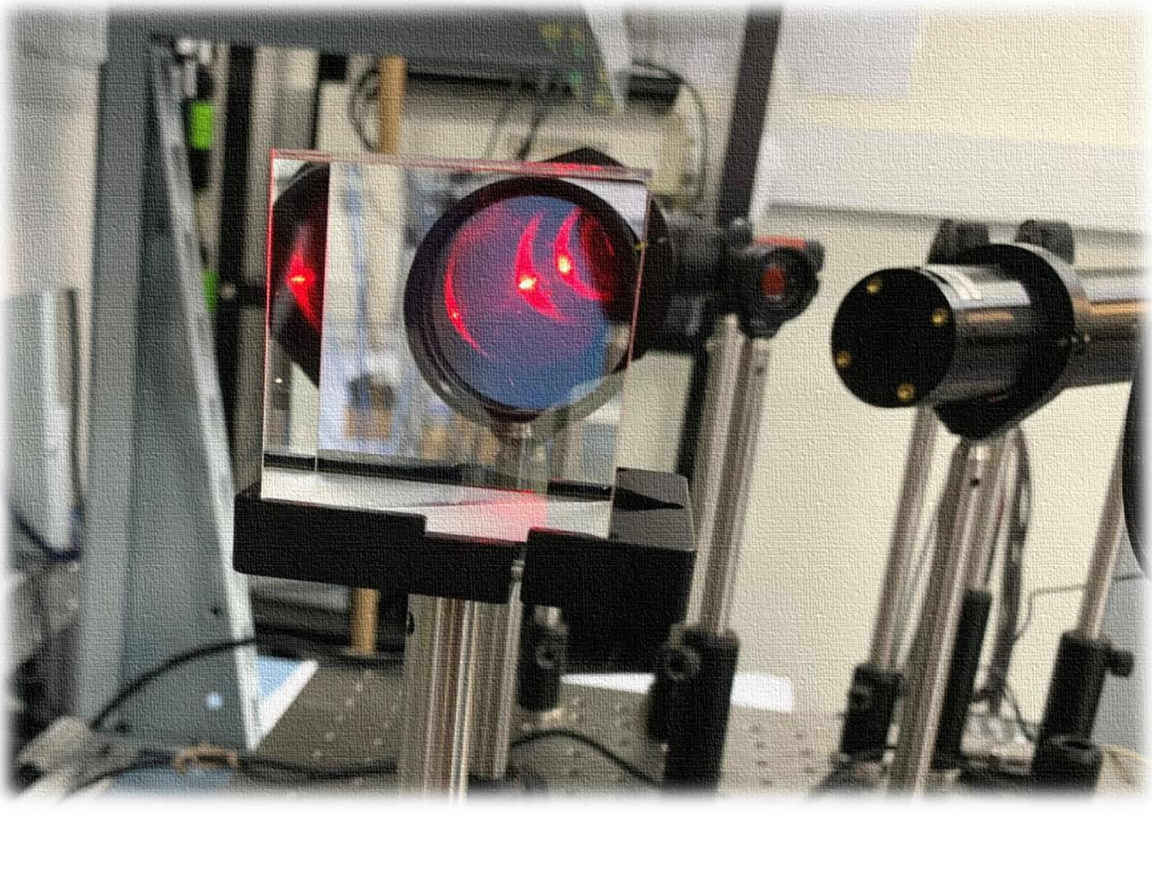
Optics for Telecommunications

Theory and Engineering

Ian G. Clarke



Contact Details:

Email: iclarke@ieee.org

Mobile: 0451 849 125

Contents

Introduction

These notes are designed to give you (almost) everything you need to know to get started in modern optical research, design and manufacturing. They are part of a 50-hour course that I've taught in various forms at the University of Sydney, Galileo Corporation, IPG Photonics, Finisar Australia and Finisar China. When you find mistakes, please let me know (iclarke@ieee.org).

In order to have a variety of material and activities each week, each chapter mixes about an equal amount of mathematical theory and engineering knowledge.

