2 Choice of Gain Medium

If you are trying to understand what kind of laser you have in your lab (or, if you are looking to buy an ultrafast laser), probably the first thing to do is to look at the kind of gain medium you have. There are basically two options, as mentioned previously: **Titanium Sapphire** (Ti:Sa), and **Ytterbium-based** media (usually Yb:KGW). Here we will briefly discuss the key differences between these two gain media.

Titanium Sapphire

As summarised back in Table 1.1, Ti:Sa has a very broad emission bandwidth (from 650-1100 nm), and is usually pumped at around 532 nm. The broad emission bandwidth means that Ti:Sa can produce very short pulses (sub 50 fs is readily achievable) straight from the amplifier without additional spectral broadening needed. Ti:Sa lasers can also reach very high pulse energies >10 mJ - higher than Yb lasers.

The downsides of Ti:Sa lasers are that the repetition rates (and thus average powers) are limited due to the large quantum defect causing thermal damage in the crystal, and they are noisier than Yb lasers as they tend to be pumped by Q-switched nanosecond pulsed lasers

[14]Which is not problematic in itself, but anyone with experience working in laser labs will know that 99% of the problems lasers have are related to chiller faults and blockages.

rather than continuous-wave laser diodes[15]. In the author's experience, most amplified Ti:Sa lasers require more regular 'tweaking' by the user to maintain good output power than the equivalent Yb laser would (though a lot of this is down to the optical engineering rather than anything inherent in the gain medium).

Ytterbium Media

In contrast to Ti:Sa, Yb gain media have much narrower emission bandwidths (again see Table 1.1). This means that the 'straight from the amplifier' pulse duration is usually longer than a Ti:Sa output by a factor of 5-10. Amplified Yb lasers can achieve much higher repetition rates than amplified Ti:Sa lasers (up to MHz level), with a corresponding increase in average power despite having a lower pulse energy (generally less than 5 mJ)[16].

The repetition rates are achieved by the smaller quantum defect of Yb gain media, which are normally pumped in the near-IR at around 980 nm, enabling efficient pumping by high-power continuous wave laser diodes. Pumping with cw diodes is inherently a much less noisy process, and the resulting laser is much more stable (both shot-to-shot and in the long term) as a result. The advantages in rep rate and stability offered by Yb systems are often more significant experimentally than having the ultrashort pulses Ti:Sa systems produce[17]. You can collect more data per second with a higher repetition rate, and greater shot-to-shot stability reduces experimental noise. The differences between both gain media are summarised in Table 4.1.

Gain Medium	Pros	Cons
Ti:Sa	Short pulses, high pulse energy, high peak power.	Low repetition rate and average power, noisier, less stable.
Yb	High repetition rate and average power, quieter, more stable.	Longer pulses, lower pulse energies, lower peak powers.

Table 4.1: Crudely summarised pros and cons and Ti:Sa and Yb formats for ultrafast lasers.

History

For those with an interest in this sort of thing, historically (until around 2010), almost all solid-state ultrafast lasers sold were Ti:Sa lasers. The broad emission bandwidth made Ti:Sa an obvious choice of gain medium for ultrafast pulse generation, and it can easily be pumped by high-power frequency doubled Nd:YAG or Nd:YLF lasers, which have historically been a workhorse laser (especially within laser spectroscopy). Speculatively in the early days

[15] Because making high-power cw diodes that emit in the visible region is challenging, although recent developments have shown diode-pumped Ti:Sa lasers achieving relatively high output powers.

[16] Yb gain media can either be set up in *crystal* format (where beams are focussed from free-space into a crystal made of the gain medium, as Ti:Sa lasers are), or *fibre* format (where the beams are coupled into an optical fibre made of the gain medium). Fibre type lasers can achieve higher rep rates, but lower pulse energies – the inverse is true for crystal type lasers.

[17] Especially when you consider that, coupled with an OPA, an Yb system can produce pulses of a comparable duration to Ti:Sa.

of ultrafast science, achieving shorter and shorter pulse durations was a key driver and Ti:Sa is the best gain medium for this due to the broad emission bandwidth.

More recently, (from 2010 until 2024), the ultrafast laser market has shifted considerably towards Yb-based lasers. Conversations with people in the industry estimate that today, in 2024, for scientific applications, Yb-based lasers are outselling Ti:Sa lasers in almost a 2:1 ratio. The higher repetition rates, lower noise, and exceptional stability are the main drivers for this – as the number of applications of ultrafast lasers that truly need a sub 100 fs pulse duration and huge pulse energy is not very high.

Anecdotally, from my own field (ultrafast spectroscopy) this switch to Yb lasers is clear outside of specific applications where very short pulse durations and high intensity are truly needed (such as in strong-field physics). The advantages in rep rate (which increase the amount of data recorded per second, and increase the SNR in most experiments considerably), plus the lower noise and greater stability, generally outweigh the disadvantages of lower bandwidth and longer pulse duration (where they are disadvantages at all).

Bibliography

[1] T H Maiman. Stimulated optical radiation in ruby. *Nature*, 187, 1960.

[2] Hooker S and Webb C. *Laser Physics*. Oxford University Press, Oxford, 1st edition, 2010.

[3] Milonni P W and Eberly J H. *Laser Physics*. Wiley, New York, 1st edition, 2010.

[4] Weiner A M. *Ultrafast Optics*. Wiley, New York, 1st edition, 2010.

[5] Paschotta R P. *Field Guide to Lasers*. SPIE Press, Bellingham, Washington, 1st edition, 2008.

[6] Paschotta R P. *Field Guide to Laser Pulse Generation*. SPIE Press, Bellingham, Washington, 1st edition, 2008.

[7] Paschotta R P. The rp photonics encyclopedia. https://www.rp-photonics.com/encyclopedia.html. Accessed: 22-12-2020.

[8] Strickland D and Mourou G. Compression of amplified chirped optical pulses. *Optics Comm.*, 56, 1985.

[9] Light Conversion. Optics toolbox. toolbox.lightcon.com. Accessed: 23-12-2020.

Chapter 5

Characterising Ultrashort Pulses

In this chapter characterisation of ultrashort pulses and laser beams is discussed. An introduction to autocorrelation and FROG for temporal characterisation is given, with suggestions for further reading about other techniques. Methods for spatial characterisation of the beam waist are discussed, and finally the differences between pulse energy, fluence, and intensity are described.

Having generated some ultrashort pulses, the next obvious question to ask is *'how short are these pulses?'*. So, we need to characterise the pulses. Full characterisation an ultrashort pulse in a laser beam can be neatly split into three different areas.

- Temporal Characterisation.
- Spatial characterisation.
- Energy Characterisation.

The second of these three parameters is, strictly, not a characteristic of the *pulse* as much as of the *beam*, but these are the three things to do that will allow the output of your laser to be fully characterised. We will treat each of these areas in turn in this chapter, and discuss how we can manipulate the properties of the laser beam and pulses in the next.

5.1 Temporal Characterisation

Knowing how short your ultrashort pulse, measuring the **pulse duration**, is self-evidently an important piece of information to have when working with ultrashort pulses. If you do not know this, you cannot know if your pulse is actually ultrashort! *Temporal* characterisation of the pulse generally means two things:

- Measuring the pulse duration.

- Measuring the shape of the pulse (especially if it's weird, and not just a Gaussian).

In this context we will also discuss measuring the chirp of the pulse (the arrival time of each colour in the bandwidth).

Measuring the duration of ultrashort pulses is an academic field in its own right, and it will be impossible to give anything approaching a full and rigorous overview here. Instead and, we will focus on giving a more practical, qualitative description of various techniques to enable effective laboratory working. I have chosen to discuss two pulse measurement techniques in detail to illustrate how the problem of temporal characterisation is solved, but there are a great many other techniques possible, some of which are listed towards the end of this section for interested readers.

5.1.1 Measure *What?*

Measuring the pulse duration seems like a straightforward concept, but it is useful to define what precisely we need to measure. Ultimately, we would like to measure the full electric field $E(t)$ of our pulse. Note specifically that this is a function of time. $E(t)$ can be conveniently separated into an amplitude, and a term accounting for the phases of the different colours in the pulse. This can be written (in the time domain) as follows:

$$E(t) \propto \sqrt{I(t)} \exp\left[i(\omega_0 t - \phi(t))\right] \tag{5.1}$$

Where we have replaced the amplitude with the square root of the **intensity profile**, $\sqrt{I(t)}$, of the electric field. Note that $I(t)$ is **not** the same thing as the intensity *of* the laser beam as defined later in this chapter – people often say 'intensity' as a synonym for 'peak power'. In this context $I(t)$ is the function that describes how the intensity of the pulse varies with time – essentially defining our pulse envelope. Generally we would measure the intensity profile rather than the electric field amplitude directly[1], and taking the square root of the intensity profile gives our electric field amplitude. We will refer to this (time-dependent) intensity as the **intensity profile** going forward, to distinguish it from the intensity, or peak power, of the beam.

The second part of the Equation 5.1 contains the central frequency ω_0, and the phase $\phi(t)$. So, to fully characterise our pulse, we need to measure the intensity profile, the central frequency, and the phase. Measuring these quantities will be the topic of the remainder of this section.

5.1.2 Measuring the Central Frequency

Considering the three quantities we wanted to measure, we first consider measuring the **central frequency**. Measuring the central frequency is very simple, and just requires that you have a **spectrometer** that is sensitive to the wavelength of your laser. A spectrometer

[1] This is because we would be measuring the square modulus of $E(t)$ - multiplying by the complex conjugate such that the effect of the quickly oscillating carrier can be separated from the slowly varying amplitude.

generally is just a device that can tell you the different frequency/energy components of some light you shine into it[2] that you put into it. In our case, we want an **optical spectrometer**, which is a device that will disperse input light into different colours, and show you a *spectrum* of the light. Generally in a lab, this spectrometer consists of a box with an optical fibre attached to it that you can point at a piece of paper that your laser beam is hitting. Inside this box are some mirrors and a diffraction grating that separates the colours in the pulse and then images them onto a detector, so the amount of each colour in the pulse can be determined.

When you do this, you should see the spectrum of your laser pulse on the computer attached to the spectrometer. You will then easily be able to see the shape of the pulse (normally a nice Gaussian), and read off both the central frequency and the bandwidth of your pulse.

Spectrometers are essential pieces of kit in any laser lab. Being able to measure the bandwidth and spectral shape of the pulse gives you a good way to quickly diagnose potential issues with the laser output. For instance, if you suspect that the laser pulses are not as short as they should be, and you then measure the bandwidth and find out it is only 50% of what is expected, then you have a very probable culprit for the cause of your broad pulses. In fact, it is common and useful to have a spectrometer permanently monitoring the laser output, for this reason. Most lasers also have inbuilt spectrometers or other systems that monitor the spectrum of the oscillator[3], but it may or may not be possible to see this spectrum, depending on your laser. Anyway - the upshot is that usually it's very easy to measure the central frequency of your laser output, no bother.

5.1.3 Electronic Measurement?

Before continuing to discuss the different aspects of temporal pulse measurement, we should first address an common question. Especially if you come from a more physical (but non-ultrafast) background, it may seem a bit odd that we are making such a point of temporal pulse measurement. Why can't we just set up a photodiode[4], attach it to an oscilloscope, and read off the pulse duration straight from the oscilloscope?

This method works fine, provided that: (a) the time resolution of the photodiode (minimum duration it can measure) is shorter than the pulse; and (b) the time resolution of the oscilloscope is also shorter than the pulse. Unfortunately, these are problems. The fastest photodiodes available tend to have a rise time (i.e. how fast it responds - essentially the time resolution) of around 10 ps. Further to this, you would need an oscilloscope capable to

[2] Usually, but you're probably also aware of things like a 'mass spectrometer' where the input thing (a cloud of ions) is decomposed into different masses.

[3] Usually to avoid something called *continuous wave breakthrough* or *cw spiking* – this is when the oscillator stops being modelocked and starts operating in cw mode. This suddenly narrows the bandwidth of the pulses from the oscillator and prevents them from being correctly stretched before amplification, which can damage the laser amplifier. This phenomenon is seen in the spectrum as a big 'spike' at one frequency, hence the name.

[4] A light-sensitive diode which converts a light pulse into an electrical voltage pulse that can be read by an oscilloscope.

measuring this, which would mean an oscilloscope with a bandwidth of about 100 GHz. This oscilloscope doesn't exist, and high bandwidth oscilloscopes can be prohibitively expensive anyway. At the time of writing a 30 GHz oscilloscope from PicoTech retails for around £25,000 (and still wouldn't do the job we need). An alternative way to measure pulses is needed.

5.1.4 Measuring the Intensity Profile: Autocorrelation

The problem with electronic pulse measurement was that to measure any pulse, we really need to have an event that occurs on a similar timescale to the pulse. To take a real-world example, if I want to take a clear photograph of something moving quickly, like a sprinter sprinting, then I need a camera where the shutter speed is faster than the timescale on which the sprinter moves. If the camera shutter is open for longer than it takes the sprinter to run through the frame, then the image is blurry. If the camera shutter has opened and shut faster than the sprinter has moved appreciably in the frame, then the image is clear. So, we need to find an event which occurs on a timescale that is at least as fast as our pulse. This is problematic, as ultrashort laser pulses are some of the fastest events in the observable universe! The solution is a neat one, and is to **use the ultrashort pulse to measure itself.**

The fundamental process is one called **autocorrelation**. An autocorrelation fundamentally involves splitting the pulse in half and then delaying one half of the pulse relative to the other. The two pulses are then recombined in a way that produces an output signal pulse that *depends on their relative delay*. This recombination is done in a nonlinear medium, such as a BBO crystal. Then, one half of our original pulse can be scanned through the other, and the output signal from the recombination will vary as a function of the delay between the two pulses. Depending on the way the pulses were recombined, and how the the output signal was measured, we can extract a variety of information about the pulse. A schematic of a generic autocorrelator is shown in Figure 5.1.

The diagram in Figure 5.1 is generic enough that it will actually suffice for most of our discussion about measuring pulse durations. The key ideas of **using a short pulse to measure a short pulse** and **detecting the signal after mixing in a nonlinear medium** are common to the majority of ultrashort pulse measurement techniques. In the case shown in Figure 5.1, an input pulse is split and used to measure itself. One half of the pulse is delayed relative to the other, and then the pulses are overlapped spatially in a nonlinear crystal. The delay is scanned and when the pulses exactly overlap in time as well as space, then a second harmonic signal (blue) is generated. This signal is measured as a function of the delay between the two pulses. Note that the overlap in the crystal is deliberately non-collinear, so that the second-harmonic is produced between the two input beams, and can be easily spatially separated.

The crucial point here is that the output signal can be measured on a *slow detector*, such as a photodiode or spectrometer. This detector doesn't have to have sub-ps time resolution. The temporal information comes from the delay between the two halves of the initial pulse

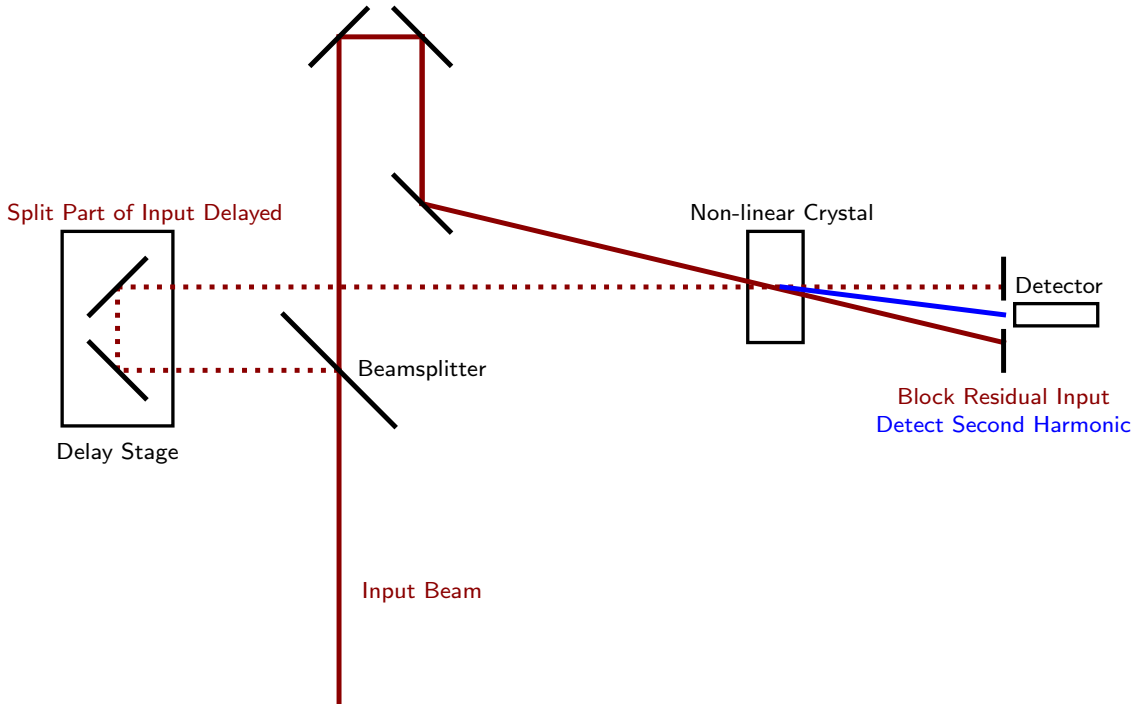


Figure 5.1: A schematic of a generic autocorrelator. Input beams (red) are split on a beam splitter, and part of the split input (dashed red) is delayed relative to the other half (solid red). The beams are recombined in a non-linear crystal, and the second harmonic (blue) is measured as a function of the delay between the split input pulses.

- and we *can* easily delay the two halves of the pulse in steps of a few femtoseconds[5]. By exploiting this, we can make our slow diode and oscilloscope or spectrometer measure our ultrafast pulse.

A final note is that autocorrelators generally function in either **scanning mode** or **single-shot mode**; as the name suggests, single-shot autocorrelation measures the entire autocorrelation signal using one laser pulse (one 'shot' from the laser). Scanning autocorrelation relies on the measurement of many pulses in the pulse train to build up the autocorrelation signal. Single-shot measurement removes the impact of any instability or jitter[6] that would impact the accuracy of the autocorrelation measurement. We will now discuss some of the different ways in which a autocorrelations can be performed, as well as some of their limitations.

[5] Note that 1 µm of difference in path length (which is easy to attain) corresponds to a time delay of around 3 fs.

[6] Jitter specifically refers to when pulses appear in a pulse train in a different place from where they would be expected to be in an ideal pulse train. However, it is commonly misused to just mean 'instability' generally.

Intensity Autocorrelation

The earliest (and arguably simplest) optical autocorrelation performed was an **intensity autocorrelation**, and a schematic of a setup to perform an intensity autocorrelation was shown in Figure 5.1, if we used a simple photodiode as a detector. An intensity autocorrelation combines the two halves of the pulse in a nonlinear crystal to produce the second harmonic of the input pulse. The efficiency of the SHG process is increased when both halves of the pulse arrive simultaneously (as the intensity is higher), and is reduced when both halves of the pulse arrive at different times. The second harmonic output signal, measured as a function of delay between the two pulses, τ, then has a distinctive intensity profile $I_{\text{out}}(\tau)$.

$$I_{\text{out}}(\tau) = \int_{-\infty}^{\infty} I(t)I(t-\tau)\mathrm{d}t \tag{5.2}$$

Equation 5.2 describes the **convolution** of the pulse with itself (see Appendix E for an explanation of what this means). The intensity profile of the input pulse $I(t)$ can be determined from $I_{\text{out}}(t)$, provided that we already know the shape of the input pulse (i.e. if it has a Gaussian , or other known temporal profile). If the input pulse has a Gaussian profile, then the FWHM width of the second harmonic intensity profile is $\sqrt{2}$ times longer than the FWHM width of the input pulse intensity profile. Therefore, by measuring the intensity profile of our second harmonic signal and dividing it by $\sqrt{2}$, we can find out the width of the intensity profile of our input pulse (if we know it is Gaussian). The width of the intensity profile is useful, and can give you a serviceable number for your pulse duration. However, there are many drawbacks and limitations to using intensity autocorrelation.

Firstly, it tells you the *intensity profile*, and does not measure the **phase** of the pulse, and therefore does not fully resolve the **electric field** of the pulse. Secondly, it requires that you assume the input pulses have a Gaussian (or Lorentzian, or sech^2, or other known) shape. This is quite a drastic assumption, as any noise around the pulse, or a weird shape, or instabilities in the pulse train can mean that the intensity autocorrelation trace is severely distorted[7]. These distortions can lead to significant errors in the final pulse duration determined.

Currently, there are much more powerful and better methods to measure the pulse duration, to the extent that many would say that intensity autocorrelation is an obsolete technique. However, the simplicity and relative inexpense of the commercially available intensity autocorrelators mean they are still widely used, despite these drawbacks. If you only want to get a serviceable number for your pulse duration (and not measure the chirp), and you know your pulse is a stable Gaussian, then intensity autocorrelation works fine.

Phase Information?

Throughout the previous discussion we have alluded to the lack of *phase information* in an intensity autocorrelation. There are other kinds of autocorrelation, like the **interferometric autocorrelation**, which do produce some phase information - however these still are not

[7] Though single-shot intensity autocorrelations negate some of these drawbacks.

entirely robust and do not provide a way to unambiguously retrieve the full electric field of the pulse. To do this, we need to measure the **spectral phase** of the pulse - i.e. the phase of each frequency component. Let us briefly recap why this is important.

In an ultrashort pulse, we know that if the spectral phase was a linear function (in time or frequency), then at some point all the colours in the pulse will arrive at the same time, and the pulse is transform limited. If the spectral phase contains higher order contributions, then the different colours in the pulse will never arrive at the same time, and the pulse is chirped and broadened. This was mathematically illustrated in both the time and frequency domains in subsection 2.3.1. Thus, if we know the form of the spectral phase, then we can tell how chirped (or not) our pulse is, and then we can unambiguously measure the pulse duration. Furthermore, if it is not transform-limited, know what kind of dispersion we need to add to make it transform-limited, because we know exactly how chirped it is.

Clearly, being able to measure the spectral phase of our pulse would be very powerful. It would mean we could measure the pulse duration *without* having to assume anything about the pulse shape; and secondly could tell if the pulse is transform-limited or not, and if it is not, measure the magnitude and sign of the chirp. This allows us to fully reconstruct the electric field of the pulse from our measurement. It turns out that performing this measurement is as straightforward as simply performing a **spectrally-resolved autocorrelation**, and one way of doing this is with **FROG** measurement.

5.1.5 Measuring the Spectral Phase: FROG

FROG stands for **Frequency Resolved Optical Gating**, and allows the full electric field of an input laser pulse to be measured. The FROG technique was developed by Rick Trebino [1, 2], and his company 'Swamp Optics' have an excellent website (www.swampoptics.com) that contains not only commercially available pulse measurement/compression equipment, but also a large number of extremely well written and accessible tutorials on the subject of ultrashort pulse measurement [3]. They are presented with a step up in mathematical complexity compared to the more qualitative descriptions here, so are an ideal next place to visit for interested readers.

As mentioned, at its heart, a FROG measurement is simply a spectrally-resolved autocorrelation. An existing autocorrelator can therefore be converted into a FROG by simply placing a spectrometer at the output rather than a photodiode. Our basic autocorrelator in Figure 5.1 becomes a FROG if we use a spectrometer as the detector. There are many possible alternative FROG geometries, based on using different non-linear processes for the measurement, which can have advantages in certain situations. These will not be discussed in detail here, but more information can be found in [1, 2].

Spectrograms and Phase Retrieval

An obvious question now is: *'how does a spectrally resolved autocorrelation allow the full electric field of the pulse to be measured?'* A spectrally resolved autocorrelation measures a **spectrogram** from which the electric field of the pulse can be retrieved. A spectrogram

is a way of representing how the frequency of a pulse varies with time, and is generally shown as a heatmap so that the intensity of various frequency components can be seen as a function of time too. Figure 5.2 shows three simulated spectrograms.

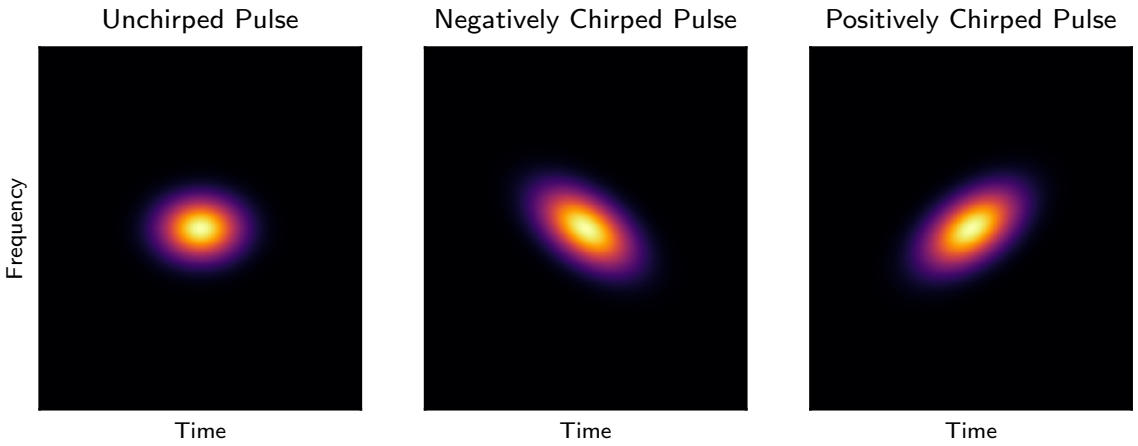


Figure 5.2: Simulated illustrative FROG spectrograms showing (from left to right) an unchirped pulse, a negatively chirped pulse, and a positively chirped pulse. Note that the two chirped pulses are slightly broadened by the chirp. These traces were simulated assuming a PG FROG, see text for details.

Spectrograms are a beautifully intuitive way to visualise your ultrashort pulse. We met the concept of a spectrogram previously in chapter 2 (Figure 2.3). In that case we simply plotted a 1D spectrogram, whereas in Figure 5.2 we have plotted the full 2D Gaussian distribution. Considering the leftmost spectrogram in Figure 5.2 (unchirped pulse), it is clear that as we travel through the pulse in time (from left to right), then the central frequency of the pulse does not change. This is because the pulse is transform limited. In the central and rightmost spectrograms, clearly the central frequency of the pulses change as we travel through the pulse in time, though at $t = 0$ the central frequency is the same as in the unchirped case (and so the 'average' central frequency that would be measured on a spectrometer is the same in all three cases, as the chirp is symmetrical). Note that, the rightmost two spectrograms in Figure 5.2 would actually look identical in a pure SHG FROG measurement (detecting the second harmonic of the input light), as in this measurement the direction of the time axis is not determined. This direction of time ambiguity can be removed by doing (for example) a polarisation-gated (PG) FROG measurement[2].

However, just having this spectrogram doesn't directly tell us the spectral phase, so we haven't answered the question we set out to answer. The way to extract the spectral phase from the spectrogram is to use a FROG inversion algorithm to reconstruct it. FROG inversion algorithms are widely available, and would come with a FROG if you buy a commercial one, but otherwise are available online. The reason that a FROG spectrogram allows the input pulse to be unambiguously determined is essentially because the sheer quantity of data contained in the pulse spectrogram makes it a quite overdetermined problem - there is a lot more data than is needed to solve the problem. This makes it comparatively easy

to reconstruct the full electric field from the FROG spectrogram. More details on how a FROG inversion algorithm actually works can be found in [1, 4, 5].

Now we have seen two ways to measure the pulse duration: an intensity autocorrelation, and a FROG measurement. The intensity autocorrelation allows us to measure the intensity profile of the pulse electric field, but does not measure the spectral phase. A FROG measurement measures the same things as an intensity autocorrelation, but also measures the spectral phase, so more completely characterises the pulse. There are also a wide range of different beam geometries for FROG measurements, each with their own advantages and disadvantages, but that enable FROG measurements to measure an extremely wide range of different input pulses. A full review of these is beyond the scope of what we do here, but looking into the tutorials from Swamp Optics [3] is an excellent starting point if you want to know more.

There are many other ways to measure ultrashort pulses besides FROG, and many tend to follow the FROG tradition of having zoological acronyms. Examples such as Spectral Interferometry for Direct Electric field Reconstruction (SPIDER) [6]; GRating-Eliminated No-nonsense Observation of Ultrafast Incident Laser Light Electric-Fields (GRENOUILLE) [7]; and Reconstruction of Attosecond Beating By Interference of two-photon Transitions (RABBIT) [8] all exist. Another notable technique is Multiphoton Intrapulse Interference Phase Scan (MIIPS) [9], which both measures and compensates for pulse dispersion using a pulse shaper. There are many review articles on ultrashort pulse measurement for interested readers[10]. You will find that the acronyms tend to get more forced as time progresses.

5.1.6 One Complication: Measuring UV Pulses

There are some instances where the aforementioned methods for pulse measurement won't work as straightforwardly as I have described here, though the general idea of 'scan two short pulses through each other in a nonlinear medium' is still basically valid. These cases are mostly exceptional cases that you are unlikely to encounter, however, one that you may encounter is when measurement of pulses in the UV region of the spectrum is desired.

The problem with trying to measure a UV pulse with either an intensity autocorrelation or a FROG lies in the nonlinear medium. If we were to take a UV pulse, split it in half, and then try to combine the two halves together in a nonlinear crystal, we would often find that we're trying to generate light that is either absorbed by the crystal, or simply not detectable by a spectrometer. For example, if we had a 266 nm pulse, using the conventional methods described we would end up trying to double the frequency and make a 133 nm pulse. This would be absorbed by the nonlinear crystal (if 266 nm was even within the phase-matching bandwidth), and we would not be able to record our autocorrelation/FROG measurement.

One solution to this is to perform a **cross-correlation** or a **cross-FROG**[11]. In this case, rather than splitting our unknown pulse on a beamsplitter and scanning it through itself, we simply scan a **known reference pulse** through our unknown pulse. The resulting measurement will be a convolution of the unknown pulse and the reference, so provided that we know the duration of our reference pulse, we can extract the duration of our unknown

pulse from the measurement. However, it is **very** important that the reference pulse comes from the same laser, as otherwise getting the reference pulse and unknown overlapped in time will be impossible – you also obviously need to know the duration of the reference pulse, by a standard FROG measurement (or equivalent).

Taking the previous example of measuring a 266 nm pulse as an example, we could mix this in a nonlinear crystal with an 800 nm pulse[8]. By selecting the right crystal and turning it to the correct angle for our polarisations, we can perform difference-frequency generation and make 400 nm light when the two pulses overlap in time. Measuring this 400 nm signal with either a diode or a spectrometer will perform a cross-correlation or cross-FROG measurement, which can then be used to extract the pulse duration of the unknown 266 nm pulse. Other ways to measure this pulse could be using either a transient-grating (TG) FROG, or a self-diffraction (SD) FROG [1]. All of these methods have their own advantages and disadvantages - there are many ways to skin a cat in ultrashort pulse measurement, and taking some time to research different methodologies is a good investment.

5.2 Spatial Characterisation

We now turn to the question of the **spatial characterisation** of the laser pulse (or laser beam). In section 1.4 we saw that for a Gaussian beam, the quality of the beam was defined using the beam parameter product (BPP), which was given by:

$$\text{BPP} = \rho_0 \theta \tag{5.3}$$

Where ρ_0 is the *beam waist*, and θ is the *divergence half-angle*. To recap, the beam waist tells us *how big* the laser beam is at a given point, while the divergence half-angle tells us *how well collimated* the beam is (higher divergence $\rightarrow$ less well collimated). We can fully[9] characterise our Gaussian beam by measuring these quantities, and can then calculate the quality factor M^2. How we practically can do this is the topic of the following sections.

5.2.1 Beam Waist

Let us start the discussion by considering how to measure ρ_0. This is a logical place to start, as to measure θ requires that we know ρ_0 in at least two locations anyway (see Equation 1.7), so once we can measure ρ_0, the rest becomes comparatively more straightforward. Note that a look in literature will reveal a large number of different ways to define the beam waist, but we talk about the $1/e^2$ width throughout.

How easily you can measure the beam waist depends a bit on how big your beam waist actually is, and how accurately you want to measure it. Are you trying to measure the

[8] As this is probably how you made the 800 nm pulse

[9] At least, as fully as we tend to need to in typical spectroscopy experiments. There are other Gaussian beam parameters like the **Gouy Phase** that can also be calculated, but these are not normally relevant except for in special circumstances. A friend of mine once designed and built an experiment to measure the Gouy phase and described it as 'utterly pointless'.

approximate size of the beam at the laser output to check that the laser is performing to specification? Or are you trying to accurately find out the beam waist of a small focussed beam so you can accurately report the irradiated area of a sample when writing a paper? In the first case, simply taking a card marked with millimetre measurements and looking at the beam on it can do a reasonably good job; but the second case is somewhat more involved, and is what we will focus on for the remainder of this section.

The simplest way to measure the size of a laser beam is to use a **beam-profiling camera**. A beam-profiling camera is just like webcam that you shine a (heavily attenuated) laser beam on, and the camera software performs a Gaussian fit to the beam spot on the camera. This can give you a fast and accurate readout of the beam waist, and you can even get a 'live' view of the beam so can monitor the size of the beam as you move a lens back and forth, or change some other optical parameters. Beam-profiling cameras with user-friendly software interfaces are commercially available and lots of companies offer good quality solutions, but it is also perfectly possible to build a beam-profiling camera using an inexpensive commercial webcam, and this has been done in many laboratories. A motivated undergraduate project student of mine made an excellent one out of a Raspberry Pi in a few months, so it is possible. A good place to start when doing this is [12, 13], as there are some subtleties to take care of when using a webcam rather than commercial profiler.

Beam-profiling cameras generally come (or are built with) mountings such that they can be mounted on an XYZ translation stage, so that the camera can be moved in 3D to allow the beam to be easily centered on the sensor. Being able to move the camera along Z (i.e. along the propagation direction of the beam) is also advantageous when measuring a focussed beam, as then you can easily ensure that the sensor is placed at the exact focus by moving the camera along Z to the point where the beam waist is smallest. As a final note, it should go without saying that an intense laser beam can easily destroy a delicate CCD camera chip, so the incident beam will generally require heavy attenuation. Attenuation by around 12 orders of magnitude is typical!

The downside of a beam-profiling camera is that if your beam diameter is very small (typically less than around 35 µm), then the beam won't illuminate enough pixels on the sensor to allow a reliable Gaussian fit to the size. In this case, an option is to use a **scanning-slit beam profiler** or **knife-edge beam profiler**. Both of these designs work by scanning either a small slit, pinhole, or sharp edge through the beam with micrometer precision. Then the transmission behind the slit/edge is measured and recorded. A Gaussian function can then be fit to a plot of the transmission against slit/edge position, and the beam waist easily read from the fit. The measurement takes slightly longer than when using as camera, but it allows smaller beams to be measured.

Building a homemade scanning slit beam profiler as can be a relatively inexpensive and easy way to get a functioning beam profiler set up in a lab. The basic principle is to have a photodiode mounted on an XYZ translation stage, and then to screw a 5 µm or 10 µm slit onto it in a rotation mount. Finding the beam in the slit, and then scanning through both the X and Y directions (by rotating the slit 90°) allows the beam waist in each of these

directions to be determined in the way described previously. This method is slow compared to using a beam profiling camera, and having to move along Z (to find the focus) and measure X and Y at each position can get rather tedious. An even faster way to get a rough idea of the size of a beam (especially useful with unfocussed beams), is simply to use a micrometer stage to scan a razor blade through the beam in front of a power meter. Monitoring the laser power as you scan the blade will allow you to obtain a rough idea of the beam size, but it is possible to burn the razor blade if you do this with a focussed beam!

5.2.2 Divergence

Having measured the beam waist, measuring the divergence half-angle is comparatively straightforward. It simply requires that we measure the beam waist at several different locations along our beam, and then use Equation 1.7 to calculate the divergence half-angle. Some commercial beam profilers are designed to be able to do this measurement automatically, so all you have to do is couple the beam into the profiler and it will measure the beam waist and divergence angle, and give you back the BPP and M^2 factor.

If you do not have a commercial instrument that can do this, then the measurement is still relatively simple, but may take some time. How long it takes depends on the Rayleigh range of the beam at the point where you are interested in it. If you want to know the divergence of the beam after focussing, then generally the Rayleigh range will be rather short (on the order of millimetres) and you can easily move your profiling camera/scanning-slit along the propagation direction of the beam using a translation stage. In this case, the measurement is comparatively fast. What can take a lot of time is if you are interested in the divergence of the beam over several metres (for example, to see how well collimated the laser output is). Then you will need to set the profiler up to measure the beam waist in several different locations that could be a metre or so apart. This is not so bad if you are using a profiling camera, but if you need to set up a homemade scanning-pinhole profiler at each location it can quickly get rather tedious!