In optics there are some arbitrary choices each author makes that changes what the mathematics looks like. Below I've listed all those choices and some of the key equations and constants. It would be great if these where consistent between books, Wikipedia articles and research papers, but they are not – so I have tried to follow the most common choices.

1. $\boldsymbol{E}(z,t) = \boldsymbol{E}_o(kz - \omega t)$ *is the standard form of the wave equation we use. Note that some books use* $\boldsymbol{E}_o(-kz + \omega t)$.

2. *Right circular polarization is defined by the electric field vector, at a fixed point in space, moving in a clockwise direction when looking back towards the source. This is the more common definition for optics texts.*

Chapter 1. Fundamental Concepts and Connectors

So we don't lose anyone at the start, I thought it would be good to briefly revise some mathematics this week. If this is not revision, but all new, then let me know and we'll take some time to go through it in detail. However, we'll start with a brief overview of the zoo of connectors available to join fibers.

1.1. Fiber Connectors

If you're not fusion splicing fibers together then you will need fiber connectors. A single-mode connector positions the fiber core to an accuracy of around 1 micron (a remarkable achievement). All common single fiber connectors introduce a loss of between 0.05 and 0.2 dB (about 4% loss). If you *really* need less loss than 0.2 dB then buy a connector from Diamond Fiber Optics (however their prices are as high as the name would suggest). By the way, one of the most amateur things you can do in a fiber optics lab is to leave the caps off your connectors – you're almost guaranteed to get the end faces contaminated.

1.1.1. Single Fiber Connectors

In laboratories the most common connector is the FC. It's robust and so is great for being connected and disconnected hundreds of times.

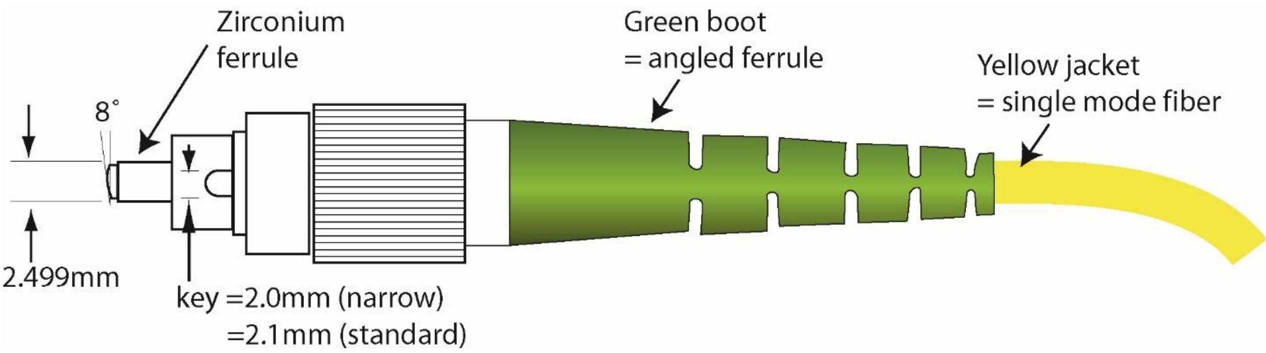


Figure 1 FC/APC connector

Within the connector, the fiber sits in the center of an extremely hard zirconium ferrule. The FC/APC has an 8° angle on the (hemispherical) end of the ferrule. The hemispherical end should mean that the fibers physically press up against each other ("APC" means angled physical contact and is indicated by a green boot).

You will also see connectors with a black boot. These connectors have a non-angled hemispherical end on the ferrule and are usually called FC/PC, FC/SPC or FC/UPC connectors (the hemispherical ends ensure that the fibers actually touch hence the name "Physical Contact" or "PC". The initial coarse polishing of the ferrule creates a region of compacted glass. If the polishing process is carefully designed, then the fine polishing can remove this compacted glass giving the connector less than 50 dB back-reflection (one hundred-thousandth of the light going back along the fiber). A less careful polish (SPC) will give < 40 dB, but FC/APC connectors can achieve <60 dB back reflection (one millionth of the light is reflected). By the way, an FC/PC connector with a blue boot (and usually a blue cable) usually has polarization preserving fiber in it – we learn about that later.