Thankfully, we often don't actually need to calculate the full M^2 factor, so don't need to calculate the divergence half-angle. Generally speaking in spectroscopy, we are often only really interested in the beam waist at our experimental sample, and measurement of this is sufficient to characterise the laser conditions of our experiment. The author has only calculated the full M^2 factor once - and then only when trying to diagnose a problem with a laser, rather than for an actual experiment.

5.3 Energy Characterisation

Having characterised the temporal and spatial shapes of our laser pulse, the final piece of the puzzle is to characterise the **energy** of our pulse. By 'energy' here we specifically mean the energy in the pulse, not the photon energy (wavelength) of the pulse. We actually discussed this in chapter 1, but will repeat some salient points of the discussion here for ease of reference.

5.3.1 Energy and Power

To measure the **pulse energy** in your laser beam (the amount of energy within a single pulse), all you need is a power meter and to know the repetition rate of your laser system. The power meter will measure the **average power** (often simply 'power') coming out of your laser (in Watts). The average power is the rate of energy flow from the laser per unit time, and is defined as the energy in a single pulse multiplied by the number of pulses produced per second (the repetition rate). The pulse energy is therefore given by:

$$\text{Pulse Energy (J)} = \frac{\text{Average Power (W)}}{\text{Repetition Rate (Hz)}} \tag{5.4}$$

We can measure the average power using a power meter, and will generally know the repetition rate of our laser system. For example, if a laser operates at 1 kHz (1000 pulses per second), and the average power is 5 W (5 Joules per second), then each pulse contains 5 mJ of energy. 5 mJ may not sound like very much energy, but when that energy all arrives in a few tens of femtoseconds then it can pack a considerable punch! Knowing the amount of energy in a single pulse also allows you to calculate the number of photons in the pulse, provided you know the central wavelength of the laser (the average energy of a single photon). The pulse energy is simply the energy of an individual photon multiplied by the number of photons present. By knowing the repetition rate, you can then easily calculate the number of photons irradiating a target per second - the **photon flux**.

5.3.2 Peak Power, Fluence and Intensity

In the previous section we showed how to calculate the pulse energy from the measured average power. However, to fully characterise our laser output, we need to account for:

1. The duration of the laser pulse, as the pulse energy arrives in a very short window defined by the pulse duration.

2. The physical size of the laser beam at the point where we want to characterise it, as clearly if the beam is concentrated into a smaller area then it will be more intense.

The first of these points can be accounted for by defining a related quantity to the average power, the **peak power**. Where the average power was defined as the rate of energy flow per unit time the peak power is defined as the rate of energy flow per individual laser pulse:

$$\text{Peak Power (W)} = \mathcal{F} \times \frac{\text{Pulse Energy (J)}}{\text{Pulse Duration (s)}} \tag{5.5}$$

Where $\mathcal{F}$ is a factor that depends on the temporal shape of your pulse, and is around 0.94 for Gaussian pulses. Using Equation 5.5 to calculate the peak power then requires that you have measured (or at least have a rough idea of) your pulse duration. If the laser used in the example above produced pulses that were 35 fs long, then the peak power in each pulse is around 140 GW! These peak powers are incredibly high, and this is one of the things that makes femtosecond lasers so useful in the natural sciences.

Looking back to our initial considerations, we can account for the second point by introducing a quantity called the **fluence** of the laser pulse. The fluence is for a Gaussian pulse defined as the energy per unit area:

$$\text{Fluence (J m}^{-2}) = \frac{2 \times \text{Pulse Energy (J)}}{\text{Irradiated Area (m}^2)} \tag{5.6}$$

The factor of two in the numerator arises from the fact that we are measuring a Gaussian pulse [14, 15]. Equation 5.6 requires that you have measured the area of the laser beam at the point of interest (as discussed in section 5.2). If our laser beam is (at least approximately) circular, then generally we can define the irradiated area as simply $\pi\rho_0^2$ where ρ_0 is the measured beam waist. This then allows the fluence to be defined, and if we take the same example system as previously, and assume that our beam is focussed to a spot with beam waist 35 µm, then the fluence can be calculated to be 2.6 MJ m^{-2}.

Finally, we can go one step further and incorporate both of our initial considerations into one quantity by defining the **intensity** of the beam. The intensity is defined in units of power per unit area, and so requires that we know both the pulse duration and the beam waist (or irradiated area). It can then be calculated in a variety of ways, such as as the fluence per unit time, or the peak power per unit area. The intensity is generally the number you need when you're writing a paper and want to document the laser conditions, and can be expressed as:

$$\text{Intensity (W m}^{-2}) = \mathcal{F} \times \frac{\text{Fluence (J m}^{-2})}{\text{Pulse Duration (s)}} \tag{5.7}$$

Where again the additonal factor $\mathcal{F}$ depends on the pulse shape and is 0.94 for a Gaussian pulse, as in Equation 5.5. Equation 5.7 gives a straightforward way to calculate the intensity of our laser beam. Useful calculators for a lot of these parameters can be found on the Optics Toolbox [16], and these also give the relevant equations in many cases.

We have now discussed how ultrashort pulses are generated, and how they are fully characterised. Accordingly, we will turn our attention to what happens in between: how do we manipulate our laser pulses so they look and behave as we desire? This requires that you build up a **beamline** from your laser output to your experiment. Doing this is the topic of the remainder of this book.

Bibliography

[1] Trebino R. *Frequency-Resolved Optical Gating: The Measurement of Ultrashort Laser Pulses*. Springer Science and Business Media, Berlin, 1st edition, 2000.

[2] Trebino R, DeLong K W, Fittinghoff D N, Sweetser J N, Krumbugel M A, and Richman B A. Measuring ultrashort laser pulses in the time-frequency domain using frequency-resolved optical gating. *Rev. Sci. Inst.*, 68, 1997.

[3] Trebino R. Swamp optics tutorials. http://www.swampoptics.com/tutorials.html. Accessed: 23-12-2020.

[4] Kane D J. Recent progress toward real-time measurement of ultrashort laser pulses. *IEEE Journal of Quantum Electronics*, 35, 1999.

[5] Kane D J. Principal components generalized projections: a review. *J. Opt. Soc. Am. B*, 25, 2008.

[6] Iaconis C and Walmsley I A. Spectral phase interferometry for direct electric-field reconstruction of ultrashort optical pulses. *Optics Letters*, 23, 1998.

[7] O'Shea P, Kimmel M, Gu X, and Trebino R. Highly simplified device for ultrashort-pulse measurement. *Optics Letters*, 26, 2001.

[8] Paul P M, Toma E S, Breger P, Mullot G, Auge F, Balcou P H, Muller H G, and Agostini P. Observation of a train of attosecond pulses from high harmonic generation. *Science*, 292, 2001.

[9] Lozovoy V V, Pastirk I, and Dantus M. Multiphoton intrapulse interference. iv. ultrashort laser pulse spectral phase characterization and compensation. *Optics Letters*, 29, 2004.

[10] Walmsley I A and Dorrer C. Characterization of ultrashort electromagnetic pulses. *Advances in Optics and Photonics*, 1, 2009.

[11] Linden S, Kuhl J, and Giessen H. Amplitude and phase characterization of weak blue ultrashort pulses by downconversion. *Optics Letters*, 24, 1999.

[12] Langer G, Hochreiner A, Burgholzer P, and Berer T. A webcam in bayer-mode as a light beam profiler for the near infra-red. *Optics and Lasers in Engineering*, 51, 2013.

[13] Keaveney J. Automated translating beam profiler for in-situ laser beam spot-size and focal position measurements. *arXiv*, 2018.

[14] Newport Optics. Newport optics tutorials. https://www.newport.com/resourceListing/tutorials. Accessed: 23-12-2020.

[15] Paschotta R P. The rp photonics encyclopedia. https://www.rp-photonics.com/encyclopedia.html. Accessed: 22-12-2020.

[16] Light Conversion. Optics toolbox. toolbox.lightcon.com. Accessed: 23-12-2020.

Chapter 6

Manipulating Ultrashort Pulses

Abstract

In this chapter, a discussion of how we can manipulate laser beams, and an overview of the common types of optical element used to do so, is presented. The structure is to first discuss how we can manipulate particular properties of the beam, followed by a description of the optics used to do it. The chapter is aimed at the new student who has no prior experience of optical building. General considerations with substrates and coatings are also discussed, and then many different types of reflective, transmissive, and refractive optics are discussed. Tips for buying optics and optomechanics are presented throughout, and the focus is on providing information that will enable effective lab working.

In the second part of this book we aim to present a guide for construction of optical setups and beamlines using ultrafast pulses. To do this, we need to use **optics** or **optical elements** to **manipulate** our beam in different ways, so that the laser can be used to perform the science we want to do. This chapter discusses how we can perform these manipulations using optics, together with discussion of different kinds of optics and how to choose what you need. chapter 7) will discuss more of the practical aspects of optical building, using the foundations described here.

Coming into a laboratory for the first time and seeing a beamline can be rather daunting, but as we will see, a lot of the individual elements are rather straightforward and just serve to manipulate properties of the light that we have already discussed. We will start by talking about how manipulate various properties of the beam, giving only brief overviews of how the various optics actually work, and then in section 6.6 talk more generally about optics and optomechanics.

6.1 Spatial Manipulation

Perhaps the most fundamental bit of optical construction you'll do is manipulating *where* the beam is, and *how* it looks, in **space**. Firstly, we will talk about how to manipulate where your beam is – **beam steering**.

6.1.1 Beam Steering

Beam steering is the act of moving a laser beam from one place to another – you *steer* the beam through the optics. For example, in a typical experiment we would need to direct a laser beam onto a sample – and it's unlikely that you can just place the laser in front of the sample. In fact, even if we could do this in theory[1], we wouldn't want to because we would have no way to manually manipulate the direction in which the beam points, or control the power and polarisation state of the beam. Most lasers don't have absolutely perfect *pointing stability*, so beams will drift over time, meaning that we need a way to steer the beam back to where it should be once it has drifted. This is shown in Figure 6.1 for a simple case. A detailed description of how to practically steer a beam through a setup is given in chapter 7.

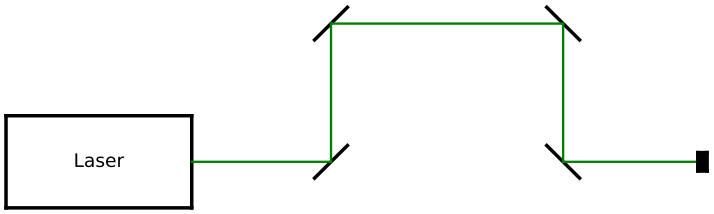


Figure 6.1: A simple example of a beam (green) from a laser being steered onto a small sample (small box) using mirrors (angled lines).

Essentially all beam steering within ultrafast optics is accomplished using **mirrors**. A **mirror**, simply, is something that will reflect any incoming light beam in a **single** direction. Most objects will reflect light to some extent – this is why you can see things – but most objects reflect light in a disordered and incoherent way, known as **diffuse reflection**. Diffuse reflection is characterised by a well-defined incoming light beam being scattered into many different directions. In contrast, a good mirror will only reflect the incoming light beam

[1] Very unlikely, because one of the beam power, wavelength, size, and polarisation would not be what is required.

in a single direction, which depends on the angle at which the incoming beam hits the mirror (the *angle of incidence*). This is known as a **specular reflection**. The difference is illustrated in Figure 6.2.

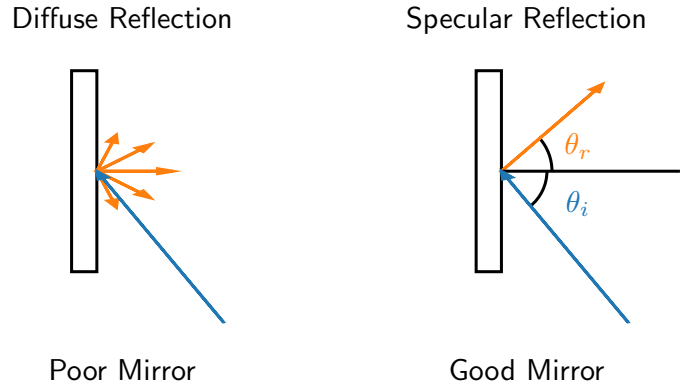


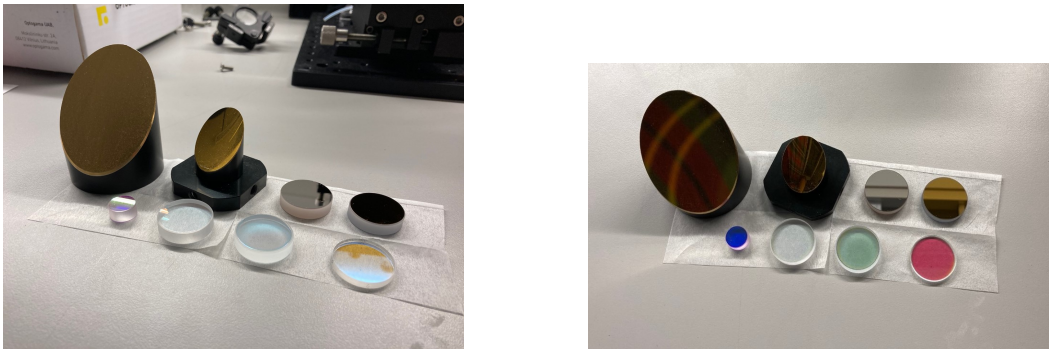
Figure 6.2: Illustration of the difference between specular and diffuse reflection. The incoming beam is shown in blue, and the reflected beam(s) in orange. The angle of incidence θ_i is equal to the angle of reflection θ_r.

So, a good mirror will reflect the incoming light out in a single direction. This is quantified by simply defining the **angle of incidence** (θ_i) and **angle of reflection** (θ_r). These are defined as the angle between the light beam and line orthogonal to the surface of the mirror. The *law of reflection* states that these two angles are the same - so the angle of incidence is equal to the angle of reflection ($\theta_i = \theta_r$). However, the **reflectivity** of a mirror can depend on the angle of incidence, amongst other things such as wavelength and polarisation state. The reflectivity is the ratio of the reflected power to the incident power:

$$\text{Reflectivity} = \frac{\text{Reflected Power}}{\text{Incident Power}} \tag{6.1}$$

And is usually expressed as a percentage. A 99% reflective mirror therefore reflects 99% of the optical power incident on it. Mirror reflectivity is usually greater than 95%, and can be as high as 99.99% for a specialised highly reflective mirror.

Optical mirrors can be made of a wide variety of materials, which all have different properties. The mirror substrate has some bearing on the performance of the mirror in some applications, as discussed in section 6.6. The mirror **coating**, however, does the bulk of the 'work' in ensuring that the mirror actually reflects the desired light. Figure 6.3 shows a variety of different optical mirrors with different coatings. The mirror coating is what turns the substrate from a flat piece of glass into a reflective mirror. There are two main kinds of mirror coatings, **metallic coatings** and **dielectric coatings**.



(a) (b)

Figure 6.3: Different kinds of commonly encountered mirrors. Clockwise from top left: 2 inch gold parabolic mirror; 1 inch gold parabolic mirror; 1 inch protected silver mirror; 1 inch protected gold mirror; 0.5 inch 635nm dielectric mirror; 1 inch ultrafast 800nm mirror; 1 inch broadband dielectric mirror; 1 inch thin harmonic separator. Note that some mirrors look different depending on the angle of incidence the photo is taken at - illustrating that often the reflectivity depends strongly on this.

Metallic Mirrors

Metallic mirror coatings will be most familiar, as these are how most mirrors you see in everyday life (in bathrooms, for example) are made. The reflective coating is simply a thin layer of metal – the three most common types are aluminium, silver, and gold – deposited onto a substrate. Metallic mirrors reflect a wide range of wavelengths[2], and are relatively insensitive to both the angle of incidence and the polarisation of the incoming beam. They also add very little GDD to ultrashort pulses (see next section), so can be used in femtosecond applications. Metallic mirrors are also relatively inexpensive, so can be a cost-effective solution.

However, the metallic coating is relatively soft and is very susceptible to damage. The surface is easily scratched (even when cleaning), and they are generally less reflective than other kinds of mirrors, and burn more easily[3]. Keeping the mirrors scrupulously clean will minimise the chance of this happening, as burning tends to happen when dirt on the mirror surface absorbs the light and burns, creating a scorch on the mirror. Whilst metallic mirrors have a very large bandwidth, there are differences between the three common metallic coatings, which are roughly summarised in Table 6.1.

Within the broad categories of 'aluminium', 'silver', and 'gold', many manufacturers produce subsidiary coatings which are very useful. Chiefly, **UV-enhanced aluminium** is a very efficient and cost-effective way to get a relatively broadband mirror with good reflectivity in the UV region (especially useful with tunable UV sources such as OPAs and dye lasers). UV-

[2]The **bandwidth** of the mirror is large.

[3]All the optical power that is not reflected by the mirror has to go somewhere – usually it will be scattered or dissipated into the mirror itself. The power that is *not* reflected is all power that could end up burning the mirror.

Table 6.1: Metallic mirror coatings and their properties.

Coating	Useful Reflectivity Range	Comments
Aluminium	400 - 600 nm, 1- 20 µm	Best in UV with extra coating. Poor around 800 nm.
Silver	500 - 10 µm	Best in the near IR. Poor in UV and far IR.
Gold	800 - 20 µm	Best for further into the IR.

enhanced aluminium extends the reflectivity of normal Al mirrors down to around 250 nm and below. **Ultrafast-enhanced silver** is also a very useful coating - as it is engineered to produce very low GDD around the 800 nm region, ideal for use with femtosecond Ti:Sapphire lasers. Other coatings designed to enhance reflectivity in other regions, or to improve the damage thresholds, are also available - but note that these overcoatings can also affect other aspects of mirror performance (such as GDD) in a negative way. A quick search around different manufacturers will reveal a wealth of different metallic coatings for many applications.

Dielectric Mirrors

Dielectric mirrors are mirrors where the reflective surface is made up of many thin layers of a dielectric material (such as glass - illustrated later on in Figure 6.20). They have higher reflectivity and are less easily damaged than metallic mirrors, however they are more expensive and for ultrafast use do not reflect as broad a bandwidth as metallic mirrors[4]. You will find that dielectric mirrors are explicitly marked with the wavelength range they are designed to reflect. Broadband dielectric mirrors tend to have quite poor GDD performance unless they are explicitly marketed as 'ultrafast' or 'low GDD', so it is worth checking documentation before buying.

To understand the need to think about the GDD added by a mirror, recall that an ultrashort pulse has a very broad bandwidth, and to reflect the pulse effectively a mirror must reflect the whole bandwidth of the pulse. Some dielectric mirrors are very narrowband, designed for use with narrowband lasers, and these should be avoided for ultrafast applications. Generally you will find mirrors marketed as **'broadband mirrors'** that are designed for this purpose. In addition, the layered construction of the mirror can cause pulse dispersion. Each colour in the pulse will pass through a different amount of the coating when it is reflected - so the different colours will pass through different amounts of material and so the pulse will broaden in time. Dielectric mirrors marketed as **'ultrafast enhanced'** or **'ultrafast mirrors'** are engineered to minimise/negate this effect, and are well worth the additional cost if you want your femtosecond pulse to still be a femtosecond pulse after reflection! The golden rule is that unless the mirror is explicitly marketed for femtosecond/ultrafast use, assume that it is not suitable.

Dielectric mirrors are also more sensitive to both angle of incidence (AOI) and input po-

[4]This can be a benefit and a drawback, sometimes you would like to only reflect a certain colour and dump the rest - a dielectric mirror can be ideal for selectively reflecting the colour you want.

larisation than metallic mirrors. For general lab use, we normally require mirrors that work well at a 45° AOI - resulting in a 90° angle between the incoming and outgoing beams; as well as mirrors that work well at a very small, ∼0°, AOI - which reflect the incoming beam back along the direction of the input beam. 45° mirrors were shown in Figure 6.1.

Dielectric mirrors will be marked as either 45° or 0° depending on the AOI. 45° mirrors are the most common, and generally you will find that mirrors with a 0° AOI are explicitly marketed as **'zero-degree mirrors'**. Regarding input polarisation, you will generally find that in a well designed 'ultrafast' mirror, the reflectivity in the useful region around the central wavelength does not enormously vary with polarisation. However, as you get further from the designed input wavelength then polarisation becomes more important. The added GDD also depends on the input polarisation, but again only appreciably when you are far from the designed input wavelength. All manufacturers give reflectance and GDD curves for various AOI and polarisations readily online - it is worth checking these before purchasing to check they will work for your application. Manufacturers will also make more data available on request if it is not readily given - such as the transmission of a coating at a wavelength far from it's designed reflective wavelength.

Mirror Coating Choice

Generally speaking metallic coatings are cheaper than dielectric coatings, but dielectric coatings are **much** more durable. Within dielectric coatings, cost can vary widely depending on the type of coating. As a general rule, coatings which have broader bandwidths, lower GDD, and are designed to reflect a larger number of wavelengths are more expensive than narrowband coatings. Additionally, common coatings are made in bulk and are cheaper - so coatings for use with typical laser output wavelengths (and their harmonics) are often affordable. Companies such as layertec have extensive catalogues of custom coatings for bespoke applications - but these are often much more expensive (as the cost of starting a production run of a bespoke coating is high). For example, at the time of writing a typical 1 inch 800 nm ultrafast mirror from a standard manufacturer costs around £100. In contrast, a more complex coating (HR 266 nm, 400 nm, and 800 nm with low GDD and >99% reflectivity at all wavelengths) costs closer to £400.

In very high power applications (such as low repetition rate beamlines driving high-harmonic generation), then the power in the beam may be close to the damage threshold of the mirror. To avoid damage, one approach is to expand the beam using curved mirrors such that the power in the beam is spread over a larger area[5]. Sometimes this is not practical, if a specific beam size is desired for a certain application. In this case, it can be more practical to either use a cheap 'economy' mirror that can be easily replaced, or to periodically rotate the mirror in it's mounting such that the beam is not incident on one spot for too long.

[5] The enormously high energy beams at the National Ignition Facility in the US are reflected off very large (∼50 cm^2) mirrors!

Beamsplitters

A **beamsplitter** is an a special kind of mirror which does exactly what you expect it to. When a beam is incident on a beamsplitter, part of the beam is reflected off the surface of the beamsplitter, and the rest of the beam is transmitted through it - it's a partially reflective mirror.

The most simple kind of beamsplitter will split light in a predictable ratio. The ratio of the reflected to transmitted light is called the **splitting ratio**, and is commonly quoted as numbers as 'R:T' - such as '50:50' or '30:70' (30% reflected, 70% transmitted). This functions as would be expected: sending 2 W of light onto a 50:50 beamsplitter will result in 1 W being reflected and 1 W being transmitted. Generally a beamsplitter will have a front surface that is coated to give partial reflectivity, and a rear surface which is anti-reflection coated. This makes sure that the only reflection you get is from the front surface, rather than there being two reflections from the front and back surface.

The most standard type of beamsplitter has a splitting ratio that is independent of both the polarisation state of the incoming beam, and the colour of the incoming beam[6]. In contrast, a **polarising beamsplitter** is one that will reflect one polarisation state (s or p) whilst transmitting the other, these will be discussed further in section 6.3. A common use of beamsplitters is to split down the full output power of a laser into more usable fractions, as shown in Example 6.1.

Dichroic Beamsplitters/Mirrors

An often important type of beamsplitter for people doing spectroscopy is the **dichroic beamsplitter**. A dichroic beamsplitter is one that will reflect one colour of light whilst transmitting another, and it is really another word for a dichroic mirror. There are many different types, but of particular use are ones designed to separate harmonics of a laser output from the fundamental. These are often marketed as 'harmonic separators', and are designed to (for example) separate the 400 nm Ti:Sa second harmonic from the 800 nm fundamental (see Example 6.2 for an example use case). Harmonic separators are available in a very wide range of configurations to suit all kinds of common laser applications. The key things to make sure of when buying dichroic beamsplitters for an ultrafast application are firstly that the coating has a broad enough bandwidth (as usual); and secondly that the substrate is thin enough to not broaden the transmitted pulse extensively.

6.1.2 Beam Expansion and Reduction

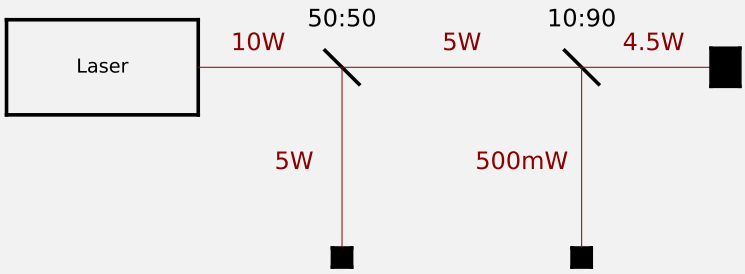
Another way we can manipulate our beam in space is to change the *size* of the beam – manipulating the beam radius, ρ. We might want to do this for a few reasons, commonly:

- To focus our light tightly onto a sample, to ensure that we maximise the interaction between the light and the sample.

[6] Provided that the colour is within the bandwidth that the coating reflects.

Application 6.1: Splitting Laser Outputs

Standard beamsplitters are commonly used directly infront of laser outputs, when it is known that the output of one laser system needs to be shared in some fixed proportion between multiple end users. A common example of this is when the output of a laser system is used to pump an external OPA, or other frequency conversion stage. Such non-linear frequency conversion stages are very sensitive to input power, so having fixed beamsplitters directing the beams to them prevents the input power being erroneously adjusted up or down. The only drift possible in this scenario is if the laser system output power drifts, and this would be quickly spotted and corrected in a working lab.



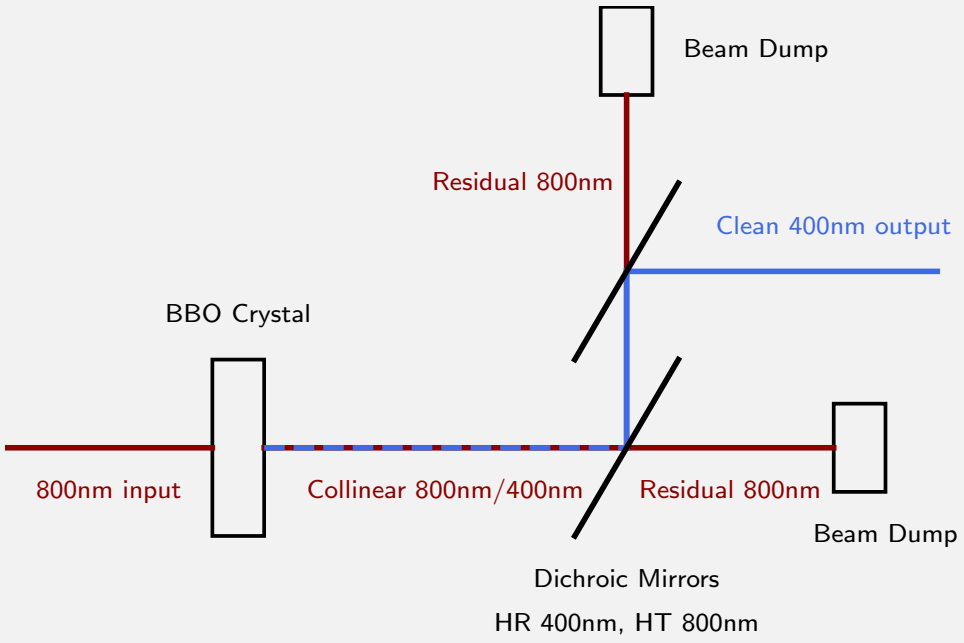
- To shrink our beam so that it fits onto the optics we are using.
- To expand our beam so that the power in it is dispersed over a larger area.

Another key area where you'd want to do a lot of beam expansion/reduction is in **optical imaging**. If you are working in a field like **multiphoton microscopy** where you use ultrafast lasers to illuminate samples and image them, then beam expansion and reduction is a crucial, but you also need to think more carefully about things like **aberrations** – some of this will be discussed later.

In general the output of our lasers will be a beam of light which is well **collimated**. This means that it has a very low divergence angle, or that it doesn't spread out from it's initial well directed beam. If we want to change the size of the beam, we need to change this divergence angle so that the beam gets smaller or larger as it propagates. In these cases we need an optical element to change the divergence angle - a **focussing optic**.

Application 6.2: Harmonic Separation

Following non-linear frequency conversion, such as second-harmonic generation, the beam will contain light of both the desired harmonic and of the input fundamental (as the conversion is generally at most around 40% efficient). A convenient way to separate out the harmonic is to reflect it off a series of dichroic beamsplitters/mirrors. This is illustrated for a simple case in the figure below, but can easily be extended to more complex systems in an analogous way.



Lenses

The most common kind of focussing optic is well known in everyday life, a **lens**. A lens is a transmissive optic, and is curved such that as light passes through it, it undergoes refraction and the divergence of the beam changes. It is possible to make lenses that are both **converging** (focussing the beam down to a point), and **diverging** (expanding the beam). Lenses come in a variety of different shapes, but for now we will consider the case where one side of the lens is flat, and the other side curved. These are called **plano-convex** or **plano-concave** lenses depending on whether or not the curved side bulges outwards away from the centre of the lens (convex), or caves back in towards the centre (concave)[7]. The main parameter that defines a lens is the **focal length**. The focal length, f, is defined as the distance from the lens that an incoming perfectly collimated beam will be focussed to - and is illustrated in Figure 6.4. Lenses with short focal lengths are often referred to as 'strong' lenses, as they focus the light beam very hard and to a very small point (and vice

[7] Lenses with one flat side are most often used when the beams entering and leaving the focussing system are well collimated, and this is most often the case in (at least) the majority of experiments encountered by the author. Other lenses may be more suitable for your application - particularly if you are using the lenses to perform imaging rather than to just manipulate beam divergence.

versa for longer focal length lenses)[8]. A diverging lens has a negative focal length, defined as the location that an image *would be formed at* if the outgoing beam was extrapolated back behind the lens. This is sometimes called a *virtual focus* or *imaginary focus*. Lenses made of material of a higher refractive index will be more powerful, and lenses that are more strongly curved will also be more powerful.

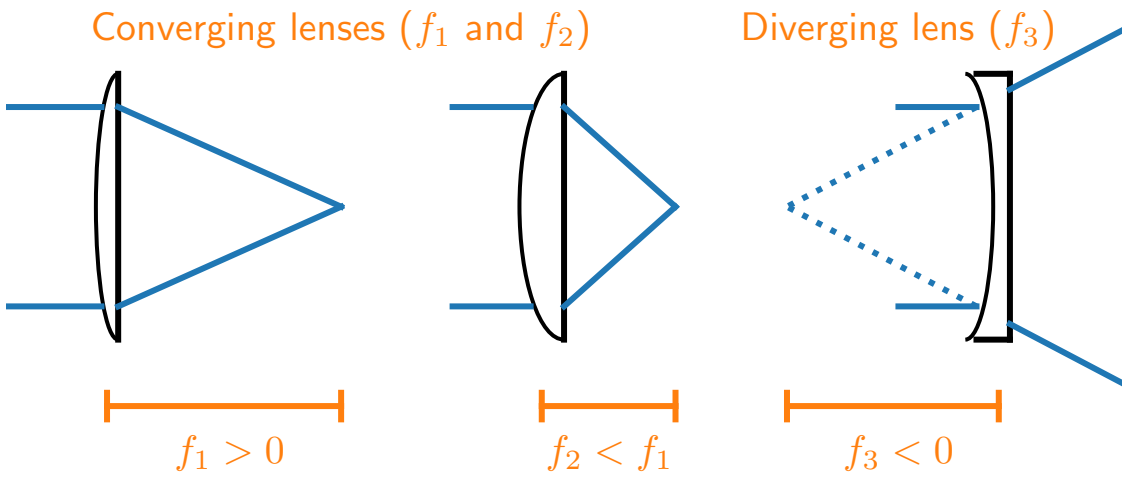


Figure 6.4: Illustration of three different focussing lenses. From left to right: a converging lens with positive focal length f_1; a stronger converging lens with positive focal length $f_2 < f_1$; a diverging lens with negative focal length f_3. In all cases collimated light enters the lens on the left and leaves on the right. Dashed lines on the diverging lens show the position of the virtual focus. Note that collimated light hits the curved side of the lens first - this is to minimise spherical aberration.

The lenses shown in Figure 6.4 are all **singlet lenses**, which are just single optics with two external faces polished or shaped. More complex lenses (**compound** lenses) can be formed by stacking multiple singlet lenses together, to make **doublets** (two lenses), or even more complicated **objectives** which can contain a number of individual singlet lenses. The purpose of compound lenses is to reduce aberrations – things that make the focussing behaviour behave non-ideally. **Spherical aberration** and **chromatic aberration** are two that are worth being aware of, but there are many more (see any introductory optics text, such as *Pedrotti, Pedrotti, and Pedrotti* [1] for more information).

Spherical aberrations manifest as light entering the lens at different positions being

[8]The **optical power** of a lens (measured in **dioptres**) is defined as the reciprocal of the focal length, and this is the number an optician prescribes you when giving a glasses prescription. Stronger lenses have higher optical powers.

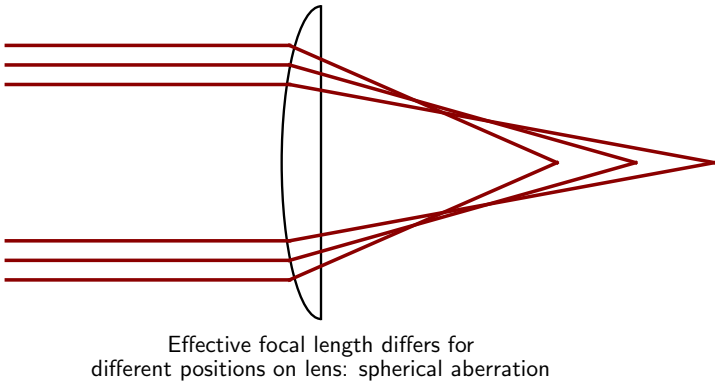


Figure 6.5: Spherical Aberration: light entering the lens at different positions is focussed to different points in space.

focussed to different points in space, as shown in Figure 6.5. Spherical aberrations are very common because making polished spherical surfaces is easy, so the curved faces of most lenses are just sections of a sphere. Unfortunately this isn't ideal for focussing, and we get the aberration below. You can get limit spherical aberration by using **aspheric lenses** (expensive), using compound lenses (also possibly expensive), or simply by reducing the area of the lens exposed to the beam using an iris or telescope (such that the beam all passes nearer to the centre of the lens where it behaves more ideally). Generally speaking, unless you are doing imaging, you don't need to worry much about spherical aberration.

We have already met **chromatic aberration** (Figure 2.7)). Recal that the focal length of a lens depends on the colour of the light being focussed, as refractive index depends on wavelength. So (for example) blue light will generally experience a higher refractive index than red light in most optical materials. This means that a lens with a certain focal length for blue light will actually have a longer focal length for red light. Put another way, the same lens will tend to focus blue light harder than red light. Chromatic aberration is (in my experience as a spectroscopist!) more problematic than spherical aberration as we often want to overlap foci of beams of different colours at the same point in space (we will discuss this in due course). However, it is not generally the case that the focal length of a lens will vary appreciably over the bandwidth of an ultrashort laser pulse, except in the most extreme circumstances.

Finally, another aberration worth being aware of is **astimagtism**. You might have also heard of this in the context of eyesight, as it is a common thing that prescription eyewear can correct for. Astigmatism is an aberration where the focal length of a lens is different for light that passes through it along the horizontal and vertical axes. The result is that the smallest focal spot for the horizontal axis of the mirror is at a different point in space to the focal spot for the vertical axis – resulting in a larger 'effective focus' (sometimes called the *circle of least confusion*) somewhere between these two foci. Astimgatism can be corrected for using specially shaped *astigmatic* lenses, but generally this is only important

if you are a) imaging, or b) in need of exceptionally tight focussing.

However, coming back to the specific case of *ultrafast* optics, lenses have one big drawback. Lenses are **transmissive** and so add dispersion to any optical assembly, because the pulse goes through some glass which adds GDD, broadening the pulses. A good upper estimate for the amount of GDD a lens will add is the same as the amount a flat piece of the same material as thick as the widest part of the lens would add. If the GDD added by a lens would be problematic for your application, then you either need to **pre-compensate** for the added GDD (section 4.3.4) using a compressor, or avoid the transmissive optic altogether and use a **focussing mirror**.

When working with lenses it is important to keep in mind that focussed high-intensity laser beams are much more hazardous than non focussed beams and will much more easily burn things (including your skin), so you should always place blocks after lenses until you are certain the beam is going where you want it to. Further information about working with lenses in practice is given in the next chapter.