1.1.2. Bulkhead Adaptors

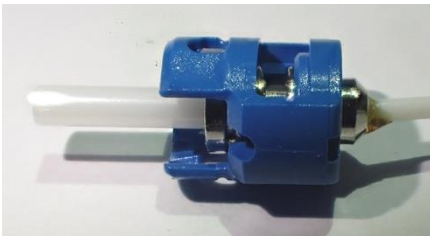
Each connector comes with its own bulkhead adaptor. These consist of a split-ring sleeve in a housing. Good sleeves are made of zirconia. Cheap ones are made of phosphor bronze (oxide dust from phosphor bronze can contaminate the connector). `

1.1.3. LC Connectors and SC connectors

Telecom companies don't like FC/PC connectors because they cannot be placed close together, since you need to get your fingers around them to undo them. Instead, they first moved to SC connectors. These have the same 2.5mm ferrule as FC connectors but have a simple push-pull engagement. The blue connectors have a flat end (SC/PC), the green connectors have an angled end (SC/APC). Grey SC connectors are designed for multimode fiber.

Figure 2 LC/PC Connector (note that a green body would indicate a LC/APC connector

In the quest to fit more connectors onto patch-panels Lucent Technologies introduced the LC family of connectors. These have a zirconium ferrule that has half the diameter, and the whole connector is half the size of the SC. The vast majority of our telecom customers use either LC/PC or LC/APC connectors. You can recognize the LC/APC connectors because they are green rather than blue.



There is an interesting variant of the LC connector which is the Micro LC bayonet style connector. This saves about 30mm of space behind the panel. However, it come with its own challenges. The zirconium alignment sleeve in the bulkhead adaptor is not well confined to the adaptor but instead comes out when the connector is removed. The short length of the bulkhead adaptor also seems to cause more problems for manufacturer (especially if they make these in small volume), but if you really need to save space, use them.

Multicolored keyed LC-connectors can be used to prevent the wrong patch-cords being inserted (and to add a little to network security).

A new connector type, the mSFP (or "mini-LC") is 1 mm thinner, but otherwise looks identical (definitely annoying).

1.1.4. Diamond Connectors

1.1.4.1. E2000

For high power applications you can't beat Diamond's E-2000 connector with a metal shutter. The shutter closes as soon as connector is detached. These are also great connectors for environments where connectors may be handled by people unfamiliar with optical fibers since it is

Figure 3 E2000 Connector (also available with metal cap) from Diamond Fiber Optics

difficult for them to accidently touch the end of the fiber. It's also more expensive and less common than either FC, SC or LC connectors (but I wouldn't use anything else than an E2000 with a metal cap with high power fiber lasers or high-power optical amplifiers.

1.1.4.2. Other Diamond connectors

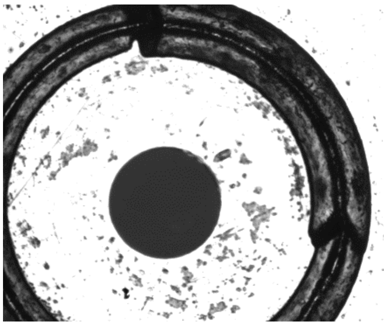


Figure 4 Diamond connector end with soft metal inner ferrule and tool mark from the centring process

Diamond also makes FC connectors but these don't have a traditional solid zirconia ferrule but have a zirconia sleeve with a soft metal inner sleeve, and a hard metal sleeve inside that. They then use a tool to indent the soft metal to center the fiber. This generally gives better than 0.1dB loss (2%), but they are not cheap, and some people say they scuff up other connectors.

1.2. Ribbon Connectors

1.2.1. MPO and MPT Connectors

Figure 5 Plastic shavings on an MPO connector from repeated matings

MPO connectors (Multi-fiber Push On connectors) allow you to join 12, 24, 36 or even 48 fibers at a time. Note that these come in a male and female version with a pair of alignment pins and alignment holes respectively. Sometimes you'll hear them called MTP connectors, but they are interchangeable (MTP connectors are allegedly better engineered). Typically, these connectors introduce about 0.4 dB of extra loss. However, these connectors have lots of problems and the number of matings should be kept to an absolute minimum. If you can avoid using them, do. You will save a lot of pain. If multiple connections are required, I find it best to attach a breakout cable made of LC or FC connectors to minimize connecting and disconnecting the MPO connector.