Beam Focussing

The simplest situation where lenses are used is in focussing laser beams onto experimental samples. This is often necessary as the beam from a laser output is often quite large, and so just shining this beam onto a sample will often not provide enough intensity at the sample to perform the measurement of interest. A simple and cheap way to increase the intensity is to make the beam waist at the sample smaller. A typical example might be taking the output of a Ti:Sa laser system with a beam waist of around 5 mm, and focussing it using a lens to a point with a beam waist of around 50 µm. Reducing the beam waist by a factor of 100 (as here) will result in the irradiated area being reduced by a factor of 100^2 - increasing the intensity by a similar factor.

To do this in a lab, you need to have a way to calculate the lens focal length needed to produce a given beam waist in the focus. More detail on the full theory of Gaussian beam focussing can be found in many standard texts [1, 2, 3], and online calculators for many of the relevant parameters can be found on the Optics Toolbox [4]. However, if we assume that our lens is placed well inside the Rayleigh range of the beam[9], then the beam waist after focussing ρ_0' is given by Equation 6.2.

$$\rho_0' = \frac{f\lambda}{\pi\rho_0} \tag{6.2}$$

Where f is the focal length of the focussing lens, λ is the wavelength of the focussed light, and ρ_0 is the beam waist before focussing. Because our lens is well inside the Rayleigh

[9] This ensures the laser beam is not wildly divergent at the lens position, so the exact position of our lens along the propagation direction of the beam does not enormously change the size of the beam on the face of the lens.

range of the laser, then ρ_0 is essentially the beam waist of beam as it hits the lens, i.e. the size of the beam on the lens. Equation 6.2 illusrates some key concepts in beam focussing.

1. Lenses with a longer focal length f will produce larger beam waists after focussing and vice versa. This is intuitive as we think of lenses with longer focal lengths as being 'weaker' lenses that don't focus as hard.

2. Longer wavelengths of light λ will focus to larger beam waists than shorter wavelengths. This is maybe less intuitive but is consistent with our knowledge that blue light will refract more strongly than red light (hence the blue parts of a rainbow are closer to the inside than the red parts).

3. A smaller beam hitting the lens (smaller ρ_0) will focus to a *larger* beam waist ρ_0'. This seems counter-intuitive, but can be understood by considering the construction of a typical lens. Close to the center of the lens, the lens is less strongly curved than it is at the edges - therefore a beam hitting the edges of a lens (a larger beam) will be refracted more strongly than a beam hitting only the very centre (a smaller beam). The stronger refraction results in a more tightly focussed beam, and a smaller ρ_0'.

Application 6.3: Sample Focussing

Say we have a laser producing a circular Gaussian beam with $\rho = 5\,\text{mm}$ $100\,\mu\text{J}$ pulse energy at a central wavelength of 1030 nm. The fluence of this beam incident on a sample would be given by (Equation 5.6):

$$\frac{2 \times 100\,\mu\text{J}}{\pi(5\,\text{mm})^2} = 2.54\,\text{J}\,\text{m}^{-2}$$

Imagine we need a higher fluence than this for our experiment, and cannot increase the pulse energy further. If we wanted to increase the fluence by a factor of 10000, we would need to shrink the beam area by a factor of 10000. This is possible if we reduce the beam waist to 50 µm (a factor of 100), i.e:

$$\rho_0' = \frac{\rho_0}{100}$$

How do we achieve this? Using Equation 6.2 and the above, we can see that we need a lens with a focal length:

$$f = \frac{\pi\rho_0^2}{100\lambda} = 0.76\,\text{m} \approx 750\,\text{mm}$$

So a 750 mm lens will do this job for us. Of course, we might have other reasons for wanting a smaller beam, like the need to focus the light into a small cuvette or a thin gas jet – the same procedure applies.

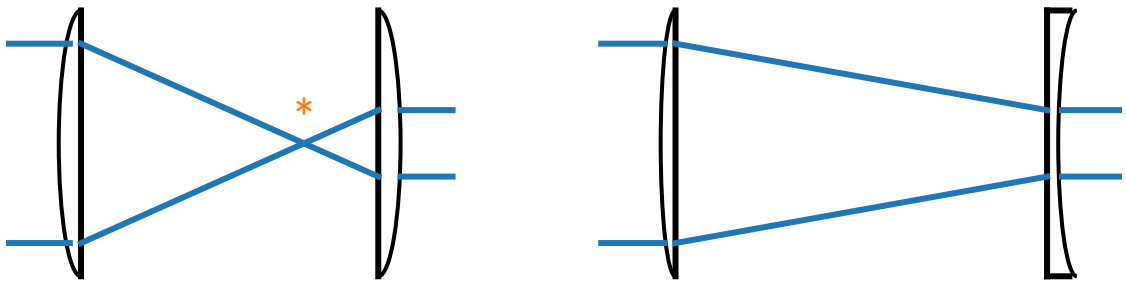


Figure 6.6: Schematics of a Keplerian (left) and Galilean (right) telescope, both set up to shrink an input laser beam (blue lines). In a Keplerian telescope the beam goes through a focus at the starred point, which can be problematic when using high-intensity beams.

Telescopes

Sometimes it is desirable to change the beam waist of a collimated beam. For example, the output beam from your laser may be large and difficult to work with, so making it smaller (whilst retaining collimation) is desirable. Alternatively, you may have a very small beam that you want to focus hard, and need to expand the beam so that it is larger when it passes through the focussing lens. In all of these cases, you want to change the beam waist whilst maintaining the collimation of the beam. This requires that you use a **telescope** of some description, either a **beam expander** (to enlarge the beam) or a **beam reducer** (to shrink the beam). A telescope made using lenses is called a *refracting telescope*, and there are a few different designs available, of which the two most common are the **Galilean Telescope** and **Keplerian Telescope**.

Figure 6.6 shows an illustration of these two geometries of telescope. The Keplerian telescope (left) uses two converging lenses, whereas the Galilean (right) uses a converging and a diverging lens to achieve the same behaviour. The major disadvantage of the Keplerian design for this purpose is that the beam goes through a focus (indicated with a star on Figure 6.6) between the two lenses. For a high intensity beam (and most of the time our ultrafast lasers produce high intensity beams), going through this focus can result in the intensity being high enough to ionise the surrounding air. This can ruin the beam quality after the focus, and if the intensity is high enough then non-linear frequency conversion can occur in the air which will make your beam look very colourful after the focus! For this

reason, we tend to want to work with **Galilean** telescopes in ultrafast optics, and this is what will be considered further in this section.

A Galilean telescope that is set up to work as a beam reducer (shrinking a beam) will have a converging lens at the input and a diverging lens at the output (as shown in Figure 6.6). Conversely, the same lenses could be used to produce a beam expander if the diverging lens was placed at the input and the converging lens at the output. Figure 6.7 shows a general design for a Galilean telescope. The **magnification** of the telescope, M, is given by:

$$M = -\frac{f_2}{f_1} = \frac{d_2}{d_1} \tag{6.3}$$

Where f_1 and f_2 refer to the focal lengths of the two lenses (f_1 is at the input, f_2 is at the output), and d_1 and d_2 refer to the diameters of the input and output beam respectively[10]. The minus sign in Equation 6.3 is present because one of the focal lengths will be negative (that of the diverging lens), but a negative magnification would be unphysical. If the input beam is perfectly collimated (and the lens focal lengths are exactly what you think they are), then placing the lenses a distance $s = f_1 + f_2$ apart will produce the desired behaviour. Practical tips for aligning telescopes will be discussed in the next chapter.

6.1.3 Focussing Mirrors

An alternative option for focussing laser beams is to use a **focussing mirror**. This is essentially a concave mirror that causes light reflected from it to be focussed down to a point. This is illustrated in Figure 6.8. There is an immediate clear advantage of using a mirror to focus ultrashort pulses, which is that the mirror is not transmissive, so it will not add GDD to the pulse in the way that a typical lens would[11]. In addition, the focal length of the mirror does not depend on the colour of the incident light in the way that the focal length of a lens does – there is no chromatic aberration. To understand this, consider that the root cause of lenses to focus different colours of light to different focal points is that the refractive index of the lens material varies with for different colours. Thus, different colours will undergo differing degrees of refraction through the lens and therefore be focussed to slightly different points.

A focussing mirror will not suffer from this problem. As the mirror is not transmissive, we don't care about the refractive index of it - and any colour of light incident on the mirror will be focussed to the same focal point. Practically, the lack of chromatic aberration means that focussing mirrors are ideal for use in places where focussing of many different colours of light is desired (for example, if you scan your laser output wavelength across a broad range, or need to focus a broad white light continuum onto an experimental target). It is also possible to make **reflecting telescopes** using a curved mirrors. These are much less

[10] One may say that this should be defined in terms of beam waists and not diameters, as we have done elsewhere in this text. However it doesn't matter in this case as we are only interested in the ratio, and it is a good lesson in dealing with irritating people that use non-standard notation.

[11] Save for any GDD accumulated during the reflection - but focussing mirrors tend to be metallic and so incur minimal GDD anyway.

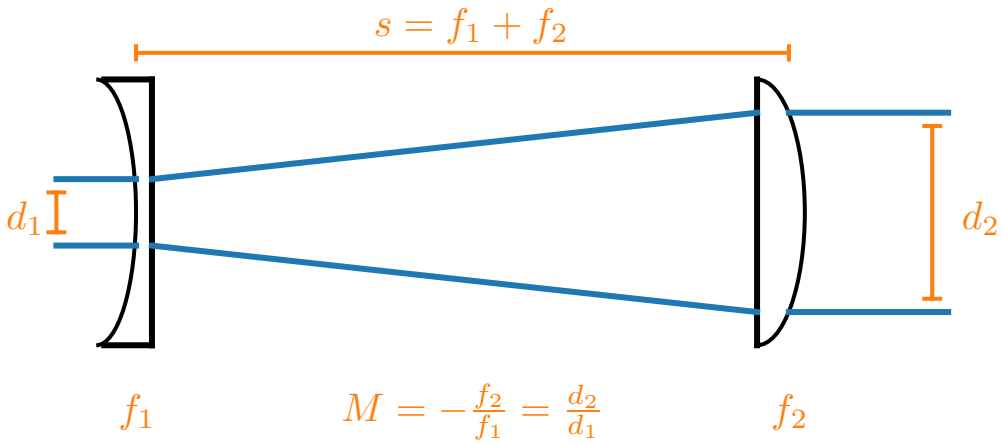


Figure 6.7: Design of a typical Galilean beam expander. Edges of a laser beam shown in blue. A diverging lens with focal length f_1 is placed at the input and a convering lens with focal length f_2 re-collimates the output. The laser enters with a diameter d_1 and exits with a diameter d_2. The magnification factor M of the telescope is given by $M = -\frac{f_2}{f_1} = \frac{d_2}{d_1}$. The lenses should be placed at a distance $s = f_1 + f_2$ to ensure that the output is collimated.

Application 6.4: Application: Beam Expansion Telescope

Imagine we want to increase the beam waist of our laser beam by a factor of two. This means that:

$$d_2 = 2d_1 \rightarrow \frac{d_2}{d_1} = 2 = M$$

Thus, the magnification required for our factor-of-two expansion is simply 2. We can then pick any combination of lenses f_2 and f_1 to achieve this, with the caveat that the sum of the focal lengths must be a positive number (otherwise the distance between the two lenses, s would need to be negative).
For example, this could be achieved with by having f_1 =−50 mm and f_2 =100 mm, such that $s = f_1 + f_2 = -50 + 100$ =50 mm. The table below gives some further examples of combinations of lenses that can be used to give a variety of beam expanding telescopes (where $M > 1$) and beam reduction telescopes (where $M < 1$).

f_1 (mm)	f_2 (mm)	$s = f_1 + f_2$ (mm)	$M = -\frac{f_2}{f_1}$
-50	100	50	2
-100	300	200	3
100	-50	50	0.5
200	-50	150	0.25

Analogous calculations can be made to create reflective telescopes using curved mirrors – discussed further below.

common than refracting telescopes as they are more fiddly to align and build, but can be useful if you cannot compensate for the accumulated GDD that a refracting telescope will introduce, or need to work with very broadband beams. Another advantage of focussing mirrors is that they can be used to focus high intensity beams that would cause self-focussing (or other non-linear effects) if focussed using a transmissive optic like a lens. Focussing mirrors are commonly specified using the **radius of curvature**, R, of the mirror rather than the focal length, f, but these are easily related with the formula $f = -R/2$.

It may sound as though using a focussing mirror is somewhat of a no-brainer when it comes to focussing your ultrashort pulses! No accumulated GDD and no chromatic aberration can be very useful, however a lot of the time you will still encounter people using lenses for focussing unless they absolutely have to use mirrors. The reason for this is two-fold. Firstly, focussing mirrors can be very fiddly to align properly. Having a single optic both reflect the beam and focus it can be annoying enough to align that many people would rather opt to decouple these effects and use a mirror followed by a lens, and then compensate for any extra GDD if necessary. Secondly, focussing mirrors tend to be made of metal, and have the same downsides as all metallic mirrors. That is, they are more easily damaged and less reflective than dielectric mirrors. For these reasons, use of lenses is still widespread.

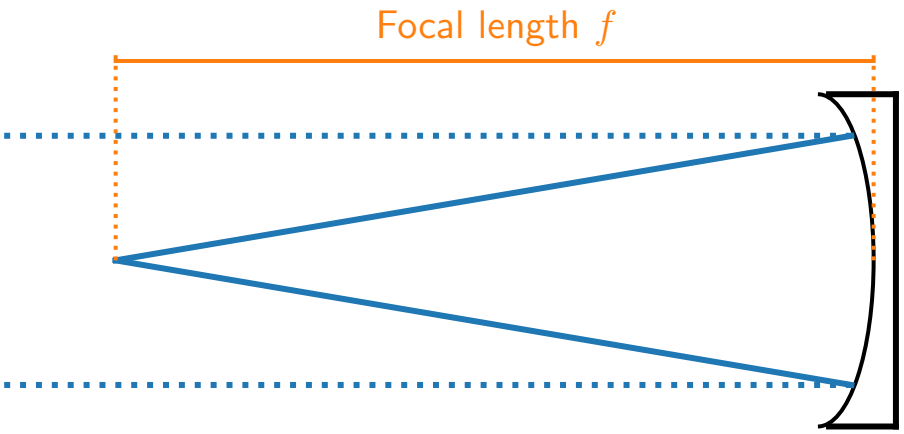


Figure 6.8: Illustration of how a focussing mirror works. The incoming beam is shown in dashed blue, and the focussed beam in solid blue. The focal length is defined from the center of the concave mirror.

Focussing Mirror Shape

As lenses come in a variety of shapes (spherical, aspheric, cylindrical...), so do focussing mirrors. The simplest case is a **spherical mirror** where the performance is analogous to a spherical singlet lens. Another common case is the **parabolic mirror**, where the curved surface is a section of a parabola.. These mirrors exhibit no spherical aberration (Figure 6.9), so it doesn't matter *where* the light hits the mirror: incoming collimated light will all be focussed to the same spot.

These are commonly used for both focussing, and for collimating light from divergent sources. The most common kind of parabolic mirror you tend to see is the **off-axis** parabolic mirror, where the shape of the mirror surface is a section of the parabola some distance away from the central optical axis. These focus beams to different angles, most commonly 90° off-axis. They are commonly available with through-holes drilled to enable other beams to be overlapped with the focussed or collimated beam. These concepts are illustrated in Figure 6.10.

Another type of focussing mirror that you might encounter is the **toroidal mirror**. A toroidal mirror has two different radii of curvature that are orthogonal to each other – a spherical mirror is a special case of a toroidal mirror where both radii are the same. The analogous transmissive optic is the **astigmatic lens**. These optics minimise astigmatism caused by light hitting an optical element at a funky angle - they are commonly found in

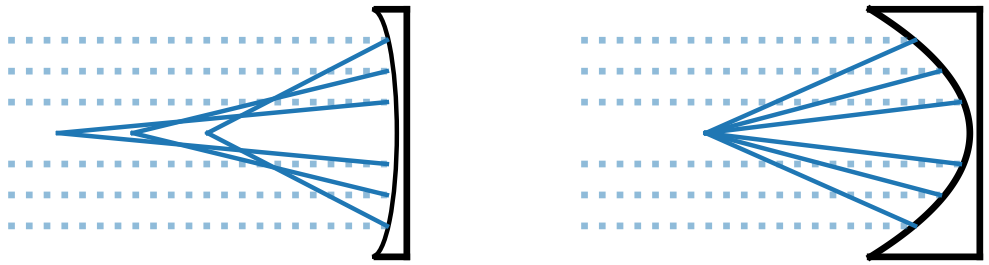


Figure 6.9: Left: a spherical focussing mirror, which will exhibit spherical aberration. Right: a parabolic focussing mirror, which will not.

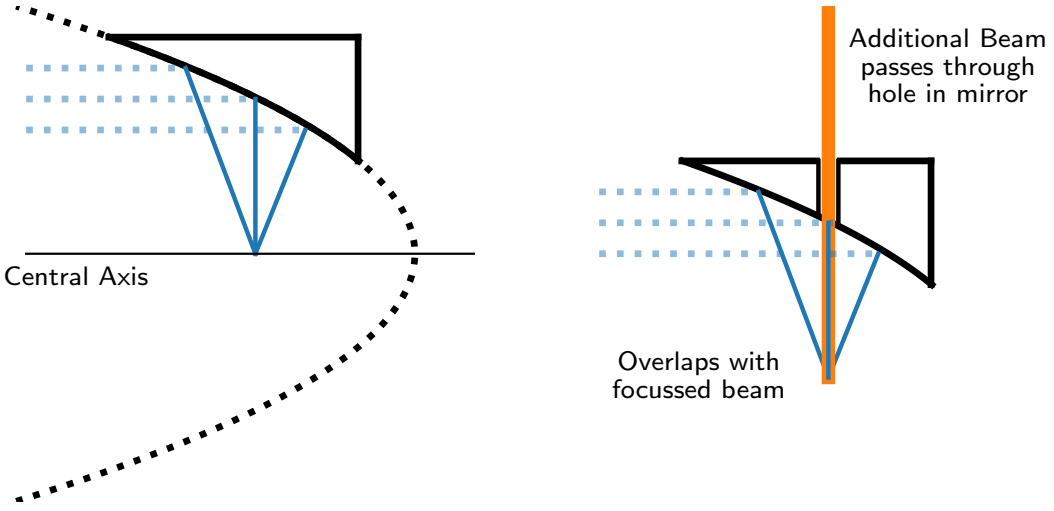


Figure 6.10: Left: an off-axis parabolic mirror, showing the definition of 'off-axis'. Right: an off-axis parabolic mirror drilled with a through-hole, so an additional beam (orange, wide line) can pass through and overlap with the focussed beam.

spectrographs (where aberrations like this can be important), and also when working with X-rays (where you normally need to have the X-ray beam hit any optic at a very shallow angle to ensure reasonable reflectivity).

6.2 Temporal Manipulation

Now we consider how we can manipulate the temporal properties of our ultrashort pulse. We start by revisiting simple pulse compression and stretching, and then introduce the more complex topic of *pulse shaping*. Finally, we discuss ways to manipulate the temporal properties of the beam, rather than the pulses. We start with a general discussion of prisms and gratings, as they feature heavily in the following discussion.

6.2.1 Dispersive Optics

When we say *dispersive optics* we mean optics that can spatially disperse different colours of light, **diffraction gratings**, and **prisms**. The dispersive action of a prism will be familiar to many from the cover of Pink Floyd's timeless album 'Dark Side of the Moon', and the rainbow effect that can be seen on the back of a CD/DVD is because the lines etched into the disc that store the data function as a diffraction grating and disperse the light. It is important to note that in this case when we say 'dispersive', we are talking about **spatial dispersion**, the spreading out of different colours *in space*; as opposed to **temporal dispersion**, which is the spreading out of different colours *in time* that we have already discussed at length (and referred to simply as 'dispersion') in this book. Going forward, I will be explicit whenever I refer to dispersion in a potentially ambiguous situation, but assume that an unqualified 'dispersion' refers to temporal dispersion, as it has for the book so far.

Being able to spatially disperse the colours within our laser pulse is a very useful thing in that it allows us to decompose our broadband laser pulse into its constituent colours. This is known as *accessing the Fourier plane* by aficionados[12], as we have physically moved from working in the time domain to the frequency domain after the spatial dispersion, just like we would if we performed a Fourier transform. Accessing frequency space like this allows us to do a plethora of useful things to our laser light such as pulse compression and stretching, as well as being very useful diagnostically.

Gratings

A **diffraction grating** (referred to as simply a 'grating' most of the time), is an optical substrate which has a large number of very fine lines physically etched into the surface of it. Light can then be bounced off of the etched surface (a 'reflection grating'), or passed through the back of the substrate and then through the etching(a 'transmission grating'). The spatial dispersion that is caused when light bounces off or passes through a grating

[12]Strictly this requires use of an extra focussing element, but we are still accessing frequency space by dispersing the colours.

depends on the density of the etched lines, and so manufacturers specify the number of lines per unit length, generally as lines per millimetre. A higher density of lines causes more severe spatial dispersion.

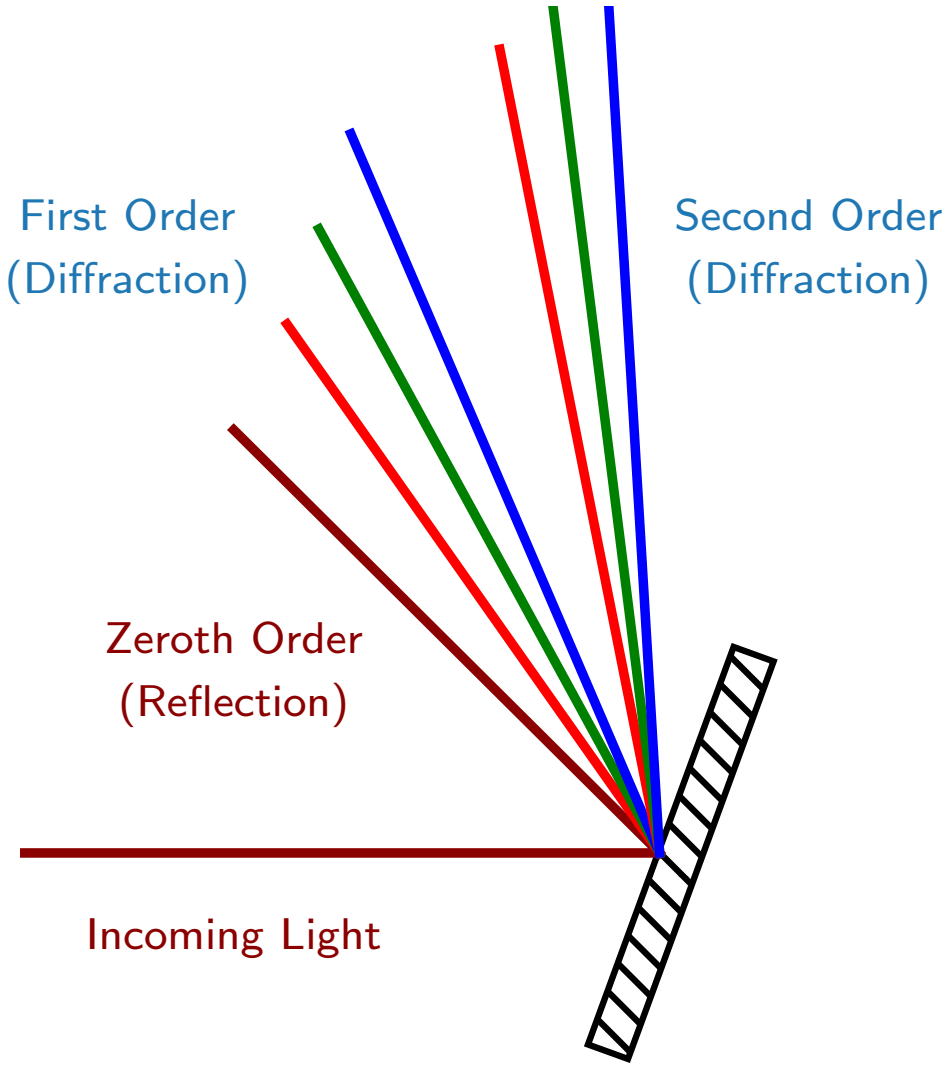


Figure 6.11: Illustration of reflection of a laser beam from a typical diffraction grating. The diffraction gratings produces various diffraction orders. The first and second orders spatially disperse the frequency components of the light (the light is diffracted), whereas the zeroth order is simply reflection from the grating surface.

Figure 6.11 shows diffraction from a typical reflective diffraction grating. The reflection consists of multiple diffraction orders of which three are shown (this may be familiar in the context of x-ray diffraction from crystal structures). The first and second order diffractions spatially disperse the light, while the zeroth order diffraction is just a reflection from the grating surface. Having multiple diffraction orders can create a lot of stray beams and power loss, so it is often desirable to suppress unwanted orders. This can be done by engineering

the shape of the etched lines in the grating, which is a process known as **blazing**. A **blazed grating** will concentrate light of a specific wavelength into a certain diffraction order, enhancing the efficiency in this region. For example, a grating may have a 400 nm blaze, which means that light incident on the grating around the 400 nm region will be concentrated into a specific diffraction order - and so the grating is most efficient when used at this wavelength.

Gratings are commonly found inside laser systems for pulse compression. They are preferable to prisms for this purpose because they can cope more easily with the high intensity beams inside a typical output compressor, and also add much more GDD. However, gratings are generally more expensive than prisms, do not function over as wide a range of wavelengths, and can cause significant power loss.

6.2.2 Prisms

A **prism** is a wedge-shaped piece of glass (shaped like a trigonal prism), which spatially disperses light. The spatial dispersion of a prism is caused (again) by the refractive index of the prism material being dependent on the wavelength of incident light. Different wavelengths of light experience different refractive indices and so undergo different amounts of refraction in the prism, which leads to spatial dispersion. Accordingly, the material that the prism is made from affects the degree of spatial dispersion (higher refractive indices disperse more strongly).

A cursory look at any manufacturers catalogue will show that a lot of prisms are intended for beam steering or imaging (binoculars and spotting telescopes tend to use roof-prisms in their construction) rather than simply spatial dispersion. Prisms are generally inappropriate for beam steering with ultrashort laser pulses as the GDD added by the prism material will broaden the pulse significantly. When working with ultrafast lasers, prisms are generally only used for their dispersive properties - to make compressors/stretchers, to tune the dispersion in a laser oscillator cavity, or to spatially disperse a beam for diagnostic reasons.

Application 6.5: :Beam Diagnostics

Using a cheap prism is an excellent way to disperse the different colours in your pulse when doing laser diagnostics.

For example, you may be using a nonlinear crystal to generate the second harmonic of your laser output. It can be difficult to see the second harmonic beam as it could be relatively weak and will overlap in space with the much more intense fundamental beam. Directing both beams into a prism and catching the spatially dispersed beams on a block will separate the colours, allowing you to clearly see the second harmonic (and optimise your crystal parameters to maximise how much you have).

Both prisms and gratings are heavily used in the creation of typical pulse compressors and stretchers.

6.2.3 Compression and Stretching

We have already discussed compression and stretching at length in this book, in the context of learning the fundamental physics that underlies ultrashort pulse generation. Recall that the fundamental ideas are:

- An ultrashort pulse is short because it contains many colours (broad bandwidth).
- The shortest duration it can have is called the **transform limit**, which occurs when every colour in the pulse arrives at the same time.
- **Dispersion** causes the spreading out of the colours in time, and thus broadens the pulse away from the transform limit. We mostly care about the second-order dispersion, or **GDD**.

So, to manipulate the pulse duration we generally need to add positive or negative GDD to our pulse. We can achieve this by simply passing it through materials, or by using prism and grating compressors (discussed earlier - see Figure 4.9). In this section we will introduce a few other ways of influencing pulse duration.

Chirped Mirrors

A **chirped mirror** is a mirror that will add GDD to a laser pulse on reflection. This kind of mirror is not typically used for general steering of a laser beam, so was not discussed earlier, but is of obvious use in pulse compression and stretching.

To understand how a chirped mirror, note that the thickness of a dielectric coating will induce some GDD in a reflected pulse, as different colours in the pulse reflect off different 'depths' of the coating, and so travel through different path lengths (see Figure 6.20 for an illustration and discussion around this phenomenon). In a standard steering mirror, this is generally an unavoidable negative effect that is often negated as far as possible in the coating design. However, in a **chirped mirror**, the coating is engineered to actively induce GDD to aid with pulse compression. A chirped mirror is most often designed to add negative GDD (to counter the positive GDD most ultrafast pulses gain as they propagate), but examples can be found that add positive GDD too.

Chirped mirrors are generally marked by how much GDD they add per reflection ('per bounce'). A typical chirped mirror may add $-50\,\mathrm{fs}^2$ of GDD per bounce, so bouncing a beam off a single mirror multiple times can add a relatively large amount of negative GDD which can help to compress the pulse. They are usually used in pairs (as shown in Figure 6.12), to add a fixed amount of GDD to the pulse. Specific mirrors can be bought that will 'undo' the GDD added by propagation through typical optics like lenses or windows. They come in different shapes, but are predominantly long and rectangular so many 'bounces' can fit onto one mirror.

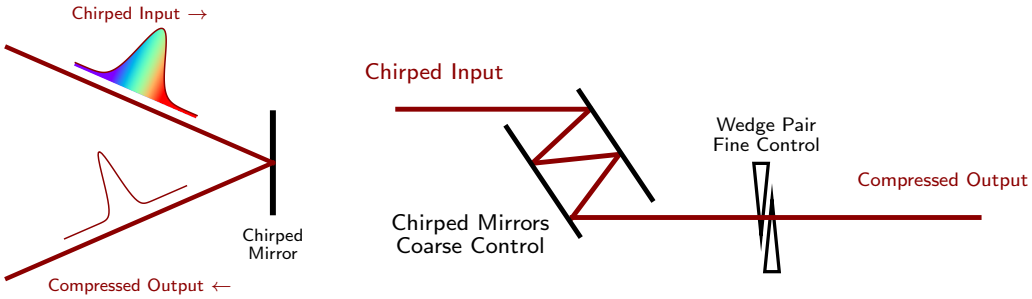


Figure 6.12: Left: basic operating principle of a chirped mirror. A bounce adds a fixed amount of GDD, which can compress a pulse. Right: practical use of chirped mirrors in a setup. A chirped mirror pair adds a large, fixed, amount of GDD, and a pair of glass wedges provide a variable small amount for fine tuning.

Fine Control of Dispersion

Chirped mirrors add a fixed amount of dispersion to a pulse. This might be ideal, but it is more usual to want to adjust the dispersion added to compensate for changing conditions and ensure optimal pulse compression at your experiment. Prism or grating compressors can accomplish this, but are fiddly to align (and can be a pain to make motorised). A simpler way to fine-tune the dispersion following use of chirped mirrors is to use a **pair of thin glass wedges**, as shown on the RHS of Figure 6.12. These wedges work a bit like a mini prism compressor. You translate them in or out of the beam to add differing amounts of material dispersion to your pulse: you need two of them, as the second wedge undoes the refraction the first one incurs which would otherwise change the pointing direction of the beam. As they are just made of thin bits of glass, they are relatively cheap (but you do need some translation stages – see later). They are commonly used in conjunction with chirped mirrors to add an ability to fine tune the compression of a pulse (see Example 6.6), and can also be used to vary the *carrier-envelope phase* of a few-cycle laser pulse.

6.2.4 Pulse Shaping

A more sophisticated way to play with the temporal shape of your pulse is via **pulse shaping**. Rather than simply making the pulse broader or narrower in time, using a **pulse shaper** allows you to change the shape of your pulse to almost anything you want: a Gaussian, double Gaussian, exponential, reverse exponential, top-hat... the list is almost endless. There are many applications in spectroscopy where being able to control the precise pulse shape is beneficial.

The shaping is physically accomplished by dispersing the colours in the pulse (like you would in a compressor), but then by applying a *different phase shift* to each colour in the pulse – see Figure 6.13 for a typical schematic. When you recombine the colours you then end up with a different pulse shape, depending on the phase shifts you added. In

Application 6.6: Fine-Tuning Compression

Imagine you are the proud owner of a new multiphoton microscope – these microscopes use nonlinear optical effects to enhance the information gained from traditional microscopy. As we know, nonlinear optical effects require a high intensity, and so ultrashort pulses are needed. A microscope objective may have several centimetres of glass in it – all of this will add a lot of dispersion to your pulse. A way to pre-compensate for this dispersion is needed. The numbers given below are illustrative of the process – you would need to look at your setup and consult manufacturer data for your application.

Let us imagine that our microscope is pumped using a 1030 nm laser, and that all the glass in between laser output port and sample adds up to about 10 cm. If we assume this is mostly BK7 glass (25 $\text{fs}^2\,\text{mm}^{-1}$ GVD), then we know that our initially transform limited pulse will gain 2500 fs^2 of GDD as it propagates. If the pulse started life as a 50 fs transform limited pulse, after all the glass it will end up with a duration of 150 fs: three times longer, so three times less intense.

We need to pre-compensate for this somehow. We could use a prism or grating compressor, but using chirped mirrors is simpler from a construction perspective. There are many different kinds of chirped mirror available, but say we find one that gives $-175\,\text{fs}^2$ GDD per bounce, so $-350\,\text{fs}^2$ for each pair of bounces. We would need to do:

$$\frac{2500}{350} = 7.14$$

Pairs of bounces on these mirrors to undo all our dispersion. Obviously we can't do a non-integer number of bounces, so what we do instead is round up (so that we overshoot and end up with a negatively chirped pulse at the end), and then add the remainder using glass wedges. If we did 8 pairs of bounces, we'd add $-2800\,\text{fs}^2$ of GDD to the pulse, ending up with $-500\,\text{fs}^2$ of GDD 'in' our pulse.

We could remove this with a pair of wedges – a typical offering might add up to 600 fs^2 of GDD if the wedges are fully inserted. In conjunction with the chirped mirror, this would allow us to bring the pulse back to the transform limit, and allow us some wiggle room if we end up needing to add extra optics to setup later that add more GDD.

practice, these are constructed using some dispersive optics to split the pulse up into it's frequency components, and then some kind of **modulator** to add the phase shift to different frequencies. This modulator could be something like an *acousto-optic modulator* (AOM) which sends sound waves into a material to modify the refractive index as the material is compressed and expanded by the sound; or a *spatial light modulator* which essentially acts like an LCD screen that shows each colour a pixel with a different opacity. In addition to allowing funky pulse shapes to be generated, a pulse shaper can also operate like a simple

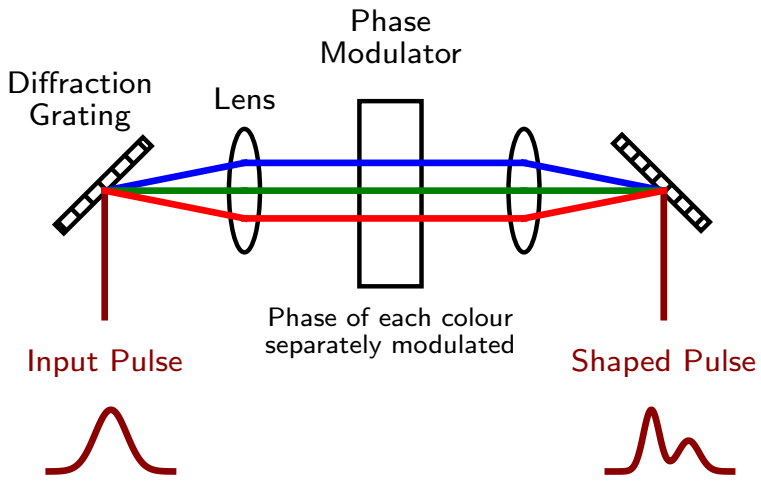


Figure 6.13: Schematic of a typical $4f$ pulse shaper. Light is dispersed and collimated using a grating and lens, and then some kind of modulator is placed in the Fourier plane to allow the phase of each individual frequency component of the light to the modulated. After modulation, the pulse is re-collimated and a shaped pulse is produced.

stretcher or compressor.

Pulse shapers can be homemade, but often are bought as commercial sub-assemblies[13]. They are particularly essential in multidimensional optical spectroscopies (Hamm and Zanni's book [5] is an excellent introduction to these), but have a wide range of uses outside of these areas. More information about building and using pulse shapers can be found in references [6, 7].

6.2.5 Repetition Rates and Chopping

A final important aspect of temporal manipulation of our laser output concerns manipulating the **pulse train** the laser produces, rather than the pulses themselves. Recall (Figure 4.4) that an ultrafast laser produces a train of pulses at a given repetition rate f_{rep}, and the time between two pulses is $T_{\text{rep}} = {}^{1}/_{f_{\text{rep}}}$.

There are many reasons why we might want to modify the pulse train and repetition rate:

[13] Check out companies like *PhaseTech* or *Fastlite* for examples.

- We might want to split our beam in two, and then modify one pulse train by chopping out every other pulse such that we can measure a signal both when pulse from the modified train is there and when it is not.
- We might have an experiment that runs with a lower duty cycle than our laser, but still want to measure in a 'single shot' mode so that we don't accumulate laser shots on the sample that don't give us a signal.
- We might want to give our experiment more time to 'relax' between pulses than T_{rep}.

There are a number of ways these modifications can be achieved.

Chopping

At low repetition rates ($<$10 kHz), we can use an **optical chopper** to accomplish this. An optical chopper is basically just a disc with slots in it that is spun at a precisely controlled frequency, you can synchronise the frequency to your laser repetition rate and use the chopper to chop out certain pulses in your train – maybe every other pulse, for example.

Optical choppers are commonly used in transient pump-probe spectroscopy where measuring 'pump on' and 'pump off' signals is required, and doing so on a shot-to-shot basis can be beneficial for the signal to noise ratio (SNR) of the experiment. They are often used in conjunction with *lock-in amplifiers*, which are electronic amplifiers that are able to 'lock' onto a signal produced at a specific frequency and extract it with an exceptionally high SNR. More mundanely, you can use them to simply reduce repetition rate (blocking a certain proportion of the input pulses), or even as optical shutters.

Pulse Picking

At higher repetition rates, we need to use a different method than simply chopping. Optical chopper repetition rates tend to max out at about 10 kHz – after this point, the disc would have to be spinning so fast that it air resistance becomes significant (producing a lot of heat) and the rotation exceptionally noisy[14]. At higher repetition rates, we can use an **electro-optic modulator** (EOM) or an **acousto-optic modulator** (AOM) to perform the same job as a chopper. We met AOMs briefly in the discussion of pulse shapers. EOMs are commonly used inside laser amplifiers to pick out pulses from the oscillator pulse train for amplification at lower repetition rates: devices that do this are known as **pulse pickers** for this reason.

Both AOMs and EOMs work by modulating the light using acoustic or electrical means. Generally they work by deflecting a beam at some angle at the desired repetition rate. In this way, you can pulse pick at high repetition rates, which is useful when using laser oscillators that operate with MHz repetition rates. You can even use them to achieve things

[14] Some companies offer optical choppers that can chop at up nearer 10 kHz, but these come with a disclaimer that you probably need to wear ear defenders while they are running!

like sending sequential pulses in the train in different directions to rapidly scan the beam across a sample.

6.3 Polarisation Manipulation

It will often be useful for us to manipulate the polarisation state of a laser beam. Any kind of light-matter interaction that is dependent on the relative orientation of the electric field from the laser with the orientation of something in the laboratory will be affected by it[15] **Polarisation optics** is a catch-all term that refers to any kind of optical element that can vary these polarisation states.

We can divide our discussion of polarisation optics into two parts. Firstly, optical elements which will allow different polarisation states to be separated out from a beam (e.g. only allowing s-polarised light to be transmitted and reflecting/absorbing p-polarised light), these are **polarisers** - elements that can turn an incident unpolarised beam into an output polarised beam. Secondly, optical elements which allow the polarisation state of a beam to be actively changed, such as turning an s-polarised beam into a p-polarised beam. These are **waveplates**. The difference between a polariser and a waveplate is that a polariser only *separates out* existing polarisation states from a beam, and does not actually *change* the polarisation state. A waveplate actively *rotates* the polarisation state of a beam (or can change it from linear to circular polarisation, for example). This is illustrated in Figure 6.14.

An accessible mathematical framework for understanding polarisation in optics uses the **Jones Calculus**, where the polarisation state of light is expressed using a vector (the *Jones Vector*), and each optical element that can affect polarisation is written as a transfer matrix (the *Jones Matrix*). The effect of a number of different optical elements on some initial input light can then be simply calculated using matrix algebra. This method is described in detail in many standard textbooks on optical physics, together with another formalism using the **Stokes Parameters** [8, 1].

6.3.1 Polarisers

Polarisers work by transmitting one polarisation state of a beam and reflecting or absorbing the other. The effectiveness of a polariser is determined by it's **extinction ratio**, which is defined as the ratio of the transmission of the desired polarisation state to the transmission of the undesired polarisation state. Very high-end polarisers can achieve extinction ratios as high as 100000:1 or more. When you buy a polariser, it will have an indication of the **transmission axis**[16]. Light that is polarised parallel to the transmission axis will pass straight through the polariser - it will transmit light that is polarised parallel to the transmission axis. The remainder of the light is either reflected out of the polariser (more common) or absorbed by the polariser (less common). In most applications, polarisers that

[15] These range from simple nonlinear frequency mixing through to more complex effects exploiting the chirality of circularly polarised light with the chirality of molecules – *circular dichroism spectroscopy*.

[16] Either directly marked on the mounting (if the polariser came mounted), or supplied with the accompanying literature.

Figure 6.14: Illustration of the function of a polariser (left), as opposed to a half-wave plate (right). The polariser acts to produce well-polarised light from an unpolarised beam, whereas the waveplate actively rotates the polarisation to a different direction.

reflect the non-transmitted polarisation are more appropriate, as then the reflected beam can be used or dumped safely, rather than simply heating up the polariser.

There is a very wide variety of different kinds of polariser, as any look through a manufacturers catalogue will show you. A thorough review of all the different types of polariser is beyond the scope of what we do here, but there are three situations where a polariser is commonly needed in ultrafast laser experiments which will be discussed in some detail. Some initial remarks about choosing and using polarisers are given below.

Polariser Choices

In an experiment using ultrafast laser pulses, adding polarisers can be dangerous as a lot of polarisers consist of very thick pieces of glass which can add lots of unwanted dispersion to our input pulse, broadening it significantly. One approach to mitigate this is to use a *thin-film polariser*. A thin-film polariser has a dielectric coating which is engineered to (normally) reflect s-polarised light and transmit p-polarised light. This coating is engineered using interference effects in the same way that a dielectric mirror coating is made. However, TFPs do not generally work over a very wide wavelength range. The range will be specified by the manufacturer, but it is not uncommon to find TFPs that only work well at one single wavelength (or a narrow range of wavelengths). In ultrafast optics, this is especially important, as a TFP marked as an 800 nm TFP may look ideal for use with a Ti:Sa ultrafast laser, but actually may not work well over a broad enough bandwidth to polarise the entire

pulse well. TFPs marketed as 'broadband' or 'ultrafast' are designed specifically for this purpose, and worth the additional cost. TFPs are also commonly supplied with AR coatings, which can be helpful.

Another important type of polariser that you may encounter is the **cube polariser**. 'Cube polariser' is really an umbrella term for a variety of different polarisers such as Glan-Taylor or Glan-Laser polarisers. These polarisers contain very thick glass, so can add a lot of GDD and are not normally ideal for ultrafast use. However, they transmit a very broad bandwidth, and are good for use in areas where the added GDD is unimportant[17]. These kinds of polarisers also have the highest extinction ratios, but tend to not be used with ultrafast laser pulses due to the GDD they add (except at wavelengths where the GDD added is low, which can be the case in the near-IR).

Polarisers are generally very sensitive to the angle of incidence of the input beam. The specific angle depends on the specific polariser, but generally the flat face of the polariser needs to be placed either at 90° or at Brewster's angle[18] to the propagation direction of the incoming beam.

Application 6.7: Polarising Beams

The simplest use of a polariser is, literally, to use it to polarise an existing beam! The output from your laser system is probably (in theory) polarised, but may actually not be incredibly well polarised if you were to measure the polarisation[a]. Some experiments are exquisitely sensitive to input polarisation, so it is often desirable to 'clean up' the polarisation before using the beam in the experiment. A simple way to do this is to insert a thin-film polariser into the beam path somewhere before the beam reaches the region of interest.

By taking the portion of the beam that is reflected by the polariser, we can clean up the polarisation whilst not adding any significant GDD to our pulse. If p-polarised light is desired for the experiment, then adding a waveplate (see below) into the reflected beam path will allow the polarisation to be varied. Alternatively, it may be possible that the substrate of the polariser is thin enough to not incur any significant dispersion (this can be easily calculated for your application using the optical toolbox), and in this case taking the transmitted portion will also work well.

[a] This can be easily done by measuring the transmission of your beam through a polariser using a power meter. If (for example) the polariser transmits 99 mW when the transmission axis is vertical, and 1 mW when the transmission axis is horizontal, then the light is 99% polarised - provided that you know the light is linearly polarised.

[17] For example, after interaction with a sample where you just need to polarise the light to couple it cleanly into a detection system.

[18] Brewster's angle is the angle of incidence where p-polarised light is totally transmitted by a material. It varies from material to material but is normally in the between 50 and 60 degrees.

6.3.2 Waveplates

A **waveplate** is an optic that actively changes the polarisation state of the light incident on it. There are two common types of waveplates , **half-wave plates** (HWP), and **quarter-wave plates** (QWP), and these are discussed in more detail below. Waveplates are sometimes referred to as 'retarders' or 'phase retarders' - a name which derives from the mechanism by which they function, where one component of the polarisation of a wave is delayed or 'retarded' relative to the other by some fixed amount. For a HWP, this is $\frac{\lambda}{2}$, whereas for a QWP it is $\frac{\lambda}{4}$, where λ is the wavelength of the incident light. This corresponds to a phase shift of 180° (in the HWP case), or 90° (in the QWP case) in the delayed component relative to the other. As a result of the delay being a fraction of the incident wavelength, it should be intuitive that waveplates generally do not work over a very wide range of input wavelengths. When trying to rotate the polarisation of a broadband ultrashort pulse, this can be problematic, so generally we use **zero-order waveplates** (see below) to mitigate this effect.

Waveplates are made of a birefringent material, so that the refractive index of light along one of the axes of the waveplate is different to that along another, producing the aforementioned phase shift. These axes are commonly called the 'fast axis' and 'slow axis', referring to the speed of light polarised parallel to the axis. An introduction to how waveplates modify polarisation states is given in section 6.3.2, but many standard textbooks will give a more detailed and mathematical treatment [8, 1].

When buying waveplates you will find that there is a distinction between lower cost 'multiple-order' waveplates, and higher cost 'zero-order' waveplates. The difference can be illustrated by considering two kinds of HWP. A zero-order HWP would produce a phase shift which is exactly half of the wavelength of the incident light. In contrast, a multiple-order HWP would produce a phase shift which is an integer number of wavelengths, plus half a wavelength. This difference arises due to the construction of the waveplates. A zero-order waveplate is constructed from a two very thin, highly polished, pieces of material that are sandwiched together. In contrast, a multiple-order waveplate is made from a single piece of comparatively thick material, so is cheaper and easier to produce. However, multiple-order waveplates are generally more sensitive to the wavelength of the incident light than zero-order waveplates, and if the incident light is far from the designed wavelength then the performance of the waveplate can decrease dramatically (the imparted phase shift becomes further from the ideal value, so your HWP might become a QWP when it is used at far from the design wavelength).

We know that when using ultrafast lasers, the bandwidth of the laser pulse is generally quite broad, and we want the entire pulse to be affected by the waveplate as equally as possible. Put another way, you do not want the polarisation components of the blue part of your pulse to experience a different phase shift from the components in the red part, as then the blue part of the pulse could end up in a different polarisation state than the red part! For this reason, zero-order waveplates are generally preferred for ultrafast work, and specific zero-order achromatic waveplates are also available for extremely broadband applications. These have exceptionally flat retardation across a very broad bandwidth, but

come at an increased cost compared to standard zero-order waveplates.

Waveplates and Polarisation

We have discussed briefly the effect waveplates have on light, but have not discussed how this can actually influence the polarisation of our laser beam in the lab. To gain an initial insight into how this works, we will consider the action of a typical half-wave plate (HWP). HWPs are commonly used as **polarisation rotators**, because linearly polarised light can be rotated through an arbitrary angle using a HWP. To understand how this works, first remember that a waveplate is made of a birefringent material, so has two axes with distinct refractive indices: the fast axis and slow axis.

Now consider what would happen if linearly polarised light was aligned exactly along the fast or slow axis of a HWP[19]. In this case, nothing would happen! As all of the light would experience the same refractive index, so no phase shift would be observed. However, if the polarisation of the light was aligned somewhere between the fast and slow axis, then the situation is more complex.

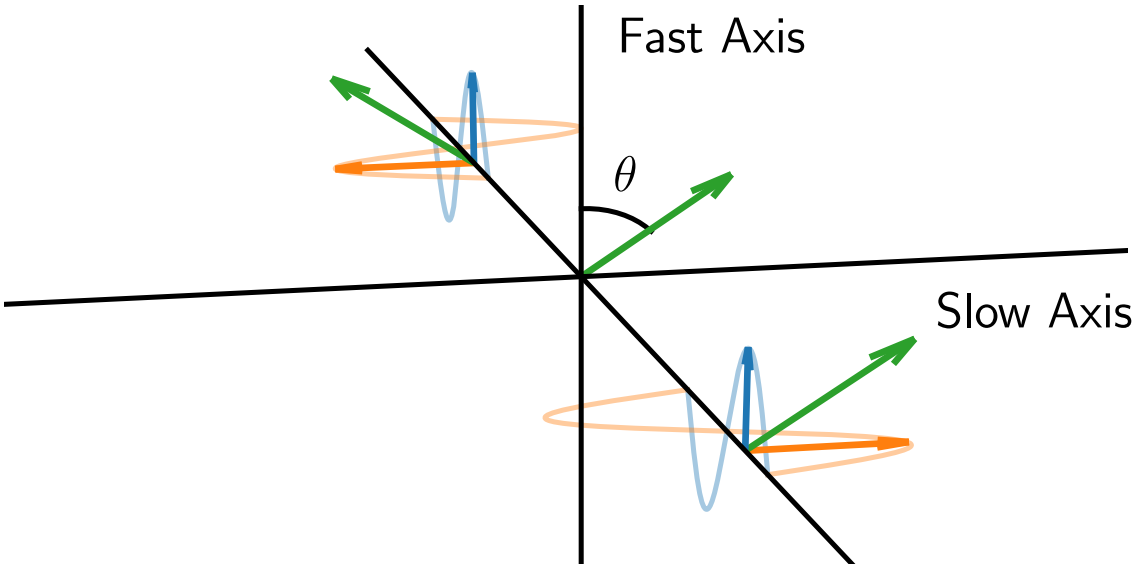


Figure 6.15: How a HWP can rotate polarisation. Linearly polarised light travels from the lower right, through a HWP at the origin, to the upper left. The overall polarisation of the light is shown in green, with the fast and slow polarisation components in blue and orange respectively. The net effect is rotation of the polarisation by 2θ, where θ is the angle between the incident polarisation and the fast axis.

Figure 6.15 shows the case that the input polarisation of the light (lower green arrow) makes an angle θ to the fast axis of a HWP, which is placed at the origin in Figure 6.15. We could decompose the overall input polarisation state into two components, one parallel to the fast axis (blue) and one parallel to the slow axis (orange). These components will oscillate

[19] So that the electric field is oscillating exactly along the axis.

as the electric field of the light oscillates, and this is shown in pale orange and blue for clarity. The component parallel to the fast axis experiences the lowest refractive index and travels fastest, and the component parallel to the slow axis experiences a higher refractive index and travels more slowly. As the waveplate is a half-wave plate, the slow component is delayed by a half-wavelength, such that the output electric field of this component (upper dashed orange line) is 180° out of phase compared to the input field (lower dashed orange line). The total effect is that the overall output polarisation (upper green arrow) now makes an angle of $-\theta$ to the fast axis. Therefore, **by rotating the waveplate by an angle θ, we rotate the polarisation of the incident light by 2θ.**

Linear to Circular

Another common use of waveplates is the use of a quarter-wave plate to turn linearly polarised light into **circularly polarised light**. As shown in Figure 1.11, circularly polarised light has a rotating polarisation vector. This arises in an analogous way to that described above, where delayed phase components cause the change when the waveplate is at certain angles. However the analogous diagram is substantially more challenging to represent in a static picture! Good theoretical treatments are given in[8, 1] and practically it can be achieved by placing the QWP with the fast axis at 45° to the incoming linear polarisation.

Turning circularly polarised light back into linearly polarised light is accomplished in the same way – pass the circularly polarised light back through a QWP. Furthermore, rotating the QWP to an angle other than 45° will result in the creation of **elliptically polarised** light, somewhere between linear and circular polarisation. The ellipse can then be rotated using a HWP before the QWP to control the direction of the light entering the QWP. These polarisation states are useful in some experiments, especially *circular dichroism* spectroscopy where the difference in absorption between left- and right-handed circularly polarised light is measured.

6.3.3 Other Polarisation Optics

Polarisation optics all fall into a category of either being a waveplate (something that modifies polarisation) or a polariser (something that selects different polarisation states). However, there are some with specific names that you'll likely encounter that it it is worth being aware of.

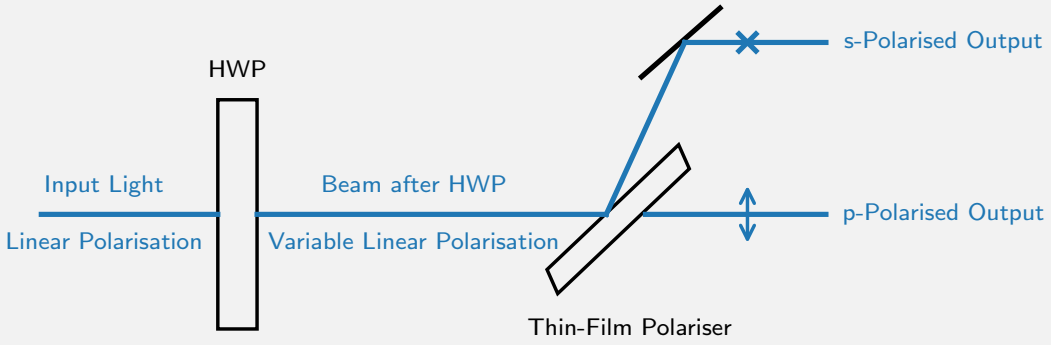
Optical Isolators

An **optical isolator** is a device that only allows light to pass through it in one direction - like a one-way gate for light[20]. They are used to isolate parts of experiments from each other, as they prevent back-reflections from travelling back along a beam path where they might interfere with or damage another part of the experimental setup (i.e. preventing feedback).

[20] Sometimes they're called 'optical diodes' for this reason, by analogy with electronics.

Application 6.8: Variable Attenuation

One of the most common uses of a polariser and waveplate in ultrafast laser labs to provide variable power control, or *variable attenuation*, of a linearly-polarised laser beam. This is easily achieved by using a HWP and a TFP together, as shown in the figure below. The incoming beam first passes through the waveplate, which can rotate the polarisation arbitrarily, and then is incident on a thin film polariser. If the waveplate is rotated such that all the incident light is s-polarised, then all of the light will be reflected from the polariser, and vice versa. The end result is that two beams are produced (one from the reflection, and another from the transmission), and the ratio of the reflected to transmitted power is controlled by rotating the waveplate.



This setup for power control is relatively easy to build, but pre-built assemblies to perform this variable attentuation are also widely available from different manufacturers. The advantage of using thin-film polarisers here is that they can withstand high power, and therefore this setup can be used to attenuate high power beams. It is also ideal for ultrafast applications provided that a broadband TFP is used. A possible drawback of this setup is that if the polariser has a 99% extinction ratio, then 1% of poorly polarised light will come through even at the 'minimum' power. This 1% could be quite a lot of laser power if the input power is high enough, so stacking multiple attenuators together such that the output of one feeds into another is a good way to get very precise control of the output power. Waveplates can also be mounted in computer-controlled rotation mounts that allow very fine, repeatable tuning of the power. For example, you could build a two-stage attenuator using a manually adjustable 'coarse' stage initially, and then a computer-controlled 'fine' stage, to obtain very fine, very repeatable control over the output power.

An optical isolator is made of two polarisers and a **Faraday rotator**. The Faraday rotator is an optical device that uses the *Faraday effect* to rotate the polarisation state of light. The

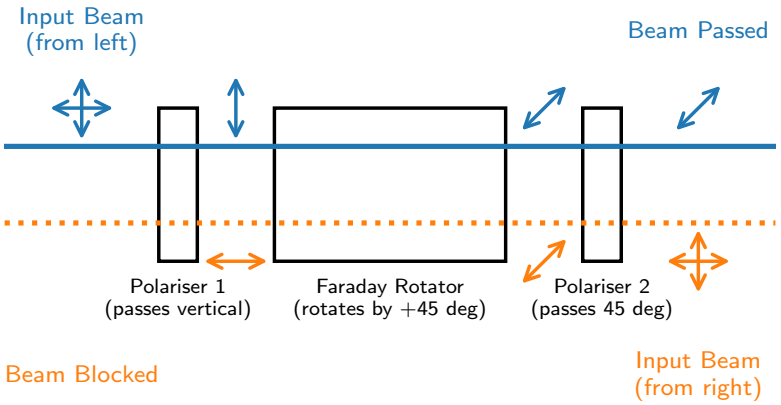


Figure 6.16: Schematic of an optical isolator. Two polarisers set to pass light at 0° and 45° surround a Faraday rotator. A vertically (defined as 0° here) polarised beam coming from the LHS (blue, solid) is passed by both polarisers. A beam of any polarisation coming in from the RHS is blocked by polariser 1.

.

Faraday effect relies on a strong magnetic field being applied to an optical material, and so Faraday rotators consist of a magnet around an optic – it is a **magneto-optical** effect. The polarisation rotation effected by a Faraday rotator is in the same direction regardless of the way that the light enters. The way an optical isolator practically works is demonstrated in Figure 6.16. The Faraday rotator is sandwiched between two polarisers, and consideration of the effect this has on light entering from the front or back of the isolator is clear: light entering from the front will pass through unimpeded, but light entering from the back will end up with it's polarisation state orthogonal to the input polariser, and be blocked.

Pockels Cells

A **Pockels Cell** is an electro-optical modulator that can function in a similar way to a Faraday rotator – an applied electric field will change the refractive index of the medium inside it and thus create different kinds of phase retardance (causing waveplate-like behaviour). Unlike a Faraday rotator (which is a static device), a Pockels Cell is usually controlled by a pulsed electronic driving circuit, allowing the polarisation state to be changed rapidly. They are commonly found inside regenerative amplifiers (Figure 4.8) as the electro-optical 'gate' that is used control when a pulse enters and exits the amplifier cavity.

6.4 Frequency Manipulation

Manipulating the frequency of an ultrashort laser pulse is often desirable so that we get the colour of light we want for our experiment. Usually, this is performed using nonlinear optical means, as discussed in chapter 3, but there are some other ways discussed here. The discussion around nonlinear optical methods is kept fairly brief to avoid excessive repetition.

6.4.1 Harmonic Generation

A very common way to generate different colours from your laser than the fundamental output is to generate **harmonics** of the fundamental. This is achieved by nonlinear frequency mixing in a crystal, the mechanism of which was discussed at length in section 3.2.

The simplest schemes involve sum-frequency mixing in a nonlinear crystal, such that two photons of the input beam(s) are converted into a single photon at higher energy. In this way the second, third, and fourth harmonic of a fundamental can be readily produced[21]. All that is necessary is an appropriate nonlinear crystal cut for the process you want to do. When you buy a crystal, you will see that it is pre-cut for a specific type of frequency mixing - such as SHG or THG at a specified wavelength. This means that when a wave is incident on the flat surface of a crystal, the optical axis is in the right plane such that rotating the crystal in a rotation mount will allow phase matching to be achieved. For example, you could buy a BBO crystal cut at 29.2° that is designed for SHG of an 800 nm beam. Having the crystal flat to the incoming beam will phase match the process. Phase matching angles are easily looked up in the Optical Toolbox, you just need to know the material the crystal is made of and the wavelengths of the beams you'd like to mix.

However, it is not necessary to always use the crystal at the angle it is cut for - you can simply rotate the crystal about the appropriate axis to find the angle needed, within reason. Remember also that the conversion efficiency you get scales with the length of the crystal (Figure 3.7), but that increased crystal length decreases the phase-matching bandwidth and makes the angular alignment more sensitive. Furthermore, if you are overlapping two different beams to generate the harmonic, then the polarisation states matter (they dictate the type of the phase matching), and so does the relative delay of the two beams, which will likely be different so some delay matching will be necessary - either through use of a flat piece of glass to slow down one pulse, or by use of a small delay stage.

6.4.2 Bandwidth Manipulation

Another way to manipulate the frequency of a pulse is to manipulate the pulse bandwidth. We saw some examples of this in chapter 3 with the SHBC (Figure 3.9), white light generation via self-phase modulation (Figure 3.3b) and with the simpler idea of spectral filtering (Figure 3.8). Here we will briefly discuss how these setups might practically be realised.

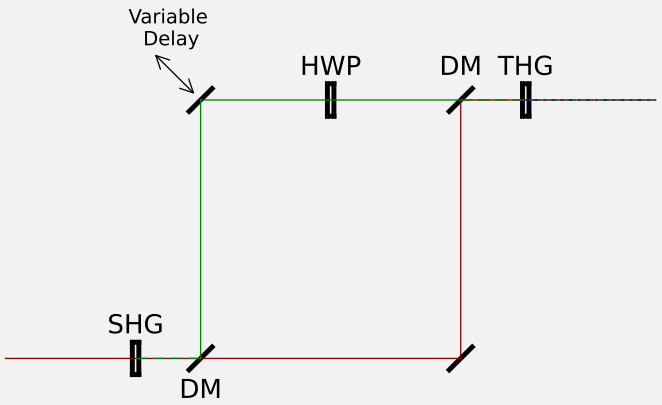
Spectral Filtering

Spectral filtering is the easiest way to reduce pulse bandwidth. There are a variety of ways it can be achieved. A simple is to pass the pulse through an **etalon**. An etalon is a finely toleranced optical cavity that, via interference of reflections from the two surfaces of the cavity, will transmit one wavelength and reflect the rest. Tilting the etalon can modify the

[21] For third and fourth harmonic generation, you would need to first generate second harmonic so that you can either combine two second harmonic photons, or combine a second harmonic photon with a fundamental.

Application 6.9: Harmonic Generation

Imagine that we have a laser emitting at 1030 nm, and that we want to generate the second (515 nm) and third (343 nm) harmonics.
A look on the Optical Toolbox will show that we can generate second harmonic in a Type-I (ooe) scheme using a BBO crystal cut at 23.4°. Such a crystal is readily available from most suppliers. We can generate this second harmonic easily. To generate the third harmonic, another look at the Toolbox will show that we need another BBO crystal cut at 32.4°, and that we need to combine one 515 nm photon with one 1030 nm photon. An experimental geometry that would facilitate this is shown below, with fundamental in red, SH in green, and TH in blue:



Note the use of dichroic mirrors (DM) to separate out harmonics, and then the use of a small translation stage to allow the delay of the 515 nm pulse to be varied relative to the fundamental by moving the mirror in the top-left corner[a], and finally a waveplate (HWP) to ensure both beams have the same polarisation heading into the THG crystal (as the generated SH light will have a polarisation orthogonal to the fundamental). Some suppliers will sell kits of optics to make setups like the one above, or 'inline' setups with all the components mounted on a rail.

[a] This is important as the different colours will not overlap in time after being split by the dichroic mirror

transmitted wavelength. Etalons are simple to use but are very expensive, and the output pulse has an exponentially decaying temporal shape (which may or may not be a good thing for you).

An even simpler way would be to pass the pulse through an absorptive or interference **filter**. A filter is ultimately just a window with a coating that will only allow certain colours to pass through it. There are a wide variety of different kinds of filter: long- and short-pass filters which will transmit all wavelengths longer or shorter than a certain cutoff wavelength; band-pass filters which will transmit all wavelengths in a specified range; wedge filters which block all wavelengths except for those in a narrow range (the 'wedge'); and notch filters, which transmit all wavelengths except for those in a narrow range (the 'notch'). Filters can work by either reflecting or absorbing the non-transmitted light. Reflective filters are less easily burned (they don't have to absorb large amounts of light), but also create more stray beams that need to be properly blocked.

The quality of a filter is determined by how much of the undesired light is able to pass, and also by the 'sharpness' around the cutoff wavelength. A higher-end filter will be able to cut much more sharply, and will block more of the undesired light. Filters are also made of glass, so the usual considerations about GDD also apply. Generally they not appropriate for high-power applications.

A more involved way to filter signals is to use a **4f filter**. This is an experimental setup that is similar to a compressor or pulse shaper, and involves the dispersion of the light into different colours. In the case of a pulse shaper, a different phase was applied to each colour to produce a different output pulse. In the case of the 4f filter, a slit, or razor blade is placed into the dispersed light to block unwanted colours. The transmitted bandwidth then depends on the width of the slit.

White-Light Generation

We discussed white-light generation in the context of self-phase modulation previously (Figure 3.3). The fundamental idea is simple: focus light into a material and let self-phase modulation[22] do the work of the broadening. The issues that can arise here are:

- Ensuring you don't burn the the material you are focussing into.
- Ensuring that you don't massively destroy the beam quality.
- Having a way to separate out the white-light from the remaining fundamental.

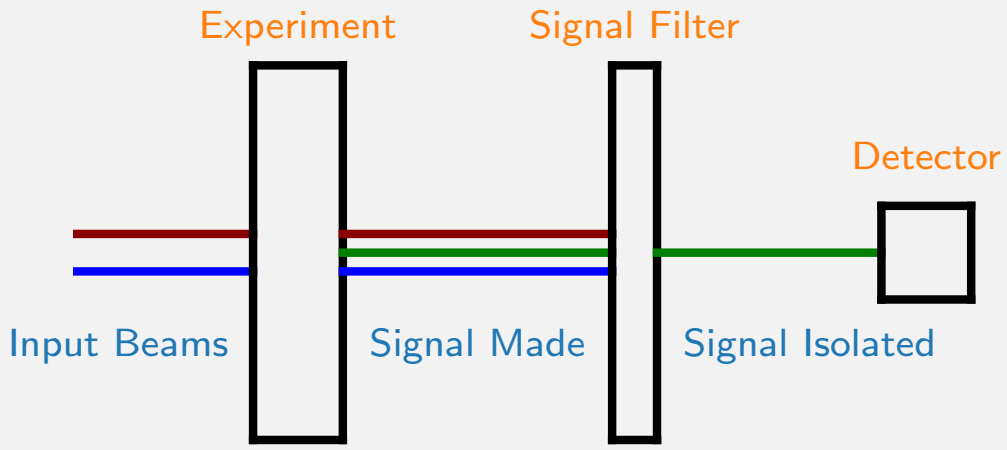
The first point is usually remedied by using a material like sapphire which has a high damage threshold – other materials can be used too, but may require periodic rotation to avoid damage. The second point is usually controlled by using an **iris** before the focussing to ensure that you generate a nice *single filament* in your generated white light[23]. Separation of the white light from the fundamental is usually accomplished with a notch filter that will knock out the residual fundamental.

[22] Among other things...

[23] If you try and build a setup for this, what I mean here will become obvious. A closed down iris and low input power will generate a nice golden 'halo' from the material. Increasing the power will turn this into a big messy colourful splat that you can't get a good beam from.

Application 6.10: Signal Separation

Many kinds of nonlinear spectroscopy require beams to be are overlapped onto a target, and their combined interaction inside the target produces a beam of a different colour as the desired experimental signal. This beam is generally a much lower intensity than the other two beams, so the other beams need to be blocked for signal beam to be recorded. A filter is ideal for this job, as illustrated in the sketch below.



As the beams here are generally low enough power to not damage an experimental sample, then a simple filter will usually suffice without causing damage. Similarly, as this filtering happens *after* the experimental interaction, normally we are not too concerned about accumulated GDD.

If you care about the pulse duration of the generated white light (and you probably do – or at least it is probably best to assume you do at first), then using curved mirrors for the focussing (especially on the output side) is preferable. It's important that the focal lengths of each mirror are the same so that the light is collimated at both the input and the output, and that you align the sapphire so that a square face is presented to the input beam (to avoid refraction). Finally note that you need to steer the generated white light

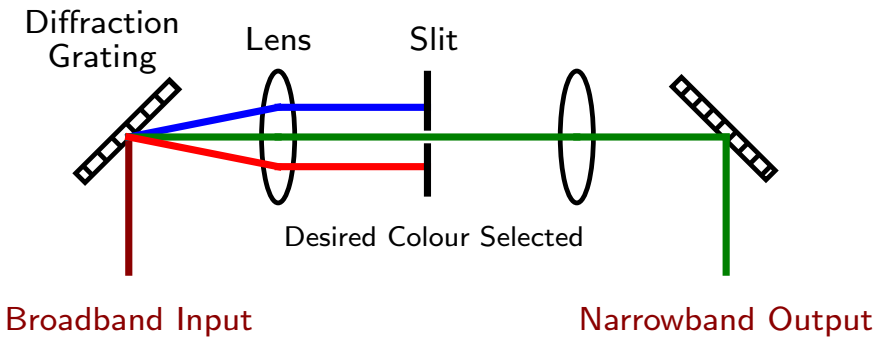


Figure 6.17: Schematic of a 4f filter. Light is dispersed into frequency components and then the desired component is selected using a slit or similar. It's called a '4f' filter because the length of it is 4 focal lengths of the lenses used. Note that the colours are chosen for illustrative reasons only!

using metallic mirrors!

SHBC

The more elegant method of bandwidth reduction – *second-harmonic bandwidth compression*, or **SHBC** – was introduced in chapter 3 (Figure 3.9). Building an SHBC from scratch is possible but probably more involved than most of us would want, so a schematic layout isn't included here.

However, one method of nonlinear bandwidth reduction that is similar to SHBC is worth mentioning. Recall that the phase-matching bandwidth of a nonlinear crystal is inversely proportional to the crystal length – so a long crystal will only phase-match a narrow bandwidth. Performing SHG on a pulse in a long crystal will actually produce a more narrowband output, because only a narrow bandwidth is effectively phase-matched[24]. This can be a cheap way of getting effective bandwidth reduction – and the resulting pulse shape is actu-

[24] See reference [9] for an example of this.

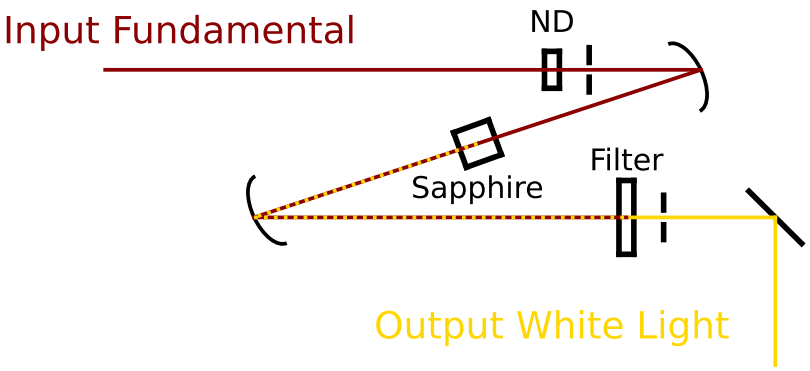


Figure 6.18: Schematic of white-light generation in sapphire (or another medium). A fundamental is focussed into the material using curved mirrors (a lens would also work, on the input side). White light is generated in the sapphire and separated from the residual fundamental using a filter and iris. A variable ND filter ('ND') and iris are necessary at the input to control the white light and ensure a stable filament is obtained.

ally a *reverse etalon* pulse shape (like a rising exponential), which can be beneficial in some applications.

6.4.3 NOPA

Finally, the Queen of methods in in nonlinear optical frequency conversion is obviously optical parametric amplification: **OPA**, as discussed at length in chapter 3. Building a homemade high power OPA is possible, but, like the SHBC, is a more involved process than most readers of this book would be interested in[25]. However, it *is* more feasible to build a device called a **NOPA** - mentioned in chapter 3.

A schematic of an example NOPA is shown above for interest, and references [10, 11] provide more concrete details if you want to build one. A NOPA consists of a WLG stage and then one amplification stage in a nonlinear crystal. Usually they're used to generate visible light (but don't have to), and so the amplification sstage pumped using the second harmonic of a Ti:Sa or Yb laser output[26]. Manually adjustable, a NOPA is a good and cheap way to generate different colours provided that:

- You don't need huge conversion efficiencies (10% would be impressive).
- You don't need to rapidly scan between different colours.
- You don't care enormously about things like beam quality.

[25] I've never made one, and I know someone (much cleverer than me) who has: to get it working properly took him the best part of 18 months. He concluded it would have been easier, cheaper, and much better to just buy it from Light Conversion.