MPO ribbon connectors have a hard plastic ferrule (liquid crystal polymer) and a softer plastic bulkhead adapter, so as the connector is inserted, plastic shavings can be created as the ferrule rubs against the bulkhead. These shavings can deposit on the connector face. Because the plastic ferrule is softer than the glass fiber, the polishing process will cause the fibers to protrude slightly so the shavings and other material tends to get caught on the fibers. When cleaning these connector wipe across the connector rather than along it so as to move debris off the connector rather than simply onto the next fiber!

Because these connectors can self-contaminate I've found it often essential to ship devices using these connectors with photos of the clean connectors (printed and in the box), otherwise they are returned as contaminated (and it's a rare customer that will believe that they contaminated the connector by simply plugging it in!)

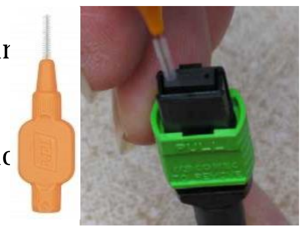


Figure 6 Cleaning an alignment hole

If repeated matings continue to give high losses then you may need to clear the alignment holes. Here's how.

Step 1. Moisten the 0.45mm intra-dental brush with isopropyl alcohol (– not acetone!).

Step 2. Gently insert the brush into each alignment hole and rotate 90 degrees (see figure 2)

Step 3. Gently remove the brush. Inspect the top of the hole for anything removed from the alignment hole.

1.2.2. HBMT Connectors

Another, much less common, ribbon connector is the HBMT (High-density Backplane MT connector). They have the same ferrule as the MPO connectors but a latch similar to the LC connector. These are specifically designed to connect from a PC board into a backplane interface. Sets of four connectors are housed in a block allowing up to 96 fibers to be connected automatically as the card slides into place. Before being placed in their outer plastic housing, the fibers are very easily snapped.

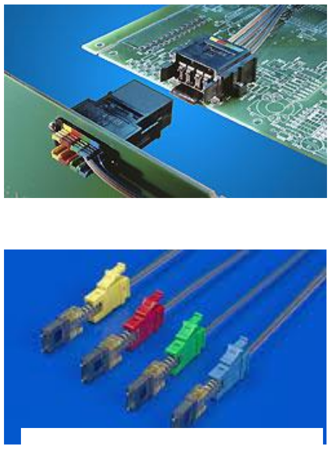


Figure 7 HBMT connectors

1.2.3. 3M EBO Ribbon Connector

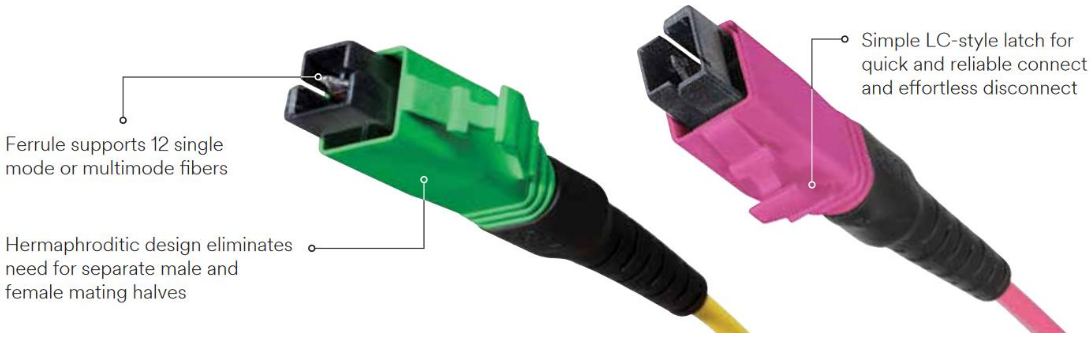
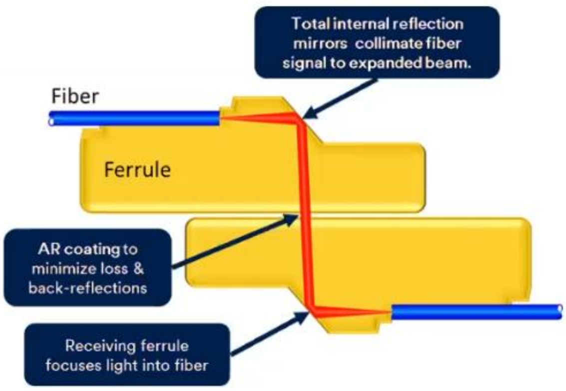