[26] Remember that an OPA will always produce photons of a *lower* energy than the pump.

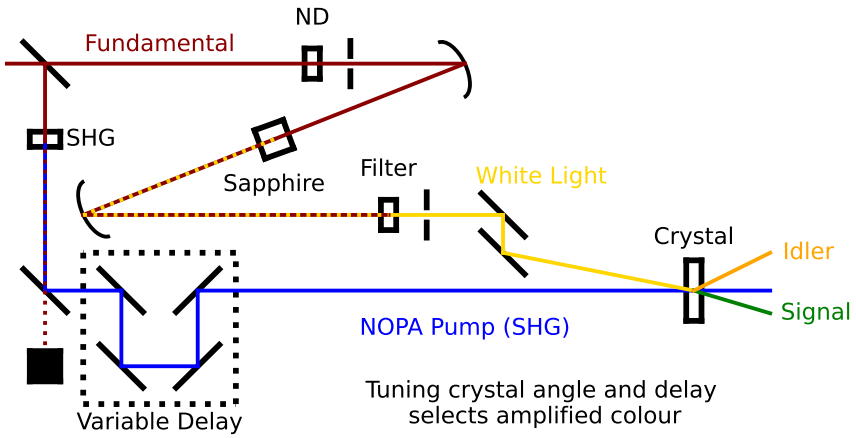


Figure 6.19: Schematic of NOPA pumped using the second harmonic (SH) of a driving laser. Note the presence of a WLG stage as shown in Figure 6.18. The white light and SH pump are mixed in a non-collinear geometry in a crystal, generating signal and idler via an OPA process. The angle between the white light and SH is chosen to maximise the phase matching bandwidth of the NOPA. The NOPA is optimised for a specific colour by tuning both the crystal angle and delay.

Of course, you could motorise your NOPA and program in some tuning curves, but normally this isn't how they're used. Aside from being a budget-friendly option for broke spectroscopists, NOPAs are also used because they can produce very short output pulses due to the wide phase-matching bandwidth they have – you'd normally need a compressor on the end of it to achieve this.

6.5 Power Manipulation

Finally, we come to the issue of manipulating the pulse energy. We have actually already mentioned some methods for doing this, like the common waveplate-polariser combination (Example 6.8). Other methods like chopping have also been mentioned, but these are less common. There are a couple of other common methods worth mentioning.

6.5.1 Neutral Density Filters

A **neutral density filter**, or **ND filter**, is a filter that acts equally across all wavelengths. ND filters act to attenuate all colours in their (very broad) bandwidth equally (hence the 'neutral'), so are ideal for use when simple and fast reduction of laser power is needed, such as for alignment or diagnostics. They too come in either reflective or absorptive variants, and can be made of a variety of substrates to suit different applications. ND filters are fairly easily burned, but are also fairly cheap and come in kits containing multiple filters. The degree of attenuation is marked on the filter as the **optical density**, and this is generally a number between around 0.1 and 10. The power of a beam passing through an ND filter with a given optical density (OD) is reduced by 10^{OD} after transmission. For example, an

ND filter marked with an OD of 2 will reduce the power by factor of 100. Multiple filters can be stacked to provide greater attenuation, and in this case the largest (highest OD) filter should be the one that the beam hits first, to minimise the possibility of damage.

A useful type of ND filter is the **variable ND filter**, or **ND wheel**. These are circular filters where the optical density increases around the circumference of the wheel, and can be used to provide a straightforward way to variably attenuate beams. This is especially useful in the UV region, where the waveplate-polariser variable attenuator previously mentioned is difficult to make and would add a lot of GDD. It is also useful if using a widely tunable source, where a waveplate-polariser attenuator would not function over a wide enough bandwidth to be useful.

6.5.2 Irises

A final method that is sometimes encountered is using an *iris* to physically block the outer portions of a beam, thus reducing the power. Irises are incredibly useful pieces of kit for alignment, and for isolating signals from messy outputs from WLG or OPA stages, but using an iris for power reduction of a beam is normally not to be recommended, especially for high power beams.

The issue is that closing an iris around a beam can cause diffraction, which can lead to hot spots in the beam and eventually damage optics. Furthermore, if the iris cuts the beam down a lot, then it can create considerable divergence that needs to be corrected. However, you can sometimes use an iris when you have a low power beam, or don't care about the beam after the iris (e.g. if it is right before a detector), or if you need to clean up very messy beams. Generally though, use irises for alignment only.

6.6 General Optical Considerations

Finally, in this chapter, we discuss some general points around optics and optomechanics. Almost all optical elements consist of two main parts: the **substrate** and the **coating**. Broadly, the substrate makes up the bulk of the element, providing structural rigidity and providing a surface onto which a coating can be applied.

The substrate can either be an optical element in it's own right, manipulating the light directly, or it can simply be the carrier of a coating which manipulates the light. An example of the latter case would be in a metallic mirror, similar to what you would find in a typical household mirror. Here a glass substrate has a metallic coating deposited on it, and this coating is what makes the mirror reflective. Without the metallic coating, the glass substrate is almost entirely non reflective, and does not work as a mirror. In contrast, an example of the former case would a lens in a pair of glasses. Here the shape of the glass in the lens produces the focussing behaviour, improving the vision of the wearer. An applied coating doesn't change or alter this primary function, but could enhance other aspects of the performance, such as an anti-glare coating. We will now discuss some important aspects of both substrates and coatings.

6.6.1 Optical Substrates

The substrate is the bulk of any optic (physically), and provides it with structural rigidity. The substrate is almost always a kind of optical glass or crystal. In some applications, the choice of substrate is relatively unimportant (the coating does all of the actual useful manipulation!) - but there are some cases where it can be important.

Transmissivity and Dispersion

A key area where the choice of substrate is critical is when transmission of a beam **through** an optic is desired. This will happen, for example, if you were focussing a laser beam using a lens, or if you want to separate a beam containing two different colours of light using a dichroic mirror[27]. The transmitted beam has to pass through the substrate, which could be relatively thick (6 mm is common for mirrors). It is therefore important that a) the substrate does not absorb lots of the light, leading to loss of power; and b) the substrate does not add a lot of GDD, leading to pulse broadening.

Thankfully this information can be easily found - most manufacturers will show the transmission characteristics of various substrates, and the GDD can be calculated in ways discussed in chapter 2. In the case where avoiding pulse broadening is important, it can be beneficial to use a very thin substrate - these are readily obtainable but substantially more fragile. Table 6.2 lists some common substrates and their transmission characteristics, with common applications.

Table 6.2: Transmission and Dispersion performance of common optical substrates.

Substrate	Transmission Range ($>$ 90% for 10mm thickness)	Dispersion
BK7	350 nm - 2.0 µm	IR: Good - UV: Poor
UV Fused Silica	200 nm - 2.0 µm	IR: Good - UV: Reasonable
CaF_2	180 nm - 8.0 µm	IR: Good - UV: Good
Sapphire	Never $>$ 90% - reasonable in UV + near IR	IR: Good - UV: Poor

Clearly, if we want to be sending ultrashort UV pulses through optics (which is generally to be avoided anyway!) then we need to use a UVFS or CaF_2 substrate. However, BK7 is generally fine for wavelengths longer than around 400 nm. You may reasonably ask why we don't just use CaF_2 for everything, given it's favourable optical properties. The reason is largely that it is more expensive, comparatively fragile, and harder to machine. For these reasons, it only tends to be used where it is truly demanded. Sapphire has relatively poor transmission qualities but a very high damage threshold (to both thermal and mechanical damage) compared to other materials, so is useful in some applications for that reason.

A final point to note is that if you want to send a beam *through* an optic, you should make sure that the rear face as well as the front face of the optic are both polished to optical

[27] A mirror that reflects one colour but allows others to pass through.

quality. Some optics are not polished on the rear face (especially mirrors), and so a beam that passes through the coating will simply scatter (diffuse reflection) off the rear face, so the transmitted beam is poor quality (if it exists at all). Conversely – if you do not need to send a beam through an optic, then having a ground rear face (not polished) is safer, as there are fewer reflective surfaces that can result in stray beams flying around the lab.

Damage

Another key property of an optical substrate is in how easily damaged it is. Damage can be mechanical: glass can be scratched, and optics will get dropped; or thermal: high power laser beams are incident on the optic and could burn it. The optical substrate needs to be able to deal with all these situations without becoming damaged.

Mechanical damage for our purposes will mostly be asking the question 'how easily can I scratch the substrate?'. The material property that governs this is the **hardness**. Most substrates have reasonably high hardness - but sapphire is exceptionally hard and difficult to scratch. This means it can be made into very thin optics - which can offset the relatively poor transmission properties. BK7 also has a reasonably high hardness (higher than UVFS and substantially higher than CaF_2).

Thermal damage to optics by the incoming laser beams is very common - and will be discussed further in the following section as it mostly damages coatings, rather than substrates. However, of importance in some applications where high energy beams are used is **thermal expansion** of a substrate via heating by a laser beam. This can cause the size/shape of the mirror to change as it is heated, which can cause position of beams to drift as the substrate heats up. One way around this is to use bigger beams, and larger optics - so that the heat is distributed over a wider area. Another is to use a specific substrate like **Zerodur**, which has a very low coefficient of thermal expansion, but comes at an additional cost.

Substrate Sizing

Optics come in a range of sizes, both in terms of substrate thickness and active surface area (the surface area that the beam is incident on). Standard sizes range from 0.25 inch (6.35 mm) to 2 inch (50.8 mm) diameter. Optic diameter is inexplicably an area where imperial units still reign supreme (with even metric optical mounts being designed to hold optics with imperial diameters).

Smaller optics are generally used in compact applications, and inside lasers where the geometries are very precise. Larger optics are used for higher power applications and larger beams, such that the power in the beam is spread over a larger area, reducing possibility of damage. In the author's experience, 1 inch diameter optics are a good compromise between cost and ease of use (a larger optic is often easier to align), and seem to be the 'standard' size in many laboratories. It goes without saying that the diameter of the optic needs to be larger than the beam diameter, and having it be around twice as large is often desirable for ease of alignment.

Optical thickness tends to increase with diameter, as the mirror needs to be structurally rigid. Thin optics are desirable for applications where transmission *through* an optic is desired (to minimise power loss and pulse broadening), but this comes at the expense of rigidity. Thin optics are prone to warping or cracking if overtightened in an optical mount - the characteristic sign of this is the beam starts to focus unexpectedly, as the optic becomes curved. Not over-tightening thin optics in mounts avoids this. Standard thicknesses are normally around 6 mm for 1 inch optics, but thinner optics designed for transmission and reflection of ultrashort pulses will generally be nearer to 3 mm.

Cost

The final property of substrates that is important is their **cost**. Optics get damaged, and will need replacing. You will find that UVFS and BK7 are among the most common substrates for optics, and this is due to a combination of reasonably high damage thresholds and reasonably low cost. CaF_2 or Zerodur substrates are more expensive. Many manufacturers also offer 'economy' mirrors, made from cheaper glass. These may work fine for some applications (or applications where mirrors need frequent replacement), but checking the specifications before buying is worth the time.

6.6.2 Optical Coatings

On many optics, the coating is what turns the substrate from a plain piece of glass to a useful optical element. Generally the flat substrate is extremely well cleaned, and then the coating is deposited via a technique such as electron beam deposition. The coating can affect a multitude of different optical properties, such as reflectivity, dispersion, and polarisation. The most immediately obvious kind of coating is a metallic coating used to create mirrors – as will be familiar from general household mirrors. Metallic coatings can be useful when making mirrors for ultrafast optics, as was discussed in section 6.1.1. The other kind of coating mentioned previously is the **dielectric coating**. Dielectric coatings can produce a very wide range of different optical effects, and so an overview of their general properties is given here.

Dielectric Coatings

At its core, a dielectric coating consists of a number of thin layers of material (normally glass or crystal) each with different refractive indices stacked onto a substrate. This construction can create a lot of different behaviours, but the working principle is best illustrated by considering how a reflective dielectric coating for a mirror can be made.

Reflection of incident light will occur off the boundary between the layers that make up the coating[28], but between any given pair of layers the proportion of the reflected light is relatively low. Stacking many layers, and then ensuring that the separation of layers is such that reflections from different layers all end up in phase, increases the reflectivity as the

[28] Reflection will occur off the boundary of any two materials with different refractive indices, but the proportion of reflected light may be very low.

reflections from different layer boundaries constructively interfere. In this way, reflectivities of >99.99% can be achieved. The operating principle is illustrated in Figure 6.20.

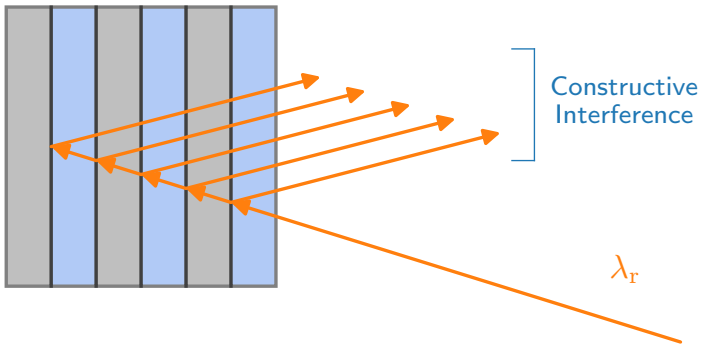


Figure 6.20: Sketch of the operating principle of a high reflectivity dielectric coating. Incoming light of a wavelength λ_r (orange lines) is reflected off boundaries between two materials of different refractive indices (grey and blue shaded). The separation of materials engineered such that reflections from multiple layers interfere constructively - resulting in high reflectivity.

Figure 6.20 is merely an illustrative sketch of this principle - in reality the layer construction is much more complex to allow a wider range of wavelengths to be effectively reflected (increasing the reflected bandwidth). A very wide range of different coatings can be produced to give almost any specification desired - a look into the catalogues for Layertec will illustrate this.

Another very common dielectric coating is the **anti-reflection coating**, or **AR coating**. This coating is designed in an analogous way to the reflective dielectric coating shown in Figure 6.20, except the layers are engineered such that the reflections from differenet layers end up out of phase with each other and interfere **destructively**, and the coating would not reflect any of the input wavelength. AR coatings are commonly applied to the surface of transmissive optics to stop them reflecting light of the desired transmission wavelength. This unwanted reflection would cause both power loss, and a safety hazard, as it is another stray beam in the lab that needs to be accounted for.

Manufacturers that produce optical coatings often produce a very wide range of them, covering many possibilities and useful wavelengths. Typically within a specific manufacturer the coatings have a systematic naming convention (such as Thorlabs' ´A' to 'G' AR coatings). These conventions can be useful, as (for example) in a lab using Ti:Sa lasers,

you'd learn that generally transmissive optics will need the 'B' coating. Speciality mirrors mentioned previously – like dichroic mirrors, beamsplitters, low-GDD mirrors, and chirped mirrors – are all dielectric mirrors with special coatings engineered for specific applications.

Damage and Cost

Generally in scientific applications optical coatings are deposited onto the substrate and then left exposed, without a further protective substrate being applied on top (in contrast to say, a household mirror, where the reflective metal surface is enclosed behind glass). This is to avoid extraneous reflections from the additional interfaces this would create, and in the case of ultrafast optics avoids accumulation of excessive GDD by passing through the front substrate. The downside of leaving the coatings exposed like this is that they are then more easily damaged, much more easily damaged than the substrate itself.

Dielectric coatings are generally much more hard-wearing than metallic coatings, both from a hardness perspective and a thermal damage perspective. All coatings are particularly susceptible to thermal damage if they are not kept clean - any dust or dirt on a coating will absorb energy from an incident beam much more readily, and can quickly result in damage to the coating. Whilst being harder-wearing (and often having superior optical properties, as discussed later in this chapter), dielectric coatings are generally much more expensive than metallic coatings.

6.6.3 Optical Labelling

Optics are usually labelled (on boxes) using the acronyms **HR** ('high reflectance'); **HT** ('high transmission'); and **AR** ('anti reflective'), depending on the coatings applied and intended applications. For example, a mirror designed to separate out 400 nm Ti:Sa second harmonic from the 800 nm fundamental may be marked 'HT800 HR400', showing that it will reflect the 400 nm light and transmit the 800 nm light. Boxes may also show things like the batch number that the coating was produced from (as manufacturers can produce reflectance/GDD curves specific to each batch for demanding applications), or the designed AOI and polarisation. They will also show information such as the substrate size, thickness, and material.

Another aspect of optical labelling is that which is written *on the optic itself*. Sometimes lots of information is laser engraved onto the edge of the substrate such that even without the original packaging all the necessary info (sometimes including batch numbers) can be found on the edge of the optic. There will also often be an arrow pointing to which side of the substrate the coating is on. This may be laser-etched, but could also be a simple pencil mark, depending on manufacturer. Good practice when taking a new optic out of it's packaging for the first time is to check whether or not it has the function of the optic marked on the edge. If it does not, then mark the function on with a pencil, as shown in Figure 6.21. Writing the *function* and not the part number is best practice here, as part numbers change, and are usually useless if the manufacturer is not also labelled. An unlabelled, or badly labelled optic is **less use than a piece of plain glass** in a working lab.

Figure 6.21: A labelled mirror showing clearly the design wavelength, designed angle of incidence, and coated side.

6.6.4 Optomechanics

The final thing to discuss in this chapter is **optomechanics**, which is a catch-all term for all of the things that hold and manipulate the optical elements on a laser table. A lot of manufacturers of optics also manufacture optomechanics, and there is a dizzying array of different things available to allow construction of different kinds of beamlines. Experienced campaigners will have their own opinions as to who makes the best optomechanics, so asking around your lab/department is well worth the time spent. Many sell pre-assembled mounts for things like periscopes and delay stages, which can save a lot of time and hassle making and aligning these assemblies from scratch (albeit at an increased cost). Exhaustive discussion of all the different types of optomechanics will not be given here, but a short overview is given together with some some potential pitfalls to avoid when buying.

Anatomy of a Mounted Optic

Figure 6.22 shows a CAD drawing of a mounted optic. In this case, it is a mirror mounted in a kinematic mount, attached to a post and post holder, which are clamped to a laser table. We will discuss this from the top down to illustrate the purpose of mounting optics in this way.

Initially, the mirror (here shown with a gold face) is mounted into the mirror-shaped recess in the kinematic mount. The mirror is secured in place using a small set screw (normally an M2 screw[29]), which is tightened with a hex key. These set screws tend to have a small piece of rubber or felt on the end of them to avoid scratching the optic. Care should be taken to not over-tighten these set screws, as doing so can cause very thin optics to warp - which can cause the beam to inadvertently focus or become elliptical. The entire kinematic mount is attached to a optical post - in this case a ½ inch stainless steel post - with either M4 set screw or M4 socket head bolt. The post has a hole in the body of it that a hex key can be inserted into to aid in tightening the mount to the post.

[29] The ISO metric screw thread system is most commonly used. The number refers to the outer diameter of the thread in millimetres. An M2 screw has a 2 millimetre diameter thread.

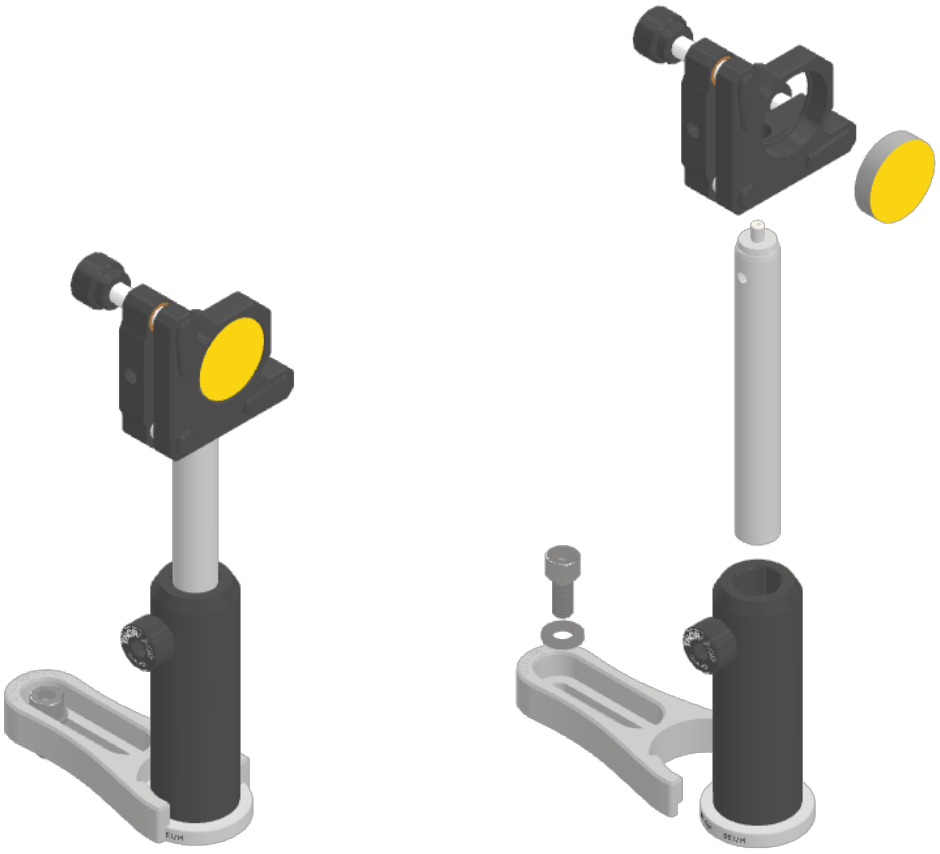


Figure 6.22: CAD drawing of a typical mounted mirror shown both complete (left) and exploded into constituent parts (right).

The post is the inserted into a post holder. The post holder is anodised black to minimise reflectivity and the post is tightened inside the holder using a thumbscrew. The post can be slid up and down within the post holder to allow the height of the mirror to be adjusted, and can also be rotated within the holder for coarse adjustment of the mirror angle (fine adjustment is taken on the thumbscrews on the kinematic mount itself). The post holder can be attached to the table in a variety of ways, but the most convenient way is through using a pedestal mount and clamping fork. The pedestal mount has an M6 stud on it which screws into the base of the post holder. The optic is then placed in the desired position on the laser table, and a clamping fork is placed over the pedestal. The fork is then bolted to the table using an M6 cap head bolt. This holds the entire mounted optic securely in place, but also makes moving the optic by undoing the bolt in the clamping fork relatively straightforward. Various lengths of fork are available to suit different requirements.

There are other ways to mount optics than this, but they all follow roughly the same principles - a holder is securely bolted to the table, and the optic is attached to a post which is placed in the holder that allows the height and angle to be easily adjusted. Some optical mounts do away with the separate post holder and simply clamp a thick optical post straight to the table. This is more stable, but does away with any easy height or angle adjustment - it is commonly found within laser cavities where the stability is necessary.

Tips for Buying Optomechanics

When buying optomechanics you should first consider what kind of stability you need for your specific setup. Most manufacturers make a range of different 'tiers' of optomechanics, often with a standard, cheaper tier, and a higher-end expensive tier[30]. Optomechanics from the higher-end tiers are made with greater precision and designed to give better long term stability, and mounts that feature kinematic adjustment (i.e. the ability to move the optics) will allow finer and more repeatable control. The downside is a vastly increased cost, so if these are not truly needed, then using cheaper optomechanics can be better[31]. It is possible to mix-and-match, so that you use a high-end ultra-stable mount for overlapping your beams (allowing you finer control of the overlap position), but cheaper mounts elsewhere. Specifications and benchmarks for the stability of things like translation stages are available from manufacturers, so you can see if they will work for your application.

Another thing to consider is ease of building. I would generally advocate that penny-pinching on cheap things that are more difficult to use can be a false economy - the cost of the time you can spend trying to build things from scratch can often outweigh the cost of buying something pre-assembled in the first place. This is especially true in optomechanics, and so spending slightly more money on (for example) a mount for delay stage mirrors that is cut at exactly 90° is generally a wiser investment that possibly having to spend many hours ensuring that your two individual mirror mounts are precisely at 90°. Of course, sometimes budgetary constraints require you to be more creative, but do not feel that spending extra

[30] Like Thorlabs' standard and *Polaris* range.

[31] In fact, I am mostly of the view that the instability of the laser itself is almost always more of a problem than drifts in the optics, and that 9 times out of 10, you don't need the fancy mounts...

money to make life easier for yourself is somehow a cop out - we are non-specialists trying to use lasers to answer our other research questions, not laser physicists!

A final thing to be aware of are the standards for optic sizes. Optics of a 1 inch diameter are commonly used, but many other sizes are also used, so ensuring that you're buying the correct mounts is important. Equally, if mounts attach to posts and holders using screws, then ensuring that you have the right screw threading is critically important. Whilst most of us in the world use ISO standard metric 'M' designated screw threads, people in the USA commonly use the ANSI standard UNC and UNF imperial screw threads - and the threads are generally not compatible with each other. Manufacturers make things for both markets, and so will normally offer both imperial/metric options on most items. Bitter experience tells me that taking the time to check you're ordering the right things can save a lot of future headaches[32].

6.7 Manipulation Miscellany

6.7.1 Getting Light to a Sample: Windows

Windows are transmissive elements that allow light to pass through them. They are a necessity if you are bringing your beams into a vacuum chamber, or passing it into some other piece of experimental apparatus – a cuvette is basically two windows that surround a liquid sample. Windows are perhaps the most simple optical element - a flat piece of glass that allows light to pass through it. Clearly, the important considerations are of the window material and the coating. The window material should be chosen so that it has good transmission properties at the desired wavelength, and does not incur too much GDD. Windows can also be AR coated, which is generally a good idea if your application allows it.

Windows are transmissive, and so will add some dispersion to an ultrashort pulse. This is generally unavoidable, as you cannot (for example) remove the window from your vacuum chamber to make the pulses shorter! As such, pre-calculating the added GDD from these elements is a good idea, and then you can more easily pre-compensate for it if necessary.

The only real complications in using windows arise because windows are often a structurally integral part of experimental apparatus, especially in high-vacuum experiments. In this kind of experiment, the window forms the interface between atmospheric pressure and an ultra-high vacuum. In this case the window needs to be thick enough to not be easily broken, and also needs to be mountable in a vacuum flange. A 3 mm window is normally strong enough without causing terrible pulse broadening. Various options for mounting windows to vacuum chambers are available from manufacturers. An additional consideration is that if you are focussing a beam into a chamber, take care to ensure that the beam is not so small on the window that it will drill a hole in it over time.

[32] For this reason, the lab I spent a few years working in during my postdoc years had a policy of 'if it's imperial, it goes in the bin'.

Bibliography

[1] Pedrotti F L, Pedrotti L M, and Pedrotti L S. *Introduction to Optics*. Cambridge University Press, Cambridge, 3rd edition, 2017.

[2] Paschotta R P. The rp photonics encyclopedia. https://www.rp-photonics.com/encyclopedia.html. Accessed: 22-12-2020.

[3] Paschotta R P. *Encyclopedia of Laser Physics and Technology*. Wiley-VCH, Berlin, 1st edition, 2008.

[4] Light Conversion. Optics toolbox. toolbox.lightcon.com. Accessed: 23-12-2020.

[5] Hamm P and Zanni M. *Concepts and Methods of 2D Infrared Spectroscopy*. Cambridge University Press, Cambridge, 1st edition, 2011.

[6] Monmayrant A, Weber S, and Chatel B. A newcomer's guide to ultrashort pulse shaping and characterization. *J. Phys. B: At. Mol. Opt. Phys.*, 43, 2010.

[7] M Weiner. Ultrafast optical pulse shaping: A tutorial review. *Optics Communications*, 284, 2011.

[8] Goldstein D H. *Polarized Light*. CRC Press, Boca Raton, FL, 3rd edition, 2017.

[9] Weeraman C, Mitchell S A, Lausten R, Johnston L J, and Stolow A. Vibrational sum frequency generation spectroscopy using inverted visible pulses. *Opt. Express.*, 18, 2010.

[10] E Riedle, M Beutter, S Lochbrunnner, J Piel, S Schenkl, S Sporlein, and W Zinth. Generation of 10 to 50fs pulses tunable through all of the visible and the nir. *Appl. Phys. B.*, 71, 2000.

[11] E Riedle. Nopa - overview and principles. http://nopa-agriedle.info/. Accessed: 06-10-2024.

Chapter 7

Building a Beamline

In this chapter we discuss the principles of building an experimental optical beamline. We start with general safety considerations, and then consider planning and construction of a typical beamline. 'Walkthroughs' for construction of common sub-assemblies – attenuators, periscopes, telescopes, and delay stages – are also given, in addition to guidance on finding temporal and spatial overlap between to beams.

This chapter focusses on a discussion of the basic principles of *building* an experimental **beamline**, using the optical elements and manipulation techniques described in the previous chapter. The term **beamline** refers to the arrangement of optics that allows you to control and direct the output of your laser system into your experiment. A beamline could be as simple as two mirrors directing the laser output into a sample cell, or could be very complex with multiple overlapping beams and lots of additional optical equipment to control various properties of the laser light. In all cases, there are some fundamentals that form the bedrock of effective optical design and construction, which will be codified here for reference.

Construction of beamlines is often a relatively infrequent part of everyday lab work - once a beamline is constructed and working, it is often left in place and used rather than continually rebuilt. As such, many students and researchers may work for several years without actually having to do substantial optical building work, and then often find themselves thrown in at the deep end when expected to know how to construct the thing they have been working with for many years! The aim of this chapter is to codify some methods of best practice into an accessible resource, so that somebody experienced with working with optics, but not necessarily experienced with beamline construction, can avoid some of the more common pitfalls.

7.1 Laser Safety

Before embarking on a discussion how to work with optics, it is prudent to make a point about safely working with lasers and laser light. Whilst lasers are an incredibly useful and

ubiquitous tool in experimental physical sciences, they are also **very dangerous if not used safely**. There is a worldwide system where lasers are rated from 'Class 1' to 'Class 4' depending on how powerful (and dangerous) they are. All lasers discussed in this work are Class 4 lasers, and the majority of lasers used in scientific research are either Class 3 or Class 4. These lasers have the potential to cause serious eye damage, if not blindness, if used incorrectly. A typical Ti:Sa laser beam is many billion times brighter than the sun, and even a small reflection could cause serious eye damage or blindness.

All research environments should provide laser safety training, and have systems in place to minimise the chance of dangerous accidental exposure such as interlocks and shielding. There will also normally be a designated laser safety officer who will be able to advise on safe practice, and will probably have to sign off any substantial new optical construction as safe for use. Basic principles such as always wearing safety glasses/goggles, always aligning optics with as low a power as possible, and always taking care to block any stray beams, should be standard practice. I have a retinal burn in my left eye from an unpredicted reflection from a polariser, despite taking all the usual safety precautions, and I was lucky to escape with only very minor lasting damage that doesn't impact every day life. The overall message is that **laser safety should be taken very seriously**. If you're a new graduate student (for example), you should have someone experienced to work with for a few months whilst you learn the ropes.

7.1.1 General Safety Procedures

There are a few things that we always do in the laser lab when *any* laser is on (even if it is all contained) to make it as safe as possible:

- Always wear protective eyewear.
- Always remove any reflective jewellery from your hands or clothing (belt buckles, watches, rings...).
- Always make people aware the laser is on - usually with a light or sign outside the lab.

Most of these should be enforced by a local safety officer. If we are working on a setup with many open beams, then there are some further best practices:

- Keep tools out of the beam. Use black anodised tools where possible to minimise any risks.
- Keep the beams contained in the laser table (using some kind of enclosure around the table).
- Keep all beams at stomach height, and never drop your head to the level of the table. Close your eyes if you need to duck below the level of the table for any reason (e.g. if you've dropped a hex key).

- Avoid vertical beams wherever possible, and keep all beams in one plane parallel to the table.
- Keep beams at the lowest power possible whilst aligning setups.

Keeping all beams in a single horizontal plane is useful both for safety, and for ease of alignment later. The temptation to put your head at the level of the table to examine something closely is something novice laser users need to train themselves not to do! Remember that even beams that are mostly invisible to the naked eye (like the 800 nm and 1030 nm fundamental outputs from Ti:Sa and Yb lasers) can easily cause blindness from a stray reflection.

7.1.2 Safety Equipment

Protective Eyewear

The most essential bit of safety kit we have in a laser lab is our protective eyewear, or 'laser goggles'. These are glasses have lenses that are designed to absorb certain wavelengths of light with a given optical density. Generally you would pick the glasses that absorb the light produced by your laser – 1030 nm for an Yb laser, for example. Of course, in a lab with many OPA systems or many lasers, there are likely many different colours of light flying around. It is impossible to have glasses that block all of these, and there are two schools of thought on the best approach:

- Use the glasses that block the light with the highest intensity (usually the fundamental from your laser).
- Use the glasses that block the light that is hardest to see.

The idea behind the second approach is that if you can see a load of bright green beams flying around (for example), then you are already relatively aware of where these are and so will know if they are somewhere dangerous[1]. The invisible IR beams are therefore much more dangerous, as you can't see them. The best approach for you depends on the kinds of light you have in your lab.

As it's often impossible to block every kind of light your setup can produce, it is important that you **don't rely on the glasses for protection**. The glasses are your last line of defence, and should never be tested. It is much better to design the setup so that all beams are contained.

Beam Blocks and Dumps

When trying to build a setup, you need to make sure all the beams are contained. Ideally you have the entire table surrounded with an enclosure, but tubes or similar for individual beams can also be useful (and help to minimise instability caused by air currents). Specifically,

[1] Visible light also stimulates your blink reflex, unlike invisible light.

usually at some point you'll need to 'dump' a laser beam safely somewhere – either residual pump from a nonlinear optical process, or after interaction with a sample.

To do this, it is best to use a specially designed beam block, or beam dump. These are engineered to be highly absorbing so stray reflections won't bounce off them. The simplest are often just a stack of razor blades, but others have more intricate designs. Some may have heat sinks, or even water cooling, if you have very high power beams. These can either be bolted to the table, or some may have magnetic bases so they can easily be moved around while testing.

7.2 Planning

Having understood the safety procedures, the first thing to do when building a new beamline is to make a plan of what you need to build. Thorough planning will help to foresee and avoid potential problems, so taking the time to make a detailed plan is worth it in the long run. There are three key questions to answer when starting to design a beamline:

1. What light do I need at my experimental target, and how different is this light from the direct output of the laser system?

2. How much space do I have to build in?

3. What budgetary constraints are there?

We will consider each of these questions in turn, and then show an example case study in the final chapter to illustrate the process.

7.2.1 What Light Do I Need?

This is the first question to ask yourself when building a beamline. As discussed in preceding chapters, there are a variety of different properties of the light that we can manipulate. The answer to this question will depend hugely on the exact experiment being undertaken, but the key idea is to determine how different the light needed at the experimental target is from the direct output of the laser system, as this will determine how much manipulation of the output is needed (if any). Clearly, the easiest situation would be one where the output of the laser can be directly steered into the experimental target using a couple of mirrors, but it is rarely this straightforward in practice! The parameters that need to be determined are as follows:

- **Beams**: How many beams do you need to overlap at your target?

- **Position**: Where is the target? Do you need to come in at a specific angle?

- **Colour**: What colour(s) of light need to be at the experimental target? Do you need to do any external frequency conversion of the laser output?

- **Size**: What size of beam do you need at the target? Do you need to focus the beam down, or expand it?
- **Pulse Duration**; How short do your pulses need to be at the target? Do you need to any external compression or stretching?
- **Intensity**: What intensity do you need at the target? With a fixed beam size and pulse duration, this will determine how much power you need in the laser beam.
- **Polarisation**: What polarisation do you need at the target? If you need multiple beams, do they need variable polarisation?
- **Sample State**: Is the sample behind a window? In a cuvette? In a vacuum chamber?

Taking these points in turn, clearly requiring more beams to be overlapped at the target increases the complexity, as the beams will need to be overlapped both spatially (so that each beam irradiates the same area), and temporally (so that the pulses in each beam arrive at the same time, or with a known delay). Strategies for this are outlined in subsection 7.4.5. Where the beams need to be is also a critical consideration, as in some cases it may be that you have a well-defined immovable target that you need to hit (such as the window into a large vacuum system), or other times it may be more flexible, and you may be able to move the sample around as desired. If the height of the experimental target is substantially different to the height of the laser output, then you may need to install periscopes to either raise or lower the beam height as needed.

Requiring a different colour of light at the target than is produced directly by the laser system requires the use of external frequency conversion. This could be as straightforward as some frequency doubling using a single nonlinear crystal (for example, if your laser produces 800 nm pulses, and you need 400 nm pulses at your target), or may require the use of an external OPA to allow the colour of the light to be more widely tunable. This is commonly the case in UV photochemistry, where photochemical pumping of a target across a relatively wide band of the UV spectrum is desirable. The conversion efficiency of these frequency conversion systems varies, but at wavelengths far from the pumping wavelength it is often very low. For example, pumping an OPA with around 3000 mW of IR light may produce only 30 mW of UV light. This is something to consider if you are buying a whole new laser system for your setup, as you may need to use some portion of the laser output to pump an OPA, whilst retaining some portion for use in other areas of the beamline. It could be an expensive mistake if you have insufficient output power from the system.

In a similar vein, requiring a different pulse duration than that produced by the laser output (or compensating for dispersion incurred as the pulse travels towards the experimental target) can create some challenges. Often, the pulses produced by the laser are sufficiently short, and dispersion through the beamline is minimal, so that no additional compression or stretching is necessary. However, especially with UV pulses, the dispersion incurred through the beamline can be severe enough to require re-compression of the pulse before it is used in the experiment. Alternatively, you may wish to stretch the pulses (for example, if you want

to reduce intensity but not change the power or spot size), and this will require construction of a stretcher. In both cases, building a compressor/stretcher can be a considerable expense - so it is worth taking the time to plan and calculate if one is needed so it can be factored into your budget.

The other parameters are more easily dealt with in the beamline construction. The required beam size at the target will depend on the nature of your experimental target. You may have a very dense sample, allowing you to focus the beam hard into the target and reduce the power in the beam. Conversely, the sample may be more diffuse (or you may want to use larger beam sizes for ease of overlapping multiple beams), so you want to irradiate a larger area of the sample to 'talk to' more of the sample with the light - but this may require more power in the beam. In all cases, the size of the beam at the target can be easily manipulated using lenses and telescopes. Similarly, the power of the laser beam can be easily manipulated using a variable attenuator such as a HWP-TFP combination or rotatable ND filter. Polarisation too can be varied using a relatively inexpensive waveplate, or combination of waveplates. There are more complex situations where very fine control of polarisation is required over a broad frequency range, but these are relatively rare.

Having established the properties of the light that we require, we can then start to plan out the arrangment of optics on the laser table.

7.2.2 General Principles

Planning a beamline is the first step of building a beamline, and important decisions are made in the planning stage that will affect the future operation. We will give an example of the planning of a beamline for a simple experiment from start to finish in chapter 8, and so here will aim to codify some principles of 'best practice' to keep in mind when planning.

Power Control

The direct output of our laser system may well produce vastly more power than we need in the experiment, so it needs to be attenuated somewhat. It is best to try to attenuate this power **earlier**, rather than later. The reasons for this are two-fold:

- You will more often be working with the beamline nearer to your experiment than the output, so attenuating the beam closer to the laser output means that high power beams are not flying around the space where you will often be working. It is much safer to work with the beams attenuated, and excess power safely dumped, in an area of the table you will not often be standing at.

- Higher power beams have more potential to damage optics, or cause unwanted non-linear effects like self-focussing. Dumping the power earlier minimises the possibility of this.

The safest way to dump excess power is to use a purpose built **beam dump**, which is rated to handle the amount of power you need to get rid of. Very high power beams may require

the use of a water-cooled beam dump.

Beam Height and Stability

- In general, it is safest to have all the laser beams propagating in a single plane that is parallel to the surface of the optical table, and is at around waist height.

This minimises the chance of stray reflections being reflected upwards towards our eyes. In an ideal world, then, the output port of the laser system and the height of our experimental sample will be the same, and all the beams will propagate in a plane at this height. In reality, this is not normally the case, and so adjusting the height of the beam is necessary. Modifying the height of the beam can be accomplished either using a **periscope**, or by simply shooting the beam out-of-plane upwards or downwards, to another mirror in a different plane. If the height difference is relatively small, then the second option can be fine, but for large changes in height use of a periscope is preferable. It is imperative that periscopes are properly shielded, as beams coming vertically up from the table have the potential to cause serious injury.

- Measure the height of the beam above the table that you need at your sample, and calculate the difference between this and the height of the beam above the table at the laser output. Install a periscope as near to the laser as possible to match these heights.

In addition, it is generally preferable that beams run closer to the table surface, as this minimises the length of the optical posts holding optics, which increases stability[2]. If the beam has to propagate a long distance between the laser output and experimental target, then it can be especially advantageous to periscope it down to a low height, and then bring it back up nearer the region of interest. As in some cases any small instability a long way 'upstream' can cause a larger instability further 'downstream', so maintaining stability over long propagation distances becomes critical. More expensive optical mounts are also more stable, but come at an increased cost. Shielding the laser and beam path from air currents in the room can also help stability enormously, and having the beam completely enclosed is much safer too.

Beam Size

- Choose optics that are at least twice as large as the beam diameter. If needed, shrink the beam to fit onto smaller optics for ease of alignment.

The size of the laser beam dictates what diameter the optics in the beamline need to have, and this can be an important consideration as large optics can be considerably more expensive than small optics. Typical Ti:Sa laser systems can produce relatively large beams, so the direct output may be too large to comfortably fit on 1 inch optics without clipping, and require the use of large 2 inch mirrors or beamsplitters. In this case, it is often

[2] As a longer post has a longer moment arm between the optic and table.

advantageous to use a telescope to reduce the beam diameter such that it can readily fit onto a 1 inch optic, as constructing an entire beamline using 2 inch optics could quickly get prohibitively expensive. A telescope would also help to control any divergence in the laser output if propagation over a long distance is desired. However, an important consideration regarding beam size is that a larger beam can be appropriate for very high power beams, as the larger beam diameter spreads the power over a larger area, minimising the chances of optical damage.

Ease of Alignment

A final consideration when making an initial plan for an optical beamline is to consider how easy it will be to align on a day-to-day basis. Laser systems drift, and the pointing stability of the laser may not stay constant over time. This means that even if you perfectly aligned all of the optics in the beamline once, they will probably not stay perfectly aligned and will require periodic readjustment. The beamline can be built to maximise the stability by running beams nearer to the table, and using heavier duty (more expensive) optical mounts; but it is also important to build the beamline such that getting it back into alignment *after* it has drifted is as pain free as possible. There are a few principles which help achieve this aim:

- The optical table has a grid of M6 tapped holes on it, and making sure that (as far as possible) beams run along the lines of the boltholes, and that beams run orthogonal to each other, will make it easier to see when the beam has drifted out of alignment, and give you a rough target to aim for to get it back into alignment.

- Using alignment targets or irises in critical areas can help you get and maintain precise alignment. Irises are especially useful as you can periodically close the iris to check that it is closing centrally around the incoming beam, providing a very fast way to gauge if things have drifted. Critical areas are places where the alignment of the beam path is paramount to the correct functioning of the experiment. Typical examples of 'critical areas' would be where the beam enters an experimental target, OPA, or delay stage. Screwing irises directly into the tapped holes on the table ensures that beams are aligned to the bolt-holes.

- Making sure that the beam is (as far as possible) centered on every mirror, such that you can more easily use the full range of the adjustment screws to fine tune the alignment.

- Having a beam which is substantially smaller than the diameter of the optics (at least half the diameter), so that the beam is able to be moved around on the surface of a mirror to a greater extent before the mirror needs to be moved.

Some of these points will become clearer once we discuss how to actually align optics in the next section.

7.2.3 Sketching a Beamline

Bearing the above principles in mind, planning out the spatial arrangement of the optics in a beamline normally starts with a simple sketch, which is an ideal way to initially get a feel for how you will fit all of your optics in, and get an idea of the rough sort of size your setup will need. Often, doing a sketch like this reasonably accurate dimensions is entirely sufficient for making sure you have sufficient space, but sometimes (in more constrained areas) being more accurate can be beneficial. Most manufacturers of optomechanics will provide dimensions of their items, and will often also provide CAD files that can be loaded into a 3D drawing program. These programs may seem like overkill for this kind of optical design, but if time permits then this is the best way to get a realistic feel for how your planned beamline will look on the real optical table. The author has used Autodesk Inventor to design an entire lab space, starting with a vacuum instrument and then including an optical table and beamline with all associated optics, as shown in Figure 7.1. While this took a long time, it meant that construction of the instrument and beamline went a lot more smoothly than it would have had it not been planned so thoroughly. If time permits, this can be beneficial - but often a simple 2D sketch will suffice for most purposes[3]. It is also important to remember that plans are made to be deviated from, and the final beamline in Figure 7.1 ended up looking slightly different than the beamline as designed here. However, having a thoroughly planned place to start makes deviation easier, and minimises the chance of you encountering a totally insurmountable obstacle when building.

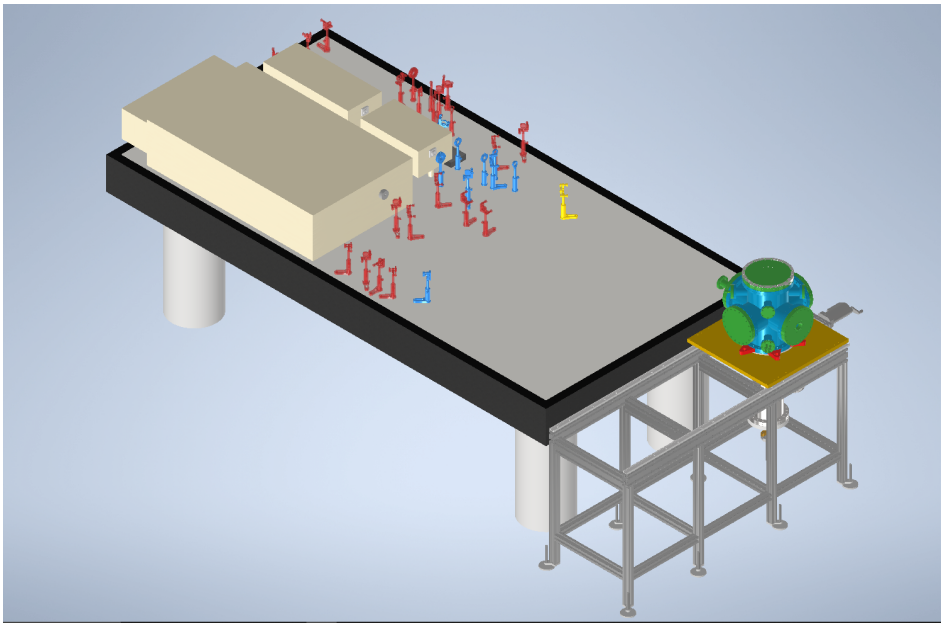


Figure 7.1: Initial CAD drawing for a lab space, including a multi-user beamline, designed by the author. Colours of the various optomechanics distinguish which part of the beamline they are intended to be used for.

[3]You could even use the Python library PyOpticalTable, if you wanted to also produce a nice figure for a report.

7.2.4 Budgetary Constraints

Often, you will be working to some kind of a budget (depending on what kind of research environment you are in!). Optics and optomechanics can be expensive, and so an essential part of planning is to shop around and find out how to build what you need in a cost effective way. There are many different suppliers of optical components (see Appendix C for a non-exhaustive starting point), and shopping around is well worth the time spent. Experienced optical campaigners will have their own opinions on the quality of various components from various suppliers - so asking around your department can also be well worth the time. I will not give many of my subjective opinions here, but taking the time to fully read the specifications of what you're buying and comparing between suppliers can be revealing!

When budgeting a beamline, it is natural to first think of the 'big' expenses. Things just as nonlinear crystals; diffraction gratings; finely controllable motorised translation stages or rotation mounts; spectrometers; and cameras are understandably going to be the biggest single items in the beamline budget. However, it is important to not underestimate the cost of simpler components that you may require large numbers of. For example, a simple dielectric mirror may cost only £100, but if you need to buy 25 of them for all the required steering (a not unrealistic number), then they quickly become expensive. Similarly, the cost of relatively simple optomechanics (such as posts and holders) can quickly add up. Thorough planning of the beamline will help you to predict these costs, so you are not taken by surprise at the end. As a ballpark figure, a beamline that the author has created for two-colour UV photochemistry cost around £16000 (in 2019), not including the cost of laser systems or OPAs. More complex systems could cost substantially more than this.

7.3 Optical Building

Now we come to the question of *how* to actually go about building and aligning an optical beamline. The best way to learn to become a proficient optical builder is simply through practice, but again there are general principles worth codifying, and common pitfalls worth avoiding.

7.3.1 Beam Steering

Laser beams travel in straight lines, and it is our job to move those straight lines so that they shine on the area we desire. We also know that we need to keep our beams travelling in a plane that is parallel to the surface of the table. If we don't do this, it is not only unsafe, but also a lot of optical elements function best when the beam hits them 'square'. That is, when the beam is collinear with a surface normal coming from the surface of the optical element. The most obvious example of this is a lens, and you can see this easily by looking through your glasses (if you wear them) at various angles and you will see that the lens behaves very oddly when you are not looking into the middle of it.

In addition to keeping the beams in the same plane, we often want to direct them along a specific vector within that plane. Mathematically, to define a vector (straight line) we need

two points, and so if we are able to align our beam such that it hits two points within our plane, then we have 'aligned' our beam to those points. These points could be irises placed at the entrance and exit of a sample cell, and it is our job to make sure we steer the beam accurately such that it hits both the entrance and exit centrally, and is therefore travelling straight[4]. To accomplish this, we use mirrors to steer our beam through both irises. In general, we will need two mirrors to steer the beam successfully through both irises. The reason for this is illustrated in Figure 7.2. A single mirror has to be perfectly placed in order to accomplish this job, and humans are generally not this accurate in placing optics. Having two mirrors allow the inevitable slight deviations to be compensated for.

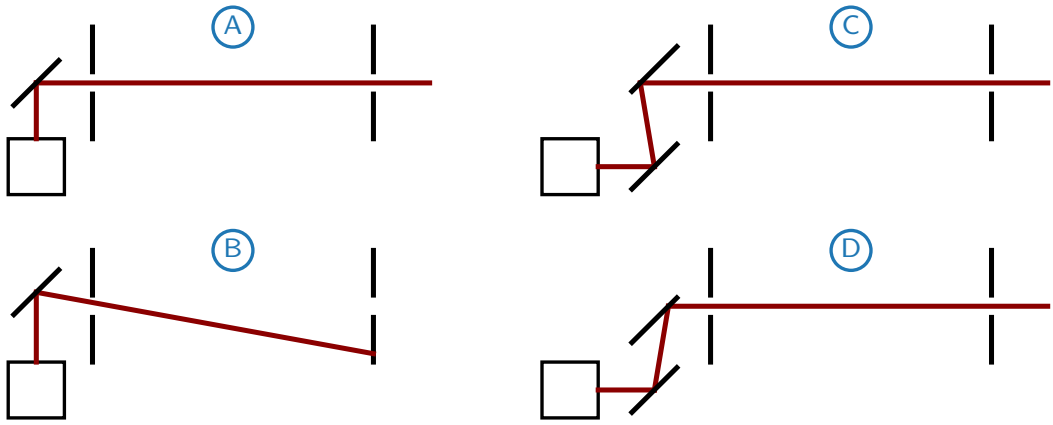


Figure 7.2: Illustration showing the utility of using two mirrors for steering a beam (red) from a laser (black box) through two irises (black). In case (A), the single mirror is aligned exactly with the iris, and no second mirror is necessary. However, slight deviation in the position of this mirror (B) means that coupling into both irises becomes impossible via rotation of this mirror. Using two mirrors, as in (C) and (D), allow slight deviations to be compensated for, as the position that the beam hits the second mirror can be adjusted using the first mirror, and then the angle of the second mirror can be finely adjusted to get the beam travelling through both irises.

Panels (C) and (D) in Figure 7.2 show how a beam can be aligned to pass through two irises using two mirrors. In practical terms, to get the beam path perfect, the alignment on the first iris (the near field) is optimised by adjusting the rearmost steering mirror, and then the alignment on the second iris (the far field) is optimised by adjusting the front steering mirror. However, moving the beam with the front mirror will slightly alter the alignment on the first iris, and so this process is repeated in an iterative fashion until the beam passes through both irises perfectly. This process is often referred to as doing a **beam walk**, and

[4]Many pieces of equipment, such as OPAs, autocorrelators, or spectrometers, contain an iris at both the entrance port and further inside so that the incoming beam can be reliably steered in straight

is an essential skill to have to be an effective optical builder. Note that the targets need not be irises, but could be marked alignment targets on beam blocks, experimental samples, or even marks on a piece of laser shielding. The fundamental process of using two mirrors to walk the beam and ensure it is travelling along the desired path is the same. Some good rules of thumb when doing a beam walk are:

- Placing the first iris as near to the final mirror as possible, and the second as far from the final mirror as possible, will minimise the number of iterations you need to do.
- Do not be afraid to 'overshoot' with one of the mirrors at first. This will normally mean you converge on the good alignment faster, and you will quickly develop a feel for how the system responds to your input.
- Start with the adjustment screws on each mirror mount set as neutrally as possible, and then get the alignment in the irises as close as you can to perfect by just moving and rotating the mirrors without touching the screws. This will make the final adjustment using the screws quicker, and will ensure that only small adjustments are needed to get the alignment perfect. Running out of screw thread when you are close to perfect alignment is a frustrating experience!

We can also use irises to create targets around beams that we already know are aligned well. A typical example of this might be if you have painstakingly aligned a beam such that it passes through and hits your experimental sample exactly in the right position. Once you have the beam aligned well like this, it is normally prescient to place irises centered around the beam, so that you can easily get the beam aligned back on to your sample simply by following the beam walk procedure described above.

Transmissive Optics

When placing a transmissive optic such as a lens into a laser beam, it is normally important that the incoming beam is collinear with a surface normal from the optic. To ensure this, the optic needs to be placed such that it is totally 'flat' with respect to the incoming beam direction. This is difficult to achieve simply by eye, but a simple trick that can help is to look for the **back reflection** from the optic. This is the reflection coming from the face of the optic that the beam hits first, and if the optic is correctly placed, then this back reflection travels back along the incoming beam direction. Whilst this reflection is normally very weak (especially if using an AR coated optic), it can normally be seen faintly if the room lights are switched off, or by using an IR viewer if you have an IR beam. Rotating the optic in its mount will move this reflection, and so it can be exactly lined up with the incoming beam - ensuring that the optic is perfectly placed.

Safe Optical Building

Building a beamline necessarily requires that there are exposed and unshielded laser beams flying around the table, at least temporarily. Some simple safety precautions can minimise any risk, described earlier, and also make the process a lot less nerve-wracking. Simple

things like attenuating the beam as much as possible before you start to align it, never letting a beam fly into empty space without knowing where it is going, and always taking care to catch unwanted reflections or parts of the beam that are dumped will help to keep you and your co-workers safe. It is important to remember that every glass surface will create a reflection[5], and this includes the front and backside of optics. Make sure that you account for all of these reflections. Another area of potential danger is when using an optic with a clear or polished backside. Beams will fly straight through the clear backsides, and this can create a major safety issue if used in periscopes where vertical beams will occur. Gaffa taping some cardboard to the back of the mirror mount to cover up the backside of the mirror is an easy, quick, cheap, and safe solution.

Another aspect of safety is keeping your expensive optics safe. Very intense ultrashort pulse laser beams can easily damage optics. **Keeping optics clean** is the best way to minimise damage. Most of the time, a burned optic gets burned because some dirt on the surface absorbs laser light and scorches the surface. Keeping dust off of optical setups by covering them up is a good first step, and periodically checking to make sure no optics are damaged is a good habit to get into. If you do find a dirty optic, the first thing to do is to blow some air on it from a clean source to try and remove the dirt. A good source of clean air is an empty and dry solvent squeeze-bottle, as in my experience canisters of 'clean' compressed air still contain trace amounts of things you don't want on your optics. If blowing air onto the optic did not remove the dirt, then cleaning it carefully with pure methanol and lens tissue should remove most of the dirt. However, be certain that the optic and its coating can tolerate methanol (or whichever solvent is used) before you do this - some optics cannot be cleaned in this way (this is especially common with things like infrared polarisers, as I once discovered the hard way).

The other main cause of damaged optics in ultrafast systems is the laser pulses being intense enough to cause nonlinear effects like self-focussing. These can lead to the creation of very high local intensities which damage even clean optics. Normally you will see an optic start to glow, or see a bright spot in it, before it is irreparably damaged, so keeping an eye on the beamline when you first allow a high power beam to enter it is a good idea. Self-focussing can be particularly pernicious, as you may inadvertently end up focussing your beam to a point somewhere you're not looking. In the worst case, this could drill a hole in a vacuum window - so it is worth being aware if you are using high power beams.

A final point to make here is to **never allow a beam to hit an optic that is not bolted securely to the table.** Aside from being very annoying when you knock something over, it is potentially very dangerous if a mirror is knocked and reflects a beam across the lab. **Always bolt your optics down**[6]**.**

[5] If you see two intense reflections from a mirror, then it is likely that you have it in back-to-front, and the uncoated backside now is at the front and reflects a lot of light, while the HR coated frontside is at the back and also reflects a lot of light at a slightly different angle.

[6] This might seem patronisingly obvious, but I once worked with a theoretician trying to gain practical experience who couldn't see the point in bolting things to the table. 'Just be careful and don't knock it once you've placed it'. As far as I know, he is still a theoretician.

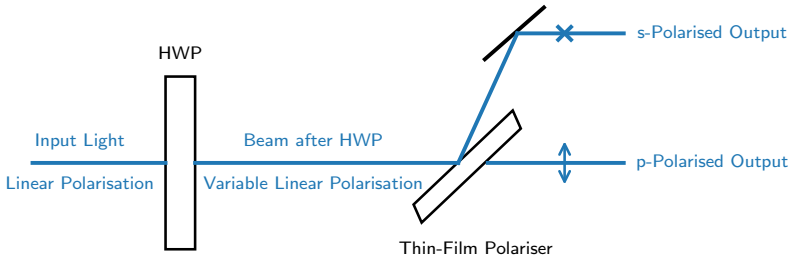


Figure 7.3: A waveplate-polariser variable attenuator. The polariser needs to be placed at Brewster's angle.

7.4 Common Sub-Assemblies

Finally in this chapter, we will walk through the construction of some common 'sub-assemblies', or things that you'll make often in a lab. There's not a substitute for actually being shown this in person, but hopefully the below can give you some useful guidance.

7.4.1 Variable Attenuator

An example of a waveplate-polariser variable attenuator was shown in Example 6.8, and the schematic reproduced in Figure 7.3 for clarity. The key points around the construction are:

- Place the polariser at Brewster's angle (the angle where s-polarised light is perfectly reflected).
- Place the waveplate so it's face is exactly orthogonal to the input beam.

To achieve this in practice, you will need the waveplate (HWP) and thin-film polariser (TFP) shown, and some beam blocks. We will assume that the beam is initially linearly polarised. The HWP needs to be in an indexed rotation mount. The steps I would follow are:

1. Reduce the power in the input beam as far as possible.
2. Place a block in the beam path, and then place the TFP after the block, and place another block behind the TFP.
3. Remove the front block, and allow light to hit the TFP. Move the TFP so the light hits it centrally. Rotate the TFP and see how the transmission varies. If the light was originally p-polarised, you should see a faint reflection and intense transmission, and vice versa if it was originally s-polarised.

4. To find Brewster's angle, rotate the TFP until the reflected (if initially s-polarised) or transmitted (if initially p-polarised) beam is minimised. You may need to dim the lights or use an IR viewer to see the exact point where the amount of light is minimised.

5. Once this is found, bolt the TFP in place. Block the beam infront of the TFP. Place the HWP behind the block, roughly in the place where the beam will hit it

6. Allow light to hit the HWP, move the HWP so the light hits it centrally.

7. Rotate the HWP around the axis of the optical post holding it so it is not orthogonal to the direction of the incoming beam. Look for the back-reflection from the HWP, and then rotate the HWP such that the back-reflection travels back along the original input direction. Lock it in place when this is found.

8. Rotate the HWP around the axis of the incoming beam to observe the power in the two arms of the polariser vary. If you are not using both, bolt a beam block behind the arm you are not using.

Once you have confirmed it is all working as intended, you can increase the input power back to the desired level. Of course, this may be something you have placed exactly at the laser output, and so you don't have an easy power control. In this case, there are some other options (depending on the laser you have):

- Some lasers have integrated power control: easy.

- Some lasers have integrated rep-rate control: you can reduce the rep rate to reduce the overall average power (but not the pulse energy - be aware of nonlinear effects).

- Some lasers allow you to change the timing of the Pockels' cells in the regenerative amplifier. I would not recommend playing with this if you are a complete novice, but you can alter the timing to release the pulse from the regen cavity before it has been amplified to it's final peak power.

7.4.2 Periscopes

Imagine the output port from your laser is 160 mm above the table surface, but that you need to eventually couple the light into a spectrometer that has an input slit only 100 mm above the table surface. In this case, we need to use a **periscope** to lower the height of the beam. You can buy periscope kits from many manufacturers which are usually worth the cost – they take out some degrees of freedom that can make alignment a headache.

Firstly, a periscope necessitates that a beam is sent vertically: **this can be dangerous, as a beam can easily reflect upwards into your eye.** Periscopes that lower the beam (as here) are less dangerous than those that raise it. In all cases, extreme care is advised and after construction the periscope should be entirely shielded, especially at the top. **I would**

strongly recommend that if you are a novice optical builder, you ask someone for help when building your first periscope.

Fundamentally a periscope is just two 45° mirrors, as shown in Figure 7.4. There are a couple of important considerations:

- The periscope drawn in Figure 7.4 has both input and output beams travelling in the plane of the page. This will preserve the polarisation state of the input beam. A *turning periscope* where one mirror is rotated so that the beam travels vertically, and then comes out at 90° to the input direction will also change the polarisation state of the beam (a good exercise for you is to work out why!). It is important to be aware of this if your application is sensitive to polarisation states[7].

- The periscope can also be aligned so that it reflects the vertically translated beam back along the input direction – useful in some situations.

- Using irises around the input and output beams to ensure they travel in a straight line is useful.

To build a periscope, you need to start by ensuring that you know the input beam is travelling along the direction you want it to – ideally the periscope will not change the direction in which the beam points, just the height. Having some markers on the table that show you the direction that the vertically translated beam should end up going in will be helpful. I would then proceed as follows, for our downward-going periscope described previously:

1. Block the beam, and place the first mirror. Ideally you are using a periscope kit so this is relatively easy. *Be exceptionally careful if you are building an upward-going periscope, as a beam will now come vertically up from the table! You need to place a block above the table to catch the beam before the next step.*

2. Ensure, as far as is safely possible, that the beam after hitting this mirror is travelling at 90° up or down from the input beam. This is much easier and safer to do for a downward-going periscope.

3. Block the beam, and place the second mirror.

4. To check that the beam is travelling straight:

    - Use a marked block to gauge the height of the beam. Check it at two locations: near the periscope, and far from the periscope. Adjust the first mirror to align it near, and the second to align it far.

    - Use another marked block or iris to gauge the horizontal position. Check it again at two locations, and adjust as above.

[7] A misaligned turning periscope can do all kinds of wacky stuff to your polarisation...

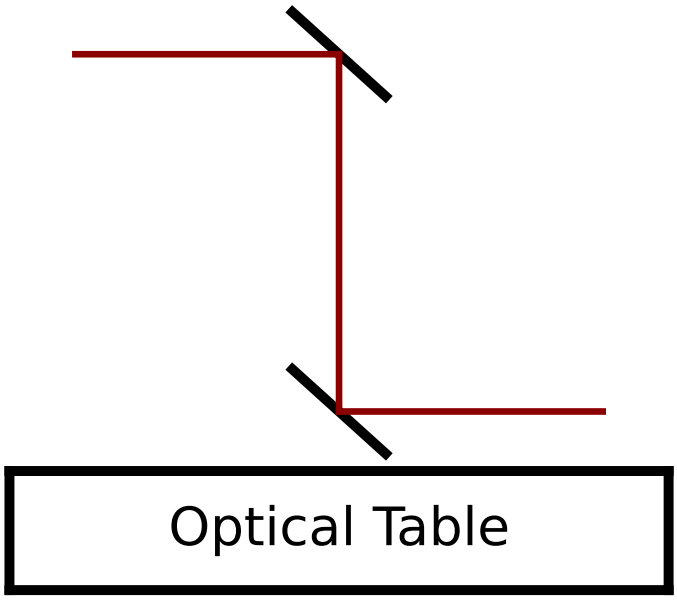


Figure 7.4: Simple periscope. The beam always travels within the plane of the page to preserve the polarisation state of the incoming beam. There is no reason why the bottom mirror could not be rotated to direct the beam back in the direction it originally came in, but lower.

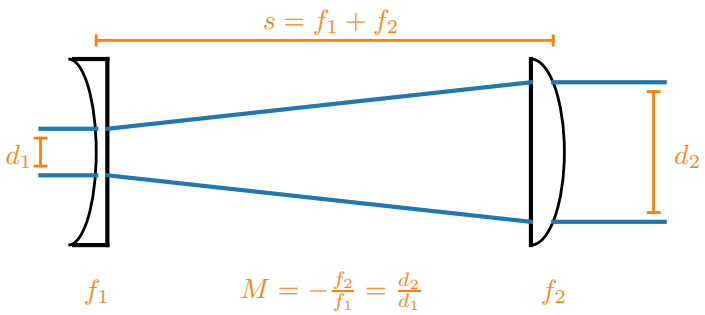


Figure 7.5: Simple Galilean telescope. From left to right: a beam expander. From right to left: a beam reducer.

5. Several iterations may be required before satisfactory alignment is reached.

Once the periscope is built, it is essential that it is properly shielded! I cannot stress this enough – periscopes are normally one of the scariest things to build in a lab.

7.4.3 Telescopes

We'll discuss how to build a simple transmissive Galilean telescope – as shown in Figure 6.7, and reproduced below. These telescopes consist of two lenses.

Constructing a telescope can be a bit fiddly, but isn't normally as terrifying as building a periscope. Mounting one lens on an inexpensive manual translation stage to allow the distance s to be adjusted so that the collimation can be fine-tuned is a good idea. As is using **collars** on the lens posts so that they can be easily rotated without changing their height.

When assembling a telescope, I would normally do the following:

1. Put the converging lens in first. This is generally easiest, as you will be able to keep track of where the beam is more easily than if you put the diverging lens in first[8], but take care to block the beam after the lens before it focusses[9]

2. Adjust the position of the converging lens so that the beam hits it centrally (in the horizontal and vertical directions).

3. Rotate the converging lens and look for the back reflection. Align the back reflection

[8] Although my colleague disagrees on the basis that it is more dangerous to have a converging beam than a diverging one. I live life on the edge, but the risk is minimal if you put a block right after the lens.

[9] If you hear a sizzling sound, see smoke, or smell burning: you need to put the block nearer the lens.

so that it travels back down the incoming beam path as far as possible.

4. Place the diverging lens, at a distance s from the converging lens. Again, adjust the position and rotation so that the beam hits it centrally and the back reflection travels back down the incoming beam path.

5. To check the collimation, look at the beam near the telescope (a card with measurements marked on it is useful here), and then far away. Adjust the manual translation stage to alter the collimation until it is collimated. If it never seems to collimate, you may need to move one of the lenses (perhaps the incoming beam was not collimated).

Ideally the direction of the beam pointing does not change on insertion of the telescope. Irises placed in the original beam path (before a telescope was added) can help confirm this.

7.4.4 Delay Stages

Next, we will talk about how to align a **delay stage**. A delay stage is an assembly that allows the path length of a beam to be changed without changing the pointing direction of the beam. Fundamentally, this consists of two mirrors arranged at 90° to each other such that an incoming beam is reflected back such that the output beam is parallel to the input. The mirrors can then be mounted on a translation stage (electronic or manual) and the position of the two mirrors can be scanned along the direction of the input beam. The amount of delay added to the beam is therefore double the distance by which the stage is moved[10]. A schematic of a simple delay stage is shown in Figure 7.6.

Alignment of delay stages can be fiddly, but ensuring that the two mirrors are placed exactly at 90° to each other simplifies this process enormously. You can do this by hand with careful alignment, but I'd strongly recommend buying a dedicated mirror mount that is precisely milled to hold two mirrors at this angle. Or, even better, buying a **hollow retroreflector**, which is a single optical element that contains three reflective surfaces at precisely 90° to each other – and is insensitive to the incident angle.

Assuming you have two mirrors at this 90° angle, it is then just a case of ensuring that the input beam is coming in exactly straight[11]. This is shown with the solid red line in Figure 7.6, and when this is satisfied, the mirrors can be translated on the stage without moving the pointing direction of the output beam. If the input beam is coming in at an angle, shown with the dashed line in Figure 7.6 for two cases, then the output beam will also be angled, and moving the stage will change the position of the beam when viewed at a distance.

To ensure the stage is aligned, I would do the following:

[10] In this example – you can also have a *quadruple pass*, or *octuple pass* stage where you send the light on more trips around the same mirrors, increasing the delay added.

[11] If you have a hollow retroreflector, even this doesn't matter as the output beam is always parallel to the input.

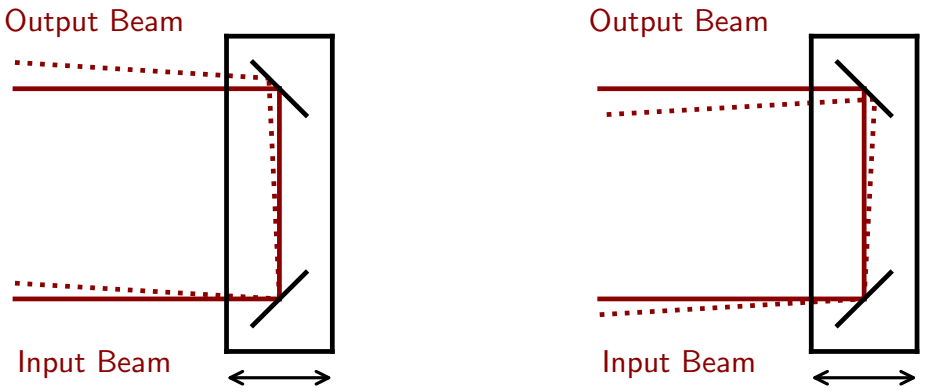


Figure 7.6: Schematic of a delay stage. The mirrors in the black box are movable in the direction indicated by the double-headed arrow. The solid red line shows a well-aligned input beam. The dashed red line shows a poorly-aligned input beam - in this case the pointing direction of the beam will not stay the same as the stage is moved.

1. Ensure the beam coming in is exactly straight. Use irises to confirm this.
2. Look at the beam a long way away from the stage – ideally project it onto a wall if you can (use low power!).
3. Move the stage, if the beam on the wall walks around, your stage is not perfectly aligned. The direction by which it walks tells you about the misalignment:
   - If the spot walks towards the center of the stage as the stage is moved back, then the angle of incidence of the input beam is greater than 45° and needs to be reduced[12].
   - If the spot walks away from the center of the stage as it is moved back, then the angle of incidence of the input is smaller than 45° and needs to be increased.

   Convincing yourself of this is a useful exercise in geometry, and your ability to sketch things. It helps to have a friend who can adjust the mirrors while you stare at the spot on a wall for this step.
4. Correct the input coupling using two mirrors as usual.

You can get the stage pretty well aligned in this way. If you need to focus very small beams at your experiment you may have to repeat this process whilst looking at the beams on a profiling camera or pinhole, or by looking at an experimental signal.

[12]Remember these angles are defined relative to the surface normal of the mirror – Figure 6.2.

7.4.5 Overlapping Beams

The final thing we will discuss in terms of optical construction is how to overlap two beams. When we discuss overlapping beams, there are two things we can mean: the **spatial overlap**, which is ensuring that the beams overlap at the same point in space; and the **temporal overlap**, which is ensuring that they arrive at the same point in time. With ultrafast lasers, finding the temporal overlap can be particularly challenging as to do this we need to make two pulses arrive at the same time as each other to within a few femtoseconds. We will address each of these problems in turn. We assume that both beams originate from the same driving laser - otherwise finding temporal overlap is impossible.

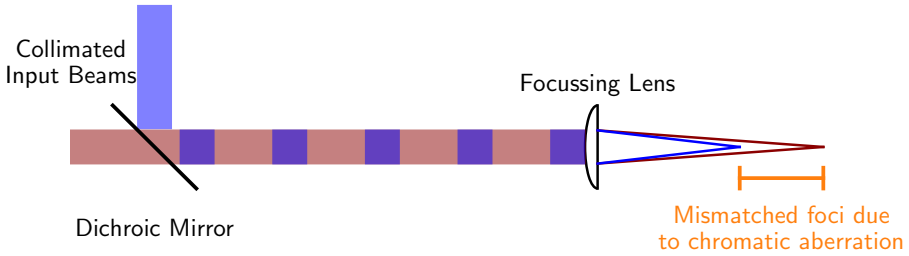
Spatial Overlap

How difficult it is to spatially overlap two beams depends mostly on how big the overlapped beam waists are (bigger is easier), and on whether the two beams are initially collinear or not. If you are using large, unfocussed beams, then overlapping two beams in space can be done by eye, simply by holding a piece of card where you want the overlap and steering both beams on top of each other. You may require some fluorescent paper, or IR paper, to see the beams if you are using pulses that are not visible immediately by eye. Overlapping two focussed beams is more challenging, however – especially if the beams are of different colours. For the purposes of the following, we will imagine we are trying to overlap an IR beam with a UV beam.