Figure 8 Connector picture and diagram from 3M presentation

A new family of connectors from 3M may well kill off the MPO ribbon connector (it can’t happen soon enough in my opinion).

The EBO connector uses a clever expanded beam connection idea:



The beam is large at the interface making it much less sensitive to contamination. The ferrules don't touch each other so shouldn't damage each other in the presence of dust and, in practice, a thousand mates caused no change in loss during testing. Typical loss was around 0.4 dB (spec < 0.7dB) which is a little higher than we would like.

1.3. What Goes Wrong with Single Fiber Connectors

You may spend more time troubleshooting connectors than any other component. Here's my top few things to look out for:

1.3.1. Contaminated Fiber (Solids)

Always use a connector microscope to check your fiber ends. Most contamination you can clean off with either a cloth connector cleaner cassette or a clean wipe with isopropyl followed by a dry wipe

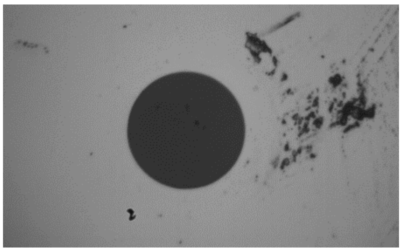


Figure 9 Soilid residue

1.3.2. Contaminated Fiber (Liquid residue)

Often attempts to clean fiber ends with isopropyl will leave rings of liquid residue. If the fibers have been mated when still wet, this will form a characteristic ring around the centre of the connector.

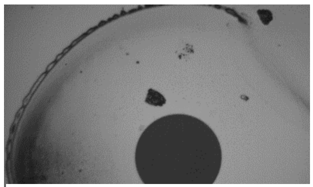


Figure 10 Dried liquid residue

1.3.3. Burnt/exploded Fiber

Contamination followed by lots of light (~100 mW) will burn the fiber end exploding off glass at the fiber core. This can't be fixed without a repolisher (Krell sells one). If you're not repolishing, please throw the connector away so no one else uses it. Powers over 100 mW can easily cause this damage.

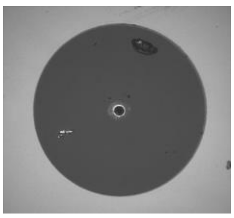


Figure 12 Burnt fiber end (from www.Thorlabs.com)

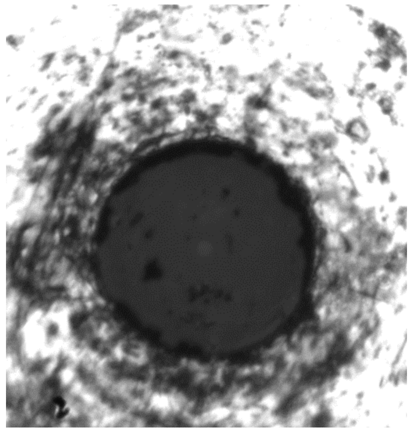


Figure 11 Scratches and pits (scratches in the soft metal of a diamond connector are fine, but notnear the fiber core

- **Scratched end face:** Scratches that don't go near the core of the fiber generally aren't too critical but any scratches on, or near the fiber core are a big problem.
- **Connecting angled to flat connectors:** This automatically creates a 4dB loss (a bit over half the light gets lost). A green boot to a black boot is obvious, but sometimes a bulkhead adaptor in a panel is far from obviously angled or

flat. If it's not labelled it's probably flat. You are unlikely to damage a flat bulkhead by plugging in an angled fiber or versa versa so try and see. If the loss improves by 4dB you've solved your problem.