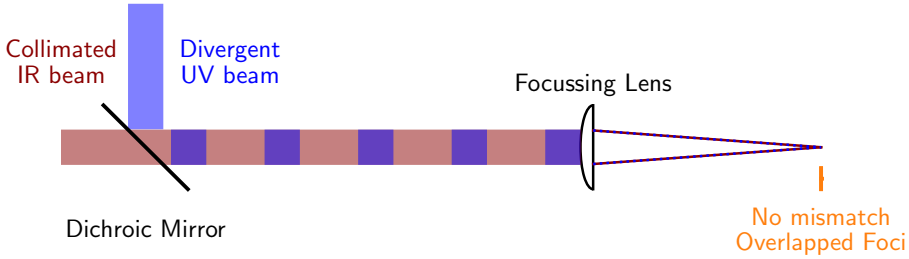
We consider initially the case where the two beams are collinear. To make these beams collinear requires that they are recombined on a dichroic mirror. Figure 7.7 shows how this is achieved. The IR beam is sent through the back of a mirror where the front side is coated to reflect the UV beam. Special mirrors to do this are available to buy and are often marketed as 'harmonic separators'. Once the beams are collinear, they are then focussed with a lens, and we need to find a way to ensure that the foci both overlap. The immediate question is then to ask how we can actually see *where* the foci are - and we normally do this using something like beam profiling camera, or photodiode and pinhole, as discussed in the context of beam characterisation in chapter 5. It is then a case of directing the focussing beams onto the camera/diode (with attenuation!), then finding one beam, marking it's position, and then steering the other beam on top of it without moving the first beam.

This sounds straightforward, but there is one further complication when focussing beams of different colours using a lens: the chromatic aberration of the lens means that the different colours will be focussed to slightly different points along the propagation direction of the laser beams, as illustrated in Figure 7.7 (a). To compensate for this, we need to adjust the divergence of one of the beams such that the focus of that beam overlaps with the other (recalling that lens focal lengths are quoted for collimated input beams, and so a diverging input beam will have a longer focal length). This is normally controlled by adjusting a telescope placed in one of the beams[13], as illustrated in Figure 7.7 (b). If the two beams are not collinear, then the process is much the same, except that it can be a **vastly** harder

[13]The problem could also be entirely circumvented by using an **achromat**, but these are relatively thick and so not ideal for ultrafast use.



(a) Collimated input beams.



(b) Collimated IR beam, divergent UV beam.

Figure 7.7: Spatial overlap of two different coloured beams using a singlet lens. (a) When the beams are both collimated, chromatic aberration causes a mismatch in the positions of the foci. (b) Making the UV beam slightly divergent (using a telescope) can compensate for the chromatic aberration, and allow the foci to be overlapped in space.

to find the overlap position, as collinear beams are already reasonably well overlapped in space by definition. Often in this case, overlapping the beams **using your experimental signal** is the best approach. If you have a clear experimental response from when the beams are overlapped in space, then using this to optimise your overlap is probably the easiest way to go about things, but it is not always the case, or the signal is too small to be easily usable in this way.

Temporal Overlap

Overlapping two femtosecond laser pulses in time can be quite challenging. Light will travel about 10 µm in 35 fs, and so you need to be able to adjust the lengths of each beam path with micrometer precision. This will require that at least one of your beams travels on a delay stage controlled with a micrometer screw, so you can accurately move your beams

in such small increments. The next thing to do is to get a rough estimate for how far apart your pulses will be in time. This can be done either by using an oscilloscope and photodiode, or by simply measuring both beam paths with string and a tape measure. Both of these methods are accurate to within around a nanosecond or so - and a useful rule of thumb is that light travels around 30 cm in a nanosecond[14]. This will tell you if you need to add vast amounts of delay to one of the beams - as may be the case if one has passed through an OPA and one has not. Having to add several metres of optical delay to a line to synchronise it with a line that has gone through an OPA is not uncommon. Getting them synchronised to within 30 cm is about as far as electronic assistance can take us, and now we need to use an optical method to determine if the pulses are temporally overlapped.

We already know from chapter 5 that nonlinear crystals can be used to measure ultrashort pulse durations, and when the pulses are overlapped in the crystal both spatially and temporally a nonlinear interaction occurs which produces a new colour of light we can detect. This is one of the most common ways to find the point in time where two ultrashort pulses are overlapped - this point in time is colloquially known as $\mathbf{t_0}$ ('t-zero'). What crystal we need depends on what nonlinear interaction we are looking for, but for illustrative purposes we will assume we are looking for the temporal overlap between an 800 nm IR beam and a 266 nm UV beam.

The first task is to find out what non-linear interaction we will need to look for. In the example here, we can do a difference-frequency mixing of 800 nm and 266 nm to produce 400 nm light. This requires that we have a nonlinear crystal that can phase match this process, and a useful tool to find out what we can use is the Light Conversion Optics Toolbox, as described in detail in chapter 3. In this case, we find that a BBO crystal cut at 44.3° is what we need. The beams need to be spatially overlapped in this crystal, and a simple way to do this (provided that your beams are intense enough) is simply to place the crystal in the beamline before the focussing lens, where the two beams are already collinear. This is illustrated in Figure 7.8.

Having placed the crystal here, we should be able to make 400 nm light when the pulses are overlapped in time. To be able to actually see this light, it is useful to disperse the collinear output using a prism, and then looking at the different colours on a white card (or similar). Then by scanning the delay stage, you should be able to see a very brief flash of blue 400 nm light when the pulses are overlapped in time. When you see this, record the position of the delay stage, as this is your t_0! Note however, that the t_0 you record here may not correspond exactly to the t_0 in your experiment, as things like vacuum windows and focussing lenses will shift the temporal overlap somewhat (normally by a millimetre or less). You can add a 'fake window' before your crystal to account for this if desired, or place a mirror between the focussing lens and window (if space permits), and send the reflected beams into a crystal to account for the effect of the lens as well.

The process described above sounds straightforward, but in reality trying to find t_0 can be a very frustrating experience if you are not lucky enough to find it in the first few attempts.

[14] Or 'a foot per nanosecond' if you live in the USA or Brexit Britain.

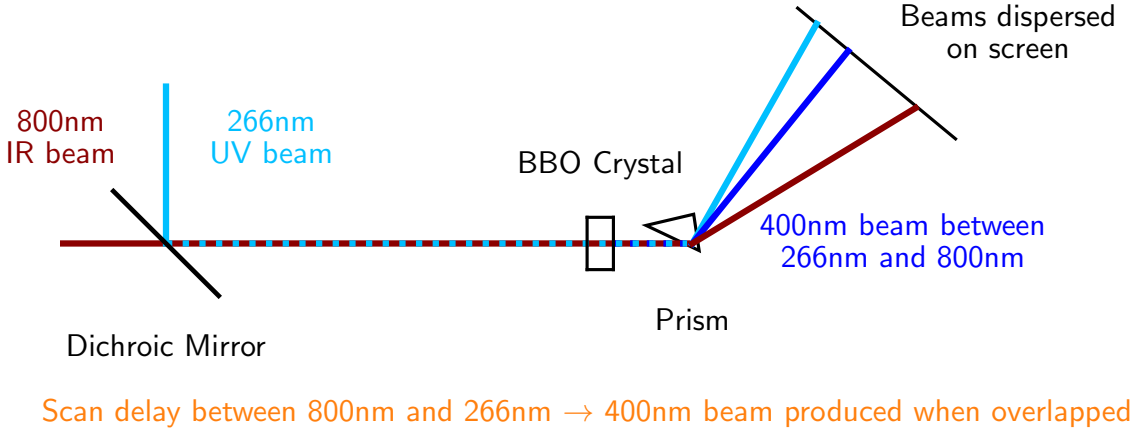


Figure 7.8: Illustration of a setup used to find temporal overlap between an 800 nm IR beam and 266 nm UV beam. The beams are spatially overlapped in a BBO crystal, and the beams are then dispersed onto a screen using a prism. The delay between the IR and UV pulses is scanned, and when the beams are overlapped temporally, a 400 nm beam will appear on the screen.

If you do not see it, then there are a few things you can try to do:

- **Optimising the crystal angle**: As the phase-matching is exquisitely sensitive to the angle of the nonlinear crystal. Iterating a process of adjusting the angle slightly, then scanning the delay stage, may help you to find the overlap if the crystal angle is not perfect when it is placed flat in the beam path.
- **Scanning more slowly**: While watching a piece of card can be tedious, the temptation to scan the stage very quickly can be problematic. Remember that you may only have a few tens of microns of stage movement where you will see the t_0 light produced.
- **Increasing the intensity**: Nonlinear processes depend heavily on having intense light. Sometimes unfocussed beams will not have sufficient intensity to drive the nonlinear process, so focussing the beams into the crystal can help. Take care not to drill a hole in the crystal!
- **Checking Polarisations**: Ensuring that you have the right combination of polarisations for your desired phase-matching scheme is important.
- **Optimising the spatial overlap**: Especially if you are focussing the beams, it is critical that the beams are well overlapped within the crystal.
- **Using a photodiode**: A photodiode is more sensitive than your eyes, so using a

diode with a filter attached to it that will only let through the light from your non-linear process can help you see the t_0 if it is very faint. A spectrometer can also be useful here.

- **Using experimental signal**: If all else fails, and you already have some experimental signal you can use, then looking for the temporal overlap within the experimental signal can be a good option. Especially if the experimental signal behaves more like a step function, which turns 'on' when the beams overlap and then stays at a high level after the overlap for some time. This can make it much easier to find the overlap, as you are not looking for a very transient flash of light.

Once you have some faint t_0 signal, then often you can make it much more intense by optimising things like the overlap of the beams and the angle of the crystal. The good news is that once you have it, it will be easy to find it again, and you are well on the way to having a fully-functioning beamline. Finding t_0 is the sort of event that generally merits a celebratory coffee-and-cake break.

To tie together all of the previous chapters, we will now complete our story in the next couple of chapters with two case studies of different beamlines: one for gas-phase UV photochemistry, another for transient absorption spectroscopy.

Bibliography

Chapter 8

Case Studies

In this final chapter we will tie together the practical knowledge discussed in the preceding chapters by walking through the design of a couple of beamlines from start to finish. Firstly, we'll look at a beamline for pump-probe gas-phase photochemistry, and then a beamline for solution phase transient absorption spectroscopy. In each case, we will focus on slightly different aspects:

- In the gas-phase photochemistry case, we will focus more on ensuring our pulses end up short, and walk through that kind of calculation..
- In the transient absorption case, we will focus more on practical tips for building.

Hopefully, between these two examples, you'll start to get the hang of how you can start to do this yourself. There is no substitute for getting in the lab and doing it, though.

8.1 Gas-Phase Photochemistry

8.1.1 Initial Equipment

As in many projects, there was not a blank sheet of paper (or blank cheque) available, and the beamline had to fit within existing constraints. Principally, these consisted of the driving laser system and OPA, which already existed prior the beamline design. The optical tables were reasonably empty, so there were no huge spatial constraints. The budget was relatively tight as it was a new research group starting from scratch, so many other kinds of equipment also had to be bought.

The beamline is driven by a Spectra-Physics Solstice Ti:Sa laser system. This produces 800 nm, 35 fs, 5 mJ pulses at a repetition rate of 1 kHz. 3.4 mJ of this output is steered into a TOPAS Prime OPA, which is coupled to a NirUVis mixer extension box. This OPA and mixer give a tunable output from 190 nm to 2600 nm, whilst maintaining approximately the pulse duration of the driving laser. The pulse energy of this beam varies substantially

with wavelength, but in the UV region we are interested in, is always greater than around 10 µJ. The height of the output of the Ti:Sa laser is around 160 mm from the table surface, but the height of the OPA and mixer output is 130 mm from the table surface. The beam from the Ti:Sa output is around 11 mm in diameter, but the beam from the OPA is substantially smaller, around 5 mm diameter.

8.1.2 Requirements

This beamline is intended for use in a gas-phase UV-IR pump-probe spectroscopy experiment. The UV pulses will be used to initiate ('pump') some photochemistry in an isolated molecular target in vacuum, and then the IR pulses will Coulomb explode ('probe') the target. Coulomb explosion is a process where a molecule is subjected to a very intense laser field and undergoes rapid multiple ionisation, which causes the molecule to 'explode' into ion fragments almost instantaneously. Knowing these requirements, we can create a list of the laser conditions we need to have in our vacuum chamber.

- **Beams**: We need to overlap two beams in space and time in our vacuum chamber, and delay them relative to each other with femtosecond time precision.
- **Position**: The target is a molecular beam inside the vacuum chamber, and the window to the chamber is around 160 mm higher than the surface of the optical table. The molecular beam is skimmed to 1 mm width.
- **Colour**: The IR pulses can be used at 800 nm, but the UV pulses need to be tunable between around 200 nm and 400 nm.
- **Size**: The IR pulses to induce Coulomb explosion need to be fairly intense, so a small 25 µm beam waist is desirable. The UV pulses can be somewhat larger, as the same intensity is not required, and having a larger beam waist means that spatial overlap is easier, and we ensure that we are only probing molecules that are actually pumped. Around 50 µm is more than sufficient for the UV beam.
- **Pulse Duration**: Measurement of sub-picosecond dynamics is desirable, so the pulses should not be longer than around 100 fs. Adding additional compression may also stretch the budget somewhat, however.
- **Intensity**: Around $1 \times 10^{14}\,\mathrm{W\,cm^{-2}}$ is required for Coulomb explosion. Several orders of magnitude less is required for the UV pump, around $1 \times 10^{10}\,\mathrm{W\,cm^{-2}}$. This translates into around 100 µJ of pulse energy for the IR pulses, and less than 1 µJ for the UV pulses.
- **Polarisation**: Variable linear polarisation of both beams is required.

We now compare these requirements to what our driving laser can provide. Working down the above list, we find the following:

- **Beams**: We will need to add a motorised delay stage to either the pump or the probe arm, to allow the temporal overlap to be found.
- **Position**: We will need to raise the UV beam by 30 mm up to a height of 160 mm.
- **Colour**: The OPA provides all needed tunability.
- **Size**: We will need to focus the beams into the vacuum chamber. The distance between the edge of the laser table and the centre of the molecular beam is around 20 cm, so using a 300 mm focal length lens for focussing would be ideal[1]. Using this lens with our direct 11 mm IR output beam would produce a beam waist of around 12.5 µm, so we would like to *shrink* the IR beam by a factor of two to increase the beam waist. The UV beam may also be too small, but adding a UV telescope may make the pulses unacceptably long, so we will see how it looks once we have beams in the chamber, as the OPA output may be quite divergent anyway.
- **Pulse Duration**: We should plan the beamline and then calculate the GDD added by all of the transmissive optics, and the air travelled through.
- **Intensity**: These pulse energies are easily obtainable using our laser system.
- **Polarisation**: Adding half-wave plates to both beam lines will be necessary.

This beamline will also be used to drive two other experiments with more or less identical requirements to this (though not simultaneously), so we should also keep this in mind when planning. It was agreed with the other groups using the beamline that we would just provide a place where beams could be easily 'picked off' and sent to another experiment, but that they would build their own beamline. Areas where a beam has been picked off and sent to another experiment are shown as dashed red and blue lines that trail off into space on Figure 8.1. We now know what requirements we have to keep in mind whilst planning the rest of our beamline.

Constraints

As shown in Figure 8.1, the vacuum chamber is not directly adjacent to the output of the laser system, and there is around 3 m of optical table in between the laser output and vacuum window. We also need to be able to provide beams to drive other experiments, so must account for this. Budgetary constraints are also such that expensive items (like the motorised delay stage) should ideally be able to be used by multiple experiments. Adding additional pulse compression stages may also be difficult, so we should keep an eye on the accumulated GDD.

[1] So the lens is not clamped right on the edge of the table.

8.1.3 Design and Construction

We will now walk through Figure 8.1, discussing how all the requirements will be met and any special considerations that need to be made. We will make reference to the orange circled numbers as key reference points throughout our discussion.

Starting at point (1), with the direct Ti:Sa output. This output needs to be split up such that it can drive the OPA, and also so that a small fraction can be sent into another experiment (the dashed line after the first beamsplitter). To accomplish this, two beamsplitters are placed at the direct output. Beamsplitter BS1 is a 90:10 beamsplitter, and so reflects 90% of the incident light whilst transmitting 10%. This 10% (500 µJ) is sent to another experiment. The remaining 4.5 mJ is directed onto BS2, which is an 80:20 beamsplitter. The reflected 80% (3.6 mJ) is used to drive the OPA, being reflected off two other 1.5 inch mirrors to steer it into the OPA. The transmitted 900 µJ will become the probe beam for our experiment, and for one other experiment. Both of these beamsplitters are 1.5 inch beamsplitters provided by Spectra-Physics with the laser. This beam is sent on a circuitous route to delay it relative to the UV beam, as the UV beam is delayed by around two metres having passed through the OPA. All the mirrors used in the 800 nm IR line are low-GDD broadband dielectric mirrors designed specifically for ultrafast pulses of this wavelength, and are all 1 inch in diameter.

We said that for our experiment that we wanted to shrink the size of the IR beam down, and so need to use a telescope. We also know that generally it is good to shrink the beam earlier rather than later, so it is not difficult to align and any divergence can be controlled. In this case, however, the two experiments being driven by this beam have quite different requirements for the beam size, so we delay shrinking the beam in a telescope until after we have split this beam further. This further splitting happens at point (2), with the typical half-wave plate/polariser combination, which also acts as a variable attenuator. The beam that is transmitted through this polariser will become our probe beam, whereas the reflected beam will become the probe beam for another experiment. As both experiments will not run simultaneously, this can also be used as a variable attenuator for both experiments without affecting the others experiment. The attenuator consists of a zero-order half-wave plate, and a broadband thin film polariser. The polariser is 2 inches in diameter, as it needs to be placed at Brewster's angle, so the beam is spread out over the surface of it, and would clip on a 1 inch polariser. After this attenuator, we have the first set of irises that can be used to aid alignment.

The IR probe beam then reflects off a zero-degree mirror and passes through a telescope. This telescope was chosen to shrink the beam down by a factor of two, so consists of a converging lens with +200 mm focal length at the input, and a diverging lens with -100 mm focal length at the output. Both of these lenses are AR coated for 800 nm to minimise loss. These lenses are placed 100 mm apart (the sum of the focal lengths), and the beam is shrunk by a factor of two (the ratio of the focal lengths). The beam is then sent steered through the back of the dichroic mirror DM, and into the vacuum chamber. The dichroic mirror DM has the front face coated to approximately reflect 250 nm to 270 nm, and the back face AR coated at 800 nm to minimise transmission losses. Note that two mirrors are

used to steer this into the chamber after the telescope, allowing easy alignment on the two irises placed before the lens into the vacuum chamber window at (5).

The UV beam is produced at the point (3) from the OPA and mixer, and is immediately raised using a periscope up to 160 mm, such that all the beams are propagating in a plane 160 mm off the table, at the same height as the vacuum window. It was chosen to do this rather than send the beam down to the table (for increased stability) and then bringing it back up, as avoiding periscopes was desirable, and the entire table is shielded from external air currents so stability was not a major concern. As the UV beam is tunable, it is important that mirrors reflect a broad range of UV bandwidth, so that they don't have to be replaced frequently when moving the UV wavelength around. For this purpose, protected aluminium mirrors were judged to be the best solution, as they have a high reflectivity down to around 250 nm, and reflect a broad bandwidth. All mirrors not specifically labelled in the UV beam line are protected aluminium.

The UV beam then is directed through some irises onto the delay stage at point (4). Irises are placed before the delay stage, as it is imperative that the light coupling into the stage is travelling exactly straight. For this reason the irises were screwed straight into the bolt holes of the laser table, so that the light was always travelling directly down the bolt holes. Substantial care was then taken to ensure that the angle between the two mirrors on the stage was exactly 90°, so that the beam coming off the stage was also travelling exactly straight. This is imperative, as if it is not, then the pointing direction of the UV beam will change whenever the stage is moved - spoiling the spatial overlap in the vacuum chamber. The best way to check this is to look at the beam coming off the stage as far away from the stage as is reasonable, and ensure that the beam does not move at all when the stage is driven.

Coming off the delay stage, the beam is then directed onto a mirror that can be rotated to send the UV beam either towards our vacuum chamber here, or the other experiment using the UV beam (dashed blue line). The beam is then sent through a half-wave plate that is designed for 266 nm in the first instance, but which will also be effective at wavelengths near this wavelength. It is then steered onto the dichroic mirror DM, where it is reflected through the irises into the vacuum chamber at (5). The UV beam was deliberately reflected off this mirror (rather than sent through it) to minimise added GDD.

Both beams are directed through a focussing lens at point (5), and this lens is a 300 mm singlet lens made of UV fused silica which is AR coated at 800 nm. Singlet lenses are preferable as they are thinner, and so add less GDD - but chromatic aberration needs to be compensated for as we need to overlap the 800 nm beam with the 266 nm beam. In this beamline, we would adjust the divergence of the IR beam with telescope TS to overlap the foci on the molecular beam. The lens is also not AR coated for the UV wavelengths, as the lens would have to be custom coated for this, which would be prohibitively expensive. This means that there will be some UV power lost to reflection (around 5%), and also an additional stray beam that needs to be safely blocked. The vacuum window the beams are focussed through is also made of UV fused silica, and is AR coated for 800 nm too. The

Optical Element	Quantity	GDD per element (fs^2)	Cumulative Pulse Duration (fs)
800nm mirrors	9	30	41
HWP	2	80	49
TFP	1	120	56
Telescope	1	160	67
Focussing Lens	1	80	72
Air	5m	100	79
Vacuum Window	1	160	**91**

window is 4 mm thick.

A final note about the point (5), where there is a mirror shown with by an orange asterisk. This mirror is removable, and is placed after the focussing lens to direct the focussed beams down towards the beam characterisation equipment. In practice, this is is where you would put a beam profiling camera or a photodiode to look at your focal spot sizes, and to make sure the beams are overlapped. Doing it as close as possible to the vacuum window ensures that when the mirror is removed, the beams remain overlapped in the vacuum chamber.

Pulse Durations

We can now see that this beamline will perform as we expect, and we have chosen all the optical elements such that it will perform to the requirements we set initially. However, what we have not yet done is consider the final pulse durations in our chamber. We know the pulse duration at our laser output, and we can estimate the final pulse duration in our chamber by calculating how much GDD will be added to each beam as it propagates. We will do this for each beam in turn.

In all cases, the numbers for GDD were either obtained from the manufacturers of different optics, or (in the case of transmissive optics) found in online databases of GDD data of certain materials, where the GDD added by a given thickness of a material can be easily looked up.

IR Beam

Ideally the IR output is a transform-limited 35 fs 800 nm pulse. We then need to calculate the GDD added to this from each optical element. This is easily done using online calculators, such as the light conversion optical toolbox, and the results are summarised in the table below. 'Worst-case' numbers for the GDD have been assumed throughout, so this gives an upper estimate for the final pulse duration.

So, in a worst case scenario, our initially transform-limited IR pulse will end up broadened by almost a factor of 3 by the time it gets to our vacuum window! We can compensate for this to some extent by moving the compressor inside the laser system, but we cannot move this too far as the OPA is very sensitive to input pulse duration, and moving it a long way will severely impair the function of the OPA. This number may seem alarming, but it

Optical Element	Quantity	GDD per element (fs^2)	Cumulative Pulse Duration (fs)
Al Mirrors	11	20	47
HWP	1	400	59
Dichroic Mirror	1	30	60
Focussing Lens	1	400	79
Vacuum Window	1	800	123
Air	3m	300	140

is in likely to be a gross overestimate, as the numbers assumed are all worst-case numbers, and if the laser output is not fully transform-limited, then the broadening will also be less severe. A pulse duration of around 70 fs in the vacuum chamber is likely to be a more realistic estimate.

UV Beam

We now perform the same analysis with the UV beam. Here we do not know the exact output pulse duration, but it is likely to be slightly broader than the IR pulse duration after all of the processes occuring in the OPA. For the purpose of this calculation we will assume we start with a 45 fs, 266 nm pulse.

As we can see, the UV pulses are substantially more affected by the broadening. It should be noted that again these are 'worst-case' numbers, as the Al mirrors probably add much nearer to zero GDD than this number, which was taken as the manufacturer's quoted upper limit. However, this illustrates that air and transmissive optics have the potential to cause lots of unwanted broadening for these pulses. An additional UV compressor may be required if much shorter UV pulses were required, but the numbers here considered to be reasonable as a place to start.

8.2 Transient Absorption Spectroscopy

8.2.1 Initial Equipment

Again, we start by considering what initially existed. The driving laser system was similar to that used in the previous example: a Ti:Sa laser system, producing pulses at 800 nm centre wavelength, 35 fs duration, and 2 mJ pulse energy at a 1 kHz repetition rate. The beam diameter is again around 11 mm, and the height from the table is 160 mm. A detection system consisting of an imaging spectrograph and camera existed and can be height-adjusted such that the entrance slit is 160 mm from the table surface.. No OPA was present in this case.

8.2.2 Requirements

This beamline is intended for use in an electronic[2] transient absorption (TA) spectroscopy experiment. Fundamentally TA is a similar pump-probe technique to that described above: we need two beams, a pump and a probe, and need to be able to vary the time delay between them. We also want to be able to tune the colour of the pump pulse. However, now our sample is in the liquid phase, we don't have an OPA to provide the tunability, and we need a broadband probe pulse so we can collect a whole spectrum in one laser shot. Thus, we can again create a list of laser conditions we need at the sample:

- **Beams**: We need to overlap two beams in space and time in our sample, and control the time delay between them with femtosecond precision.
- **Position**: The target (after the sample), is to focus our probe beam into an imaging spectrograph. The entrance slit to this spectrograph is 160 mm higher than the surface of the optical table.
- **Colour**: The pump pulses will ideally have tunability across the visible range (400-650 nm). The probe pulses will be as broadband as possible within the visible range.
- **Size**: The intensities in this TA setup do not need to be as high as in the Coulomb Explosion imaging case, because we are not relying on a highly nonlinear strong field ionisation process to create our signal. Larger beam waists can therefore be used, generally on the order of 100 µm[3]
- **Pulse Duration**: 'as short as possible' is obviously nice, depending on the planned experiments. Sub 100 fs would certainly be nice.
- **Intensity**: As mentioned, this does not need to be as high as in the previous case. Typically, somewhere around 1×10^9 W cm^{-2} is sufficient. At a beam waist of 100 µm and 100 fs pulse duration, this corresponds to around a 20 nJ pulse energy[4]
- **Polarisation**: Variable linear polarisation of the pump beam is required. The probe beam should be linearly polarised (and ideally matched to the preferred polarisation state of the gratings in the detection system[5].

Again, comparing these requirements to what we initially have:

- **Beams**: We will need to add a motorised delay stage to either the pump or probe beamline..

[2] Although many of the ideas are applicable to other kinds of TA – such as IR or THz TA.

[3] A large beam waist will 'talk' to more molecules in the sample, but require more photons (= pulse energy) in the pump and probe beams. A smaller beam waist will require fewer photos.

[4] Depending on the absorption coefficient of the sample, this would create a 1-10% excitation of the sample. Such excitation is around what is generally wanted in TA: it would produce a differential absorbance of between 5 and 50 mOD. REF

[5] Which is normally when the polarisation is orthogonal to the lines on the grating, but it can depend on the grating.

- **Position**: No adjustments to height are needed. An appropriate lens will be needed to focus the probe light into the spectrograph.
- **Colour**: We can create a manually-tunable NOPA to provide the tuning needed in the pump pulse. White-light generation of the fundamental beam in a medium like sapphire will create the broadband probe spectrum needed.
- **Size**: Larger beam waists can be used, but the large beam that comes out of the laser is slightly cumbersome to work with, and will focus down to a much smaller spot than we want on our sample unless we use a lens with a very long focal length (even a lens with a focal length of 1 m would focus this large beam too tightly). Furthermore, a large beam requires that optics like waveplates and nonlinear crystals have a large clear aperture – which can become expensive. Thus, it will likely be best to shrink this beam by a factor of two using a telescope.
- **Pulse Duration**: Minimising transmissive optics (especially for broadband pulses) will mitigate much damage here. Given the short pulses that come from the laser, we don't expect that pulse lengthening will be a significant issue.
- **Intensity**: Easily attainable, but will require variable attenuation in both pump and probe beam paths.
- **Polarisation**: Adding a half-wave plate to the pump path will be necessary, and it will need to be achromatic to accommodate varying pump wavelengths. Polarising the broadband probe (if needed) is best accomplished using a wire-grid polariser.

8.2.3 Design and Construction

The layout of the setup is shown in Figure 8.2. As before, we will walk through this setup. It can be divided into four areas, annotated with circled numbers on Figure 8.2:

1. **Input optics**: dividing up the pump beam and manipulating the size.
2. **Probe generation**: generating a broad white-light continuum to use as the probe beam.
3. **Pump generation**: generating tunable visible pulses to use as the pump beam.
4. **Sample interaction**: interaction with the sample and focussing into the detection system.

We will discuss each of these areas in turn.

Input Optics

The laser produces pulses with an energy of 2 mJ. Not all of this energy is needed for this application, so the first optic in the beamline is a 50:50 beamsplitter, BS1, which should be

a broadband beamsplitter, ideally optimised for ultrafast use. We take the reflected beam from BS1[6].

After BS1, we immediately shrink the beam by a factor of two using a telescope – we opt for a transmissive telescope here as the pulse energy is not too high (so we do not expect self-focussing to be a problem), and to make alignment easier. To shrink by a factor of two requires the ratio of focal lengths to be 0.5 (seeExample 6.4 for an example). Thus, we can choose any two lenses to accomplish this, noting that the distance they must be placed apart is the sum of the focal lengths (as we may have spatial constraints in our setup). A flick through a manufacturer's catalogue will show the lenses you can usefully buy, and define the available options. A plano-convex (converging) lens with focal length 200 mm and a plano-concave (diverging) lens with focal length of −100 mm would fit the bill nicely. Note some tips:

- Ideally we mount one of these on a translation stage, to make alignment easy.
- The curved faces[7] of each lens point towards the collimated beam to minimise aberrations.
- Ensure that the beam hits the center of the lens - check the back-reflections to make sure they are aligned square to the incident beam.

After this, we have a beam containing 1 mJ pulses. We split this on a 90:10 beamsplitter (BS2), sending 100 µJ to generate the probe pulse (stage 2), and the other 900 µJ to generate the pump pulse (stage 3).

Probe Generation

Initially considering the probe generation, we take our 100 µJ and immediately send it through a motorised delay stage – giving us our variable delay. It's important that this stage is motorised, because you don't want to be standing around for hours moving this as you measure a spectrum. Delay stages are expensive and there are lots of different kinds:

- Here, it's important that this stage is pretty 'good'. We don't want the beam to be walking around as it is scanning, and want a good repeatability.
- We want to scan delays up to around 1000 ps. The stage length thus needs to be at least 150 mm (in a double-pass geometry). Using longer stage adds cost but makes finding t0 easier.
- The minimum repeatable incremental motion needs to be smaller than the time resolution you want. 0.1 µm repeatability in distance corresponds to 0.67 fs in time (again, assuming a double-pass geometry).

[6] Leaving the slightly broadened pulses for whoever is building a setup next to us.

[7] A handy tip to spot these is to hold the lens under a ceiling light – the flat side will show you a mirror-like reflection, the curved side will show you a fairground hall-of-mirrors type reflection.

- Being able to scan this stage quickly is useful to maximise the time spend collecting data, and minimise the time spent moving a stage.

The mirrors on this delay stage can be either dielectric or metallic, but I would err towards metallic on the grounds that:

- The pulse energy is low, damage is unlikely to be an issue.
- It gives us the option of generating the second harmonic before the stage if we want to use that to generate our white-light continuum (see later).
- It's cheaper.

But either would work fine. If I was building this, I'd buy a corner-cube retroreflector or similar to remove the hassle of aligning two mirrors. Make sure you set up some alignment irises on the input to the stage so you can ensure it is running true.

After the delay stage, we steer into the white-light generation (WLG) part of this stage. A variable reflective ND filter (VND) is essential here, as the WLG is very sensitive to input power. Ensure that you get a continuously variable VND filter, rather than one that has discrete steps. Again ensure this is aligned properly by looking at the back-reflection – and ensure that you have a beam dump in place (if needed) to catch the back reflection, as at high attenuations this filter will reflect a considerable amount of the input power.

After the VND, an iris is necessary to help control the generated white light. The beam then is focussed into a sapphire plate (around 3 mm thickness is sufficient) using a spherical curved mirror[8], and the generated beam is then recollimated using a second curved mirror – a curved mirror is essential here because you need to effectively collimate all the colours in the (now very broad) pulse. The focal lengths of these mirrors will depend a little on the space you have available, but somewhere around 100 mm is in the right ballpark. Notes here:

- Figure 8.2 shows these mirrors aligned with quite a large angle of incidence. In reality, you want the angle of incidence here to be as shallow as possible to minimise aberrations caused by the mirrors.
- Getting a stable white light continuum requires a bit of fiddling with the input power, input iris, and focussing. Try to focus the light at the far end of the sapphire, and use as low a power as possible initially. The priority here is a *stable* continuum, not the most colourful one.
- Materials other than sapphire (such as CaF_2) are also possible to us here, as discussed earlier. Sapphire is preferred because it has a high damage threshold.
- By placing a BBO crystal and filter in the beam path before the VND, you could

[8] A singlet lens could also be used for the input focussing, provided we aren't worried about dispersion.

pump the WLG with 400 nm instead – this can give you a more blue continuum, which may be useful for some applications.

After the white light is generated, we pass through a notch filter (NF) to remove residual fundamental light, and then through another iris to help clean up the continuum and give us a nice probe pulse. From here on, until interaction with the sample, we want to avoid any transmissive optics to keep our probe pulse as short as possible: being so broadband, it will be affected greatly by dispersion through materials. It's useful to have a little USB spectrometer about that you can use to monitor what the probe pulse spectrum is.

The probe pulse is now sent into the sample, and we will pick up it's journey later.

Pump Generation

Going back to BS2, we sent 900 µJ of light to generate our tunable pump pulses. We are going to generate these using a NOPA. Comparison the area around stage 3 in Figure 8.2 with Figure 6.19 will hopefully show some similarities: a NOPA seeded by a white-light continuum and pumped by the second harmonic of our fundamental.

Immediately after BS2, we hit another 90:10 beamsplitter, BS3. This sends the majority of the pulse energy to go and create a 400 nm pump beam through the BBO crystal BBO1. In this case:

- BBO1 would be cut for SHG at 800 nm, so at 29.2°.
- The mirrors following BBO1 will be dielectric mirrors that reflect the second harmonic and not the fundamental - ensure you place beam blocks behind those that have a lot of fundamental leakage passing through them.
- We then have a manual variable delay stage. This doesn't need to be anything fancy or with a long travel (25 mm is plenty), unlike the delay stage in the probe generation stage. We won't be adjusting this mid-measurement, but will change it to optimise the pump wavelength before each experiment.

After the variable delay, the beam is gently focussed into the crystal BBO2, via a half-wave plate (HWP) and through a focussing lens. The focal length of this lens will depend mostly on the geometry of your setup, but something like 200 mm would be where I'd aim to start.

Coming back to BS3, the smaller fraction of the pulse energy enters a WLG stage exactly analogous to the one in stage 2, the discussion around this will not be repeated here. In this case, however, the WLG is focussed gently into BBO2 to overlap with the pump. This is where the OPA process occurs. It's essential that you have BBO2 mounted in a precise manual rotation mount – allowing rotation of the crystal around an axis orthogonal to the table, and around the axis the beams are incident on it at.

As this is a NOPA, it's important that these two beams are non-collinear. As we know, a non-collinear geometry means we can maximise our phase-matching bandwidth and produce

a huge range of colours. The angle the beams cross at depends on the wavelengths and the crystal you use – in this case, let's assume we are using a crystal[9] cut at 32.5°. The angle the two beams need to cross at *inside the crystal*[10] is then about 4.5°. The angles *inside* the crystal are different from the external angles the beams come in at *outside* the crystal because of the refractive index of the crystal. But you can figure out the right external angle by sketching and using Snell's law, an instructive exercise for you :) Once you have the angle right, then the pump beam alone should generate a broad cone of parametric superfluorescence[11], which probably looks green-ish. You want to overlap the seed beam with this cone – which might take some wrangling but should be easy enough as the beams are loosely focussed.

Ensuring you have temporal overlap could be more tricky, but measuring the two beam paths after BS3 (use a photodiode, or the budget option: string) will tell you if you are close. Match the paths as best you can and then move the variable delay stage until you can see a flash of light after the crystal that depends on the delay. It will again, likely take some fiddling, unless you have *The Knack*[12]. Once you have it, you'll get some bright output that you can optimise – using a spectrometer to make sure you have the light you actually want is essential here. You should be able to tune the light from this NOPA over a relatively broad range by tweaking the delay and crystal angle[13].

Once the NOPA is running, our pump pulses go through an iris (to clean up any residual crap from the NOPA), then enter the sample interaction stage.

Sample Interaction

Our pump and probe beams now need to overlap in a sample. We'll assume the sample is a simple cuvette, but it could equally be a flowing jet or liquid film, or something. Usually you want to ensure the cuvette is thin so that the pulses remain ultrashort and don't get broadened in the cuvette.

Staying with the pump beam at first, after BBO2, it is cleaned up using an iris, and then has it's polarisation controlled using a HWP. This waveplate needs to be achromatic (if we are going to vary the pump wavelength), and is there to ensure that the angle between the pump and probe polarisation is 54.7° – or the *magic angle*. This is the angle where effects due to rotational anisotropy in the sample cancel out[14] – of course you may not want this, and would want to measure the anisotropy. Either way, you need a HWP there. Then, there is an ND filter to control the power (this could be variable, depending on budget), and then an optical chopper (CH) – which will be synchronised to the laser repetition rate

[9] One commonly available from Eksma Optics.

[10] See Figure 31 from Manzoni and Cerullo [1], for example.

[11] Sort of the unseeded equivalent of the output you want.

[12] https://www.youtube.com/watch?v=g8vHhgh6oM0

[13] There is a limit to how much guidance you can transfer in something as practical as OPA alignment through a book, though!

[14] It's the angle where a second-order Legendre polynomial, $3\cos^2\theta - 1 = 0$. Rotational dependence of the interaction with the sample depends on this polynomial - so setting the magic angle removes this dependence.

and knock out every other pump pulse: enabling the differential absorption:

$$\Delta A = A_{\text{pumped}} - A_{\text{not pumped}}$$

to be measured. Finally, a curved mirror collimates the beam from the BBO. The beam is then steered off a zero-degree mirror, through a lens, and into the sample. The focal length of this lens is chosen to ensure that the pump spot is larger than the probe spot on the sample. Given our requirements:

- We assume our beam size is unchanged from the telescope in stage 1, so a beam waist of around 3 mm (a radius).
- We want a spot size of around 100 µm – or a little larger.
- Plugging these numbers into the Light Conversion toolbox will show that a fairly weak lens of around 750 mm will do the job for us.
- Mounting this lens on a translation stage will help to optimise the pump beam size on the sample, and correct for any chromatic aberration as the pump colour is varied.

The pump beam is blocked by an iris after interaction with the sample.

The probe beam then leaves stage 2 and the notch filter (NF). It then passes through a tilted thin window where a reflection is incident on a photodiode (PD). The signal from this photodiode is used to correct for shot-to-shot fluctuations in the probe pulse energy, due to the instability of the WLG process – which can cause significant noise in spectra[15]. Note that the window needs to be thin to minimise dispersion. Then, the probe pulse is focussed using a parabolic mirror (PM - so no spherical aberrations) onto the sample:

- The probe beam should be focussed to a smaller spot than the pump beam – just as was shown in the laser amplifier in Figure 4.1. Ensuring that the probe is only probing molecules that have been effectively pumped.
- The positioning of the parabolic mirror PM, or of the sample, can be adjusted to optimise the probe beam size at the sample. I would place PM first and then decide where to put the sample.

The probe beam after the sample needs to be focussed effectively into the spectrograph and detector. The exact lens to use here depends on the focal length of your spectrograph, but the idea is basically to ensure that you fill the grating on your spectrograph as much as possible to maximise resolution (but without losing too much light over the sides of it)[16].

You may find that you need another ND filter after the sample, if the probe beam is intense

[15] A better solution is to spectrally resolve the WLG for every laser shot and reference the measured absorption to it – but this requires another (or a different) detector.

[16] A nice tutorial on this is given in the Newport Optics Technical note '*Getting Light Into a Monochromator*' [2]

and you have a sensitive detector. Always start using the lowest power you can get away with, and then turn things up!

Now, you can start to look for temporal and spatial overlap in the sample. Spatial overlap should be relatively easy to find by eye with these large beams, but temporal overlap is likely to be most easily found if you replace the sample with a BBO crystal or similar – you can then easily look for sum-frequency generation between the pump and probe beams to define your t0. You should now be ready to do some spectroscopy!

8.3 Other Beamlines

Finally, just to note that other kinds of beamlines have often described in literature, in varying levels of detail. Recently there have been descriptions of beamlines for sum-frequency generation spectroscopy [3], femtosecond-stimulated Raman spectroscopy [4], and others in literature. Even setups for imaging and microscopy[17] have been described [5] – searching literature is always useful. Usually you'll find someone has made something similar, or at least you'll get some helpful and useful ideas. Hopefully, with a combination of this book and some literature, you'll quickly feel confident working in the lab and building your own setups.

Bibliography

[1] Manzoni C and Cerullo G. Design criteria for ultrafast optical parametric amplifiers. *J Opt*, 18, 2016.

[2] Newport Optics. Newport optics tutorials. https://www.newport.com/resourceListing/tutorials. Accessed: 23-12-2020.

[3] Pickering J D, Bregnhoj M, Chatterley A S, Rasmussen M H, Roeters S J, Strunge K, and Weidner T. Tutorials in vibrational sum frequency generation spectroscopy. ii. designing a broadband vibrational sum frequency generation spectrometer. *Biointerphases*, 17, 2022.

[4] Batignani G, Ferrante C, Fumero G, Martinati M, and Scopigno T. Femtosecond stimulated raman spectroscopy. *Nat. Rev. Methods Primers*, 4, 2024.

[5] Young M D, Field J J, Sheetz K E, Bartels R A, and Squier J. A pragmatic guide to multiphoton microscope design. *Advances in Optics and Photonics*, 7, 2015.

[17] Which lie far outside of my comfort zone!

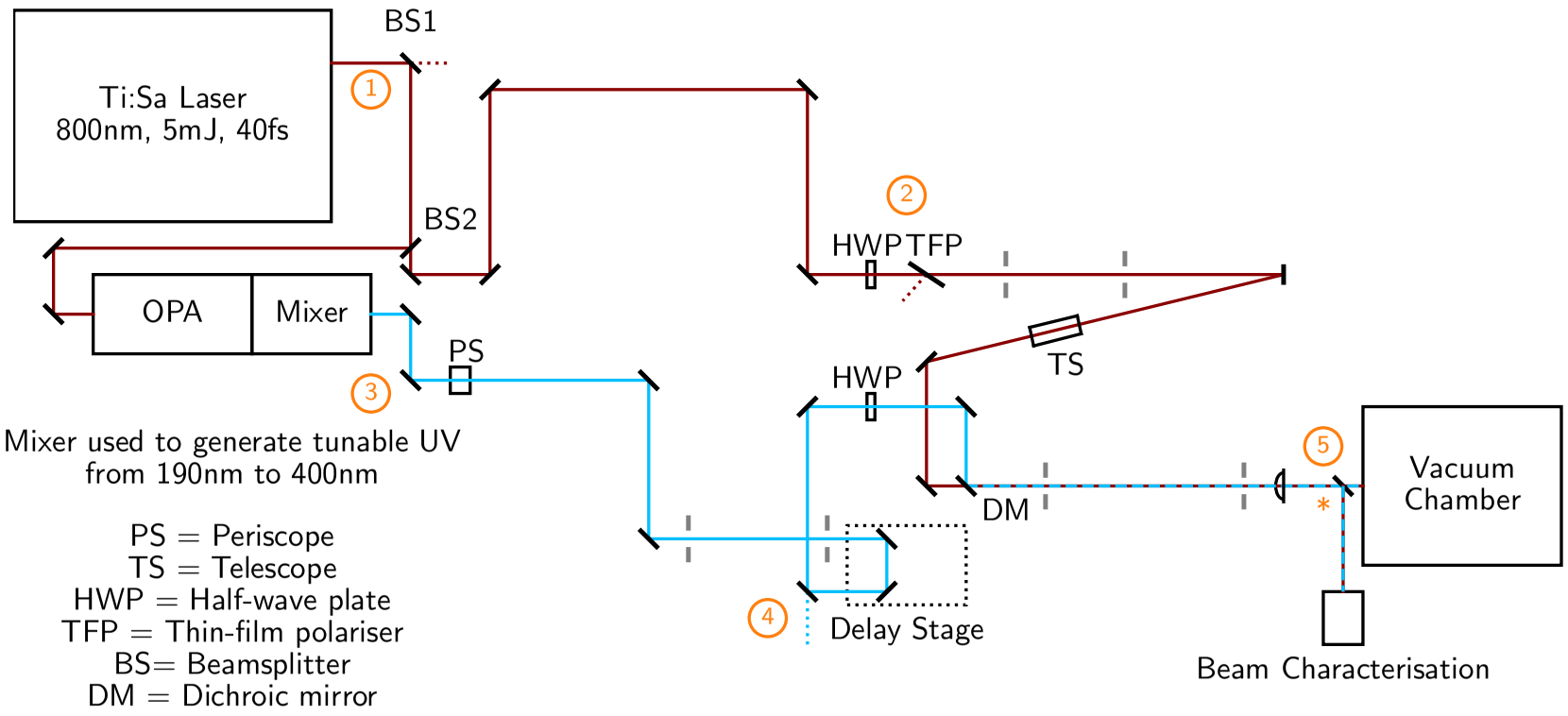
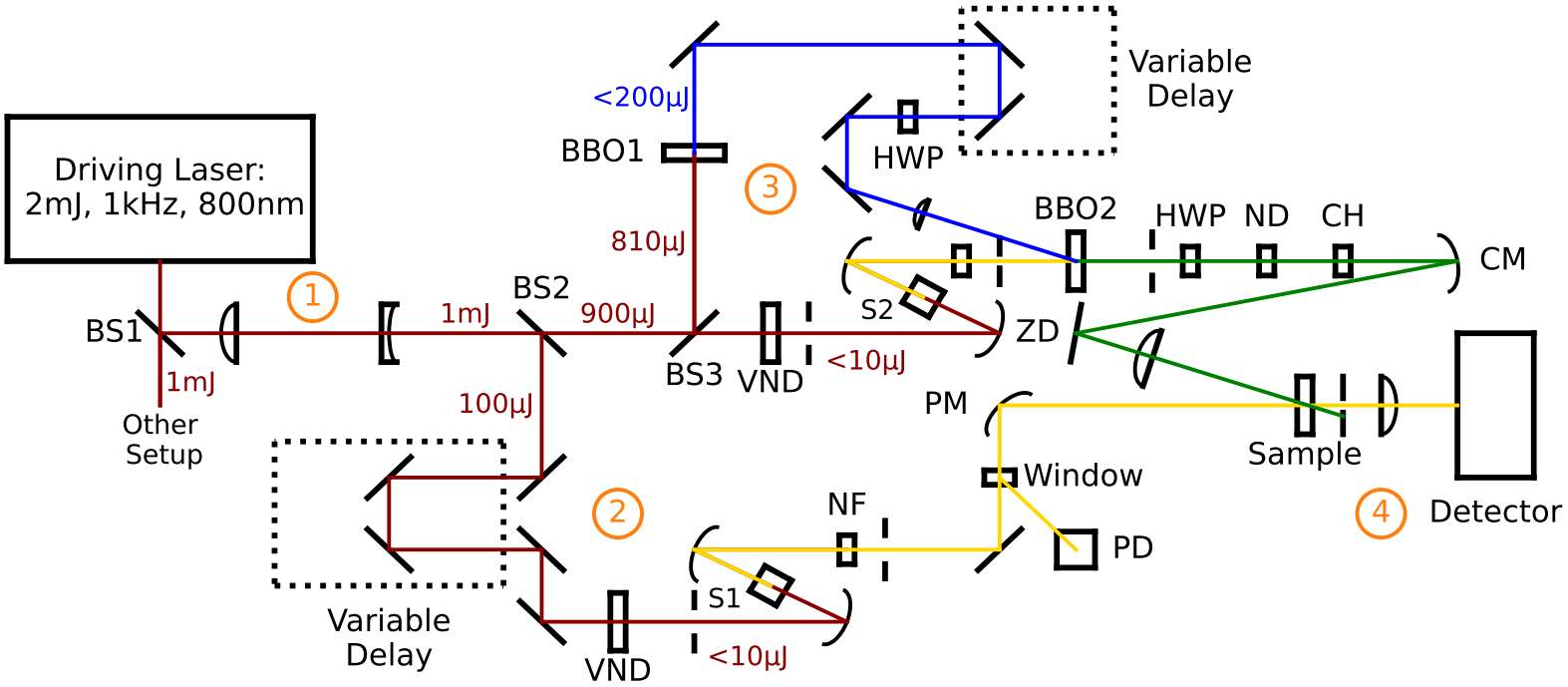


Figure 8.1: A UV-IR pump-probe beamline. See text for details



BS = Beamsplitter, BBO = BBO Crystal, HWP = Half-Wave Plate, ND = Neutral Density Filter, PM = Parabolic Mirror, NF = Notch Filter, VND = Variable ND Filter, CH = Chopper, PD = Photodiode, S = Sapphire, ZD = Zero Degree Mirror

Figure 8.2: An transient absorption spectroscopy beamline. See text for details.

Appendix A

Electromagnetic Waves

In my experience people entering this field from a non-physics background can struggle with some of the concepts surrounding the mathematical description of electromagnetic waves. For a trained physicist, this is generally just assumed knowledge and so it can be demoralising[1] for a new student to not understand what is meant by terms like 'phase', or 'k-vector'. However, it is rather straightforward if broken down simply. A more mathematical treatment can be found in a standard text on electrodynamics, such as Griffiths [1].

A one-dimensional travelling electromagnetic wave $E(z,t)$ can be expressed as follows:

$$E(z,t) = E_0 \exp\left[\mathrm{i}(kz - \omega t)\right] \tag{A.1}$$

Physically, this wave could be the electric field of some laser light. The wave is 'travelling' because it is moving in both space (the coordinate z), and time t. Let us consider the meaning of each of the terms in Equation A.1 in turn.

- E_0 refers to the **amplitude** of the electric field - the maximum height of the peaks in the wave. This would have units of volts per unit length.
- t is time, with units of time.
- ω is the **angular frequency** at which the wave oscillates *in time*. It has units of radians per unit time, such that the product ωt has units of radians, which are dimensionless[2]. This is normally just called the **frequency**. *The more cycles the wave completes per unit time, the larger the frequency*
- z is the position of the wave in space along the z-axis, with units of length.
- k is the **angular wavenumber** of the wave. This has units of radians per unit length, such that the product kz has units of (dimensionless) radians. It is normally

[1] It was for me, at least.

[2] A radian is defined as the ratio of arc length to radius length of a circle, thus the units of length cancel out and the radian is dimensionless.

just known as the **wavenumber**. This can be thought of as a **spatial frequency**, where ω was a **temporal frequency**. *The more cycles the wave completes per unit length, the larger the wavenumber.*

- i is the imaginary unit, defined such that $\mathrm{i}^2 = -1$. This will be discussed further below.

The wave is written in an exponential form, but is really just a sinusoidal wave, as Euler's formula states that:

$$\exp(\mathrm{i}\theta) = \cos(\theta) + \mathrm{i}\sin(\theta) \tag{A.2}$$

So if we took the real part of our wave $E(z, t)$, we get:

$$\mathrm{Re}[E(z, t)] = E_0 \cos(kz - \omega t) \tag{A.3}$$

Which is the sinusoidal form we expect. We use the exponential form as it is the most general, and it makes manipulating the wave much simpler when we try to add phase factors and things – but remember it just refers to something sinusoidal. To visualise all the parameters above, it is easiest to plot the wave, but before we do this there is some mathematical complexity that needs to be cleared up.

You will probably encounter multiple definitions of the wavenumber, unfortunately. Within spectroscopy and chemistry it is normally thought of as the reciprocal of the vacuum wavelength of a particular spectroscopic transition, with the symbol $\tilde{\nu}$:

$$\tilde{\nu} = \frac{1}{\lambda} \tag{A.4}$$

This definition is useful in chemistry where desire is really just to have a number that is linked to the transition wavelength but is directly proportional to transition energy. However, in the context of laser physics, we define the wavenumber, k, as:

$$k = \frac{2\pi}{\lambda} \tag{A.5}$$

Which has units of radians per unit length, as mentioned above. This is really the 'angular wavenumber', but it is normally just called the 'wavenumber', like the angular frequency is just called the frequency. We do this because we are always talking about waves that are periodic, and what is interesting is how often the wave completes a complete periodic 'revolution' around 2π radians (a circle). So how many radians our oscillating wave moves through in a propagation length, or propagation time, is what interests us.

With this in mind, we should think about how we can link together *the number of oscillations per unit length* (k) with *the number of oscillations per unit time* (ω). It seems natural that these should be connected: if the wave is oscillating through a certain number of radians in a certain time, and is also moving through space, then the **speed at which it moves through space** will dictate how many radians it oscillates through in a given distance. That is:

$$k = \frac{\omega}{v_\mathrm{p}} \tag{A.6}$$

Where v_p is the **phase velocity** of the wave, which is how fast it is moving in whatever medium it is travelling in. If the wave moves faster, v_p is larger, and the wave won't have time to oscillate through as many radians in a given distance than it would if it was moving more slowly. This is simply what Equation A.6 expresses mathematically.

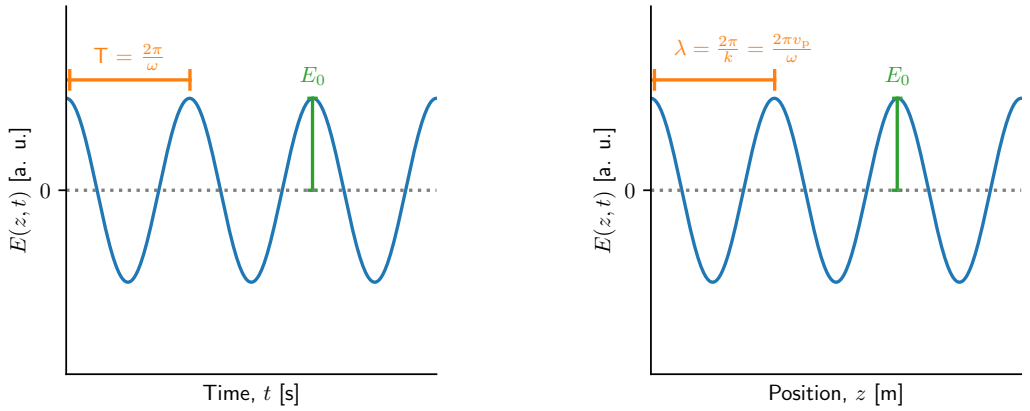


Figure A.1: Illustration of our wave $E(z, t)$ plotted as a function of time (left) and position (right). Quantities discussed in the text are annotated.

Figure A.1 shows graphically all of the quantities we have discussed, and how they relate to one another. We have also defined the reciprocal of the frequency, the **oscillation period**, $T = {}^{2\pi}/_{\omega}$, in the leftmost plot on Figure A.1; and the **wavelength**, λ, in the rightmost plot. Both of these quantities represent the time taken (T) for, and the distance travelled (λ) in one full oscillation (through 2π radians). It is clear from the figure that the wavelength, λ, can be defined in terms of previously met quantities as:

$$\lambda = \frac{2\pi}{k} = \frac{2\pi v_p}{\omega} \tag{A.7}$$

However, at the start of this section we specifically said that we would clarify the meaning of the **phase** and of the **wave vector** - two of the concepts that cause most confusion in this topic in my experience of teaching it. Before we discuss the phase in the following section, we will briefly discuss the wave vector, **k**.

A.1 Wave Vectors

The wave vector **k** looks like the wave number k, but is in bold. This is because the wave vector is a **vector**, so has a direction and a magnitude. **The magnitude of the wave vector is simply the wave number**. The wave vector points in the direction of propagation of our electromagnetic wave (the direction of propagation of a laser beam, for example). In the case that the wave is travelling in 3D space, rather than the 1D case shown above, then the direction of travel can be split into three components, corresponding to movement along the x, y and z axes. In this case we make the substitution:

$$kz \rightarrow \mathbf{k} \cdot \mathbf{r} = k_x\hat{\mathbf{x}} + k_y\hat{\mathbf{y}} + k_z\hat{\mathbf{z}} \tag{A.8}$$

Where $\mathbf{r}$ is our position vector in 3D space, which can be split into three components in terms of the unit vectors $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$. The magnitude of each of these components is k_x, k_y and k_z respectively. The total magnitude k_{3D} of our 3D wave $\mathbf{k}$ is given straightforwardly by:

$$|\mathbf{k}| = k_{3D} = \sqrt{k_x^2 + k_y^2 + k_z^2} \tag{A.9}$$

The value k_{3D} given in Equation A.9 is simply the three-dimensional analogue of the one-dimensional k used Equation A.1. To summarise, the important points are:

- The wave vector, $\mathbf{k}$, is a vector that points in the direction that the wave is propagating in.
- The magnitude of the wave vector is the wave number, k, and tells you how many cycles the wave completes in a unit propagation distance (units of m^{-1}).

In the context of lasers, people often talk about **k-vectors**, and these are just another name for wave vectors. Specifically, we talk about them in the context of momentum conservation in phase-matching in nonlinear optics. This is because we can link the magnitude of the k-vector k to the momentum of the wave p:

$$p = \hbar k \tag{A.10}$$

A simple dimensional analysis illustrates this. Momentum has units of $\mathrm{kg\,m\,s^{-1}}$. The magnitude of the k-vector has units of $\mathrm{rad\,m^{-1}}$. The reduced Planck constant $\hbar$ has units of $\mathrm{kg\,m^2\,s^{-2}\,rad^{-1}}$. So, a larger k-vector (or shorter wavelength, or higher frequency) corresponds to a higher momentum for the wave. We will now turn to a discussion of **the phase**.

A.2 Phase

The concept of phase causes confusion among students in my experience, but it need not. Fundamentally, **the phase of a wave tells you which part of the cycle it is in.** For example, a wave with a phase of π rad is halfway through a cycle, and a wave with a phase of 0 rad is at the beginning of a cycle. As such, **the phase is an angle**, given in radians. Recall however that a radian is dimensionless, so you will equally see it said that the phase is dimensionless. The phase is generally given between 0 and 2π, but sometimes the phase can be greater than 2π, if a wave finishes a cycle and goes onto the next one[3].

So the phase is just an angle that tells you which part of the oscillation cycle you are in. If it seems odd that we use an angle to define this, remember that our wave is a periodic sinusoidal function, as shown in Equation A.3. The argument that this function takes is an angle, so it's natural that we use an angle to define where 'on' this function we are. Now, we need to distinguish between the **absolute phase** of the wave, and a **phase shift** or **accumulated phase** that is added to the wave.

[3]Sometimes it's then useful to *unwrap* this phase so it stays between 0 and 2π. This will produce discontinuity in a plot of the phase, which is sometimes informative.

The **absolute phase** tells us exactly where we are in the wave cycle overall. To find the absolute phase of our one-dimensional travelling wave, we note that we wrote Equation A.1 in exponential form deliberately. This exponential form may be familiar from study of complex numbers in mathematics, where a complex number N can be written as:

$$N = |N| \exp(i\Theta) \tag{A.11}$$

Where $|N|$ is the modulus (magnitude) of the complex number, and Θ is the argument (phase) of the complex number. So we can identify the stuff in the exponent next to the imaginary unit as our phase. This means that the absolute phase of the wave in Equation A.1, which we will call Θ, is given by:

$$\Theta = kz - \omega t \tag{A.12}$$

Both quantities kz and ωt are angles as discussed above, so the difference $kz - \omega t$ is also an angle, and defines the absolute phase of our wave. You may sometimes see it written as:

$$\Theta = kz - \omega t + \arg(E_0) \tag{A.13}$$

Which accounts for the possibility that the amplitude E_0 is also a complex number with it's own phase. The argument of the complex number is just the phase, so $\arg(E_0)$ is the phase of the amplitude which also contributes to the total absolute phase. But in the examples we consider here E_0 is simply a number, so does not have a phase (or has a phase of zero).

We can further exploit the beauty of the exponential form of complex numbers to understand what we mean by a **phase shift** or **accumulated phase**. A phase shift is when we move our wave along in its cycle by a given angle. For illustration, we will call this angle ϕ. Ultimately what we are doing with a phase shift is:

$$\Theta' = \Theta + \phi \tag{A.14}$$

Where Θ' is the absolute phase of the wave after the shift, Θ is the absolute phase of the wave before the shift, and ϕ is the added phase. The exponential form of complex numbers makes this trivially simple. To shift our initial travelling wave $E(z, t)$ by ϕ radians, we simply do:

$$E(z, t) \times \exp(i\phi) = E_0 \exp\left[i(kz - \omega t + \phi)\right] \tag{A.15}$$

So multiplication by the **phase factor** $e^{i\phi}$ caused a phase shift of our wave by ϕ radians. This is what we mean when we talk about phase shifts or accumulated phase, here we have accumulated a phase of ϕ radians.

We can get a feel for what this looks like by considering what happens if we shift our wave by some different values of ϕ. This is illustrated in Figure A.2. Here we still plot our waves as a function of time, but you can see that the added phase has the effect of moving the part of the cycle that a wave is in at a *given time* around. That is, if you were to pick a specific point on the time axis (such as that shown with the dashed line on the rightmost plot in Figure A.2), then all the waves are clearly at different points in their cycle. The orange curve is $\pi/2$ radians ahead of the blue curve, and so on. This kind of plot can be

difficult to visualise as it looks as though the waves with a positive phase shift are being moved backwards. The best way to think about it is to look at the unshifted wave (blue), and then look at where it is at the dashed line. Now consider where the blue wave would be in $\pi/2$ radians time - this is where the orange wave is at the dashed line.

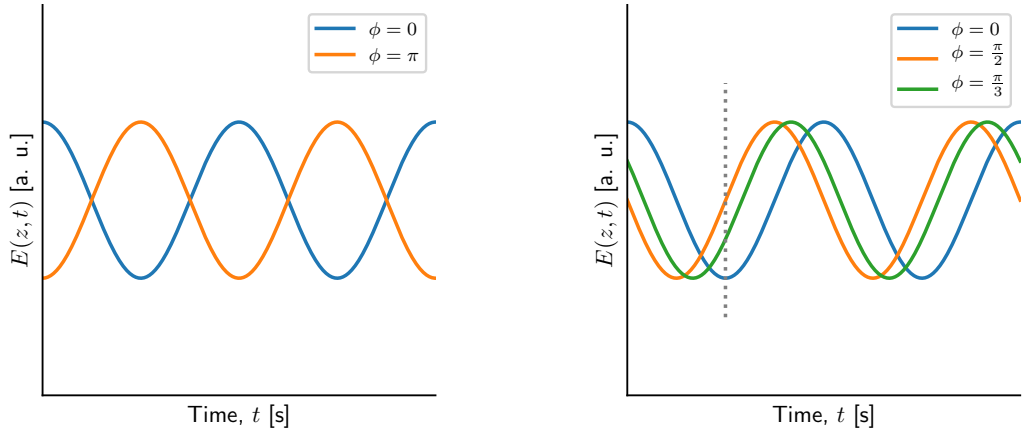


Figure A.2: Illustration of our wave $E(z,t)$ plotted as a function of time with different phase shifts, ϕ, applied.

To end with a laser based example, often we talk about the spectral phase $\phi(\omega)$ that is accumulated by a pulse as it propagates through a medium. This just means that a wave in the pulse with frequency ω is moved in its cycle by $\phi(\omega)$ on propagation through the medium. $\phi(\omega)$ could be a complicated function of ω, which is what gives rise to the GDD and higher order dispersions that cause our pulse to broaden during the propagation.

Bibliography

[1] Griffiths D J. *Introduction to Electrodynamics*. Cambridge University Press, Cambridge, 4th edition, 2017.

Appendix B

Useful Resources

The internet contains a wealth of useful resources that are often much more easily accessible to a working scientist than a traditional textbook, and can present ideas at a more accessible level than many journal articles. Here a selection of those that are most useful for people working within optics and lasers is listed - all of these have helped me in the past, and many of them have been drawn on when writing this book.

- **RP Photonics Encyclopedia**: https://www.rp-photonics.com/encyclopedia.html
- **Light Conversion Optics Toolbox**: toolbox.lightcon.com
- **Newport Optics Technical Notes**: https://www.newport.com/resourceListing/technical-notes
- **Thorlabs Technical Resources**: https://www.thorlabs.com/navigation.cfm?Guide_ID=2400
- **Edmund Optics Technical Notes**: https://www.edmundoptics.com/knowledge-center/technical-literature/
- **Online Refractive Index Database**: https://refractiveindex.info/

Appendix C

Suppliers

Sometimes, you may not be in a research group with a lot of experience using optics and lasers, so knowing what suppliers exist for various items of equipment can be challenging. Here a list of commonly-used suppliers of optical components is provided for general reference[1]. This list is intended to be a useful starting point rather than be exhaustive, and a much larger directory of suppliers can be found at https://www.rp-photonics.com/bg_suppliers.html. I was raised in a lab where the saying was: 'if in doubt, trust the Lithuanians to make it properly'. That advice has never let me down.

- **Thorlabs**: www.thorlabs.com – very wide product range, which seems to keep getting wider. Generally very competitively priced. Best lab snacks.
- **Newport Optics**: www.newport.com – very wide product range, do some things Thorlabs don't do. Mediocre lab snacks.
- **Eksma Optics**: www.eksmaoptics.com – wide product range, including nonlinear and laser crystals. No lab snacks – but I did get a branded canvas tote bag last time I bought some mirrors.
- **Standa**: www.standa.lt – optomechanics galore. Very attractively priced for things like delay stages relative to Thorlabs/Newport.
- **Layertec**: www.layertec.de – very wide range of coated optics. Excellent quality but can be expensive – will make custom mirrors (and chirped mirrors).
- **Edmund Optics**: www.edmundoptics.co.uk – very wide product range, often seems to be on the expensive side.
- **Alphalas**: www.alphalas.com – laser crystals and polarisation optics.
- **4lasers**: www.4lasers.com – wide range of optics and crystals.

[1] With a few subjective opinions as of 2024 – your mileage may vary!

- **Crystran**: www.crystran.co.uk – wide range of windows, lenses, and prisms.
- **United Crystals**: www.unitedcrystals.com – nonlinear crystals from China, I hear good things about them..
- **Lattice Electro Optics**: www.latticeoptics.com – wide range of coated optics.
- **Ultrafast Innovations**: www.ultrafast-innovations.com – wide range of chirped mirrors.
- **Laser 2000 Photonics**: https://www.laser2000.co.uk/ – UK-based supplier of the excellent Semrock filters, among other things.
- **Redbourn Express**: https://www.redbournexpress.com/ – if you're in Europe and need a laser table installing, these are the guys.

In addition to these suppliers of smaller optical 'consumables', you might be in the market for bigger things:

- **Light Conversion**: https://www.lightcon.com/ – there's a reason these guys dominate the OPA market, and are increasingly dominant in Yb-based lasers too. Complete systems for ultrafast transient absorption spectroscopies too.
- **PhaseTech**: https://www.phasetechspectroscopy.com – pulse shapers, complete 2D-IR spectrometers, mid-IR detectors.
- **A.P.E Optics**: https://www.ape-berlin.de/en/ – all kinds of fun stuff. Nice single-box autocorrelators and FROGs.
- **Swamp Optics**: https://www.swampoptics.com/ – in addition to a load of useful resources, they also sell pulse measurement and compression devices. Cool logo.
- **Ultrafast Systems**: https://ultrafast.systems/ – mostly things for time-resolved spectroscopy, including complete systems and microscopes.

Appendix D

Frequency Mixing Detail

Taking as our input two plane waves with frequencies ω_1 and ω_2, our overall electric field can be written as:

$$E = E_1 e^{i\omega_1 t} + E_2 e^{i\omega_2 t} + E_1^* e^{-i\omega_1 t} + E_2^* e^{-i\omega_2 t} \tag{D.1}$$

Where the complex conjugate is explicitly included (in standard physics books this would generally be abbreviated by writing $+ c.c.$ after the first two terms).

The second order polarisation generated by this field in a material with susceptibility $\chi^{(2)}$ is:

$$P^{(2)} \propto \epsilon_0 \chi^{(2)} E^2 \tag{D.2}$$

Expanding E^2 produces ten terms:

$$P^{(2)} \propto E_1^2 e^{i2\omega_1 t} + E_2^2 e^{i2\omega_2 t} \tag{D.3}$$

$$E_1^{*2} e^{-i2\omega_1 t} + E_1^{*2} e^{-i2\omega_1 t} \tag{D.4}$$

$$2E_1 E_2 e^{i(\omega_2+\omega_1)t} + 2E_1^* E_2^* e^{-i(\omega_1+\omega_2)} \tag{D.5}$$

$$E_1 E_2^* e^{i(\omega_1-\omega_2)t} + E_2 E_1^* e^{-i(\omega_1-\omega_2)t} \tag{D.6}$$

$$E_1 E_1^* + E_2 E_2^* \tag{D.7}$$

The terms correspond to:

- Second harmonic generation (SHG) (D.3 and D.4).
- Sum-frequency generation (SFG) (D.5).
- Difference frequency generation (D.6).
- Optical rectification (D.7).

It is impossible to satisfy the phase-matching condition for all of these processes simultaneously, so the dominant process observed in a given situation depends on what has been

set up. A similar process can be followed to identify the possible processes following the mixing of three waves, but in that case there are 21 terms in the expansion! As it is good for the soul, the reader is encouraged to try this for themselves.

Appendix E

Convolution

Equation 5.2 described an **autocorrelation**, which we said was the 'convolution of the pulse with itself':

$$I_{\text{out}}(\tau) = \int_{-\infty}^{\infty} I(t)I(t-\tau)\mathrm{d}t$$

Stating *'it's the convolution of the pulse with itself'* and then quoting the above equation don't really explain the physical nature of what is happening here. So let's do that now, because convolution is a useful thing to understand. We'll use the autocorrelation example, but it applies generally too.

Generally a good way of trying to understand the physical 'thing' happening behind a complicated-looking integral is to remember that the integral is basically just a summation, but one where the gaps between the things we are adding up are infinitesimally small. So, if we rewrite this integral:

$$I_{\text{out}}(\tau) = \int_{-\infty}^{\infty} I(t)I(t-\tau)\mathrm{d}t \rightarrow \sum_{t=-\infty}^{\infty} I(t)I(t-\tau)$$

Written out more explicitly, with only the most minor abuses of notation:

$$I_{\text{out}}(\tau) \approx I(t_0)I(t_0-\tau) + I(t_1)I(t_1-\tau) + ... + I(t_n)I(t_n-\tau)$$

Where t_0 is the earliest time we measure at and t_n is the latest time we measure at. So basically, what we do is take the intensity profiles of our two pulses at a certain delay τ, and measure their product at every single time point possible (from $-\infty$ to ∞). In the autocorrelation example, we repeat this measurement for many different relative delays τ, so that we build up a picture of $I_{\text{out}}(\tau)$.

Ok, but this still isn't hugely clear. Let's imagine what happens in some limiting cases. Firstly, If the pulses are very far apart in time, so $t_0 - \tau >> t_0$, then we will have no point between t_0 and t_n where both of the pulses have a non-zero intensity (because they never overlap). Every term in the summation above is thus zero, and $I_{\text{out}} = 0$.

Now imagine a case where $\tau = 0$, so there is no delay between the two pulses. In this case, any time point where our pulse has non-zero intensity will contribute to the summation above, and we will have I_{out} being non-zero and a maximum.

Finally, imagine that there is a small delay between the two pulses, one that is less than the width of the pulse in time. in this case, there will be some time points in our summation where one of our pulses has non-zero intensity, but the other has zero intensity. Thus, fewer time points in the summation above will contribute to I_{out} than in the previous case, making it smaller but still non-zero. The combination of all of these measurements at different values of τ gives you the overall convolution $I_{\text{out}}(\tau)$. Another way to think of the convolution is as 'fixing' one function and 'scanning' the other through it, measuring the product of the two functions as you go. There are some nice animations of it on Wikipedia, among other places.