- **Standard key connector in a narrow key bulkhead**

This is only an FC connector problem. For crazy historical reasons these connectors come in a narrow key (2.0 mm wide) and a standard key configuration (2.1 mm wide) as shown in Figure 1. A narrow key connector with fit in a wide-key bulkhead adaptor, but a wide-key connector will not fit in a narrow key adaptor! Most of the connectors you will find lying around the lab are narrow key. The only way to tell this is to get out digital calipers and measure the key width. If the key on your connector is 2.1mm then check that the bulkhead adaptor doesn't have a slot width of less than 2.1mm (narrow is about 2.05mm).

- **Cracked fiber inside the ferrule**. If you shine a red laser into the ferrule and see a lot of light spill out then it may be that the fiber is cracked inside the ferrule, this is a manufacturing fault that sometimes shows up in cheap connectors.

1.4. Cleaning Connectors

Don't clean connectors until you know they are dirty (use a fiber microscope). If you see contamination on the ferrule you can either use a cleaning cassette or a cotton bud (Q-tip) dipped in isopropyl followed by a dry Q-tip). Contaminated bulkheads can be cleaned using a "One-click" cleaner from Fujikura or similar.

1.5. Non-contact connectors

For continuity tests (a quick and dirty "is it still working?" test, it's common to use a non-contact LC connector. These have a distinctive red boot. When you mate a normal connector to a non-contact connector there is a gap between the fibers that creates a loss that varies sinusoidally with wavelength (we'll learn why later). It is not ideal, but it does stop your fiber end getting contaminated just before you ship your device.

1.6. Background mathematics

1.6.1. Unit Vectors

A vector is simply a number that has a direction. Can you think of some examples? Rather than write out vectors it's often easier to use vectors of length 1 in the x, y and z directions. These are traditionally called $\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}$ and $\hat{\boldsymbol{k}}$. $\hat{\boldsymbol{\imath}} = (1,0,0)$, $\hat{\boldsymbol{\jmath}} = (0,1,0)$ and $\hat{\boldsymbol{k}} = (0,0,1)$ so the vector

$$(5,3,2) = 5\hat{\boldsymbol{\imath}} + 3\hat{\boldsymbol{\jmath}} + 2\hat{\boldsymbol{k}}$$

1.6.2. Vector Field

We can represent the water velocity in a pool of water, or an electric field, or a magnetic field in a volume as a function

$$\boldsymbol{F}(x,y,z) = \left(F_x(x,y,z), F_y(x,y,z), F_z(x,y,z)\right) = F_x(x,y,z)\,\hat{\boldsymbol{\imath}} + F_y(x,y,z)\hat{\boldsymbol{\jmath}} + F_z(x,y,z)\hat{\boldsymbol{k}}$$

1.6.3. Cross product

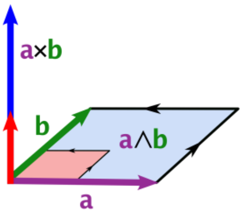


Figure 13 Cross product (from Wikipedia)

A cross-product of vectors **a** and **b** has magnitude $ab\ sin(\theta)$ and a direction orthogonal to **a** and **b.**

$$\left(a_x, a_y, a_z\right) \times \left(b_x, b_y, b_z\right) = (a_y.b_z - a_z.b_y, a_z.b_x - a_x.b_z, a_x.b_y - a_y.b_x)$$

We can write this as the determinate of the matrix

$$\begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix}$$

Cross products are important when two quantities interact only to the extent that they are orthogonal. Torque (or twisting force) is a good example of this. The torque produced by a force ($\boldsymbol{F}$) acting a distance ($\boldsymbol{r}$) from the axial is $\boldsymbol{\tau} = \boldsymbol{r} \times \boldsymbol{F}$**.** Try For example calculating the torque that can be applied by a pusher to a glue joint where the position of the pusher is at (0.02,0.02,0)m relative to the joint and applying a force of (200,-200,0)N (and sketch the vectors).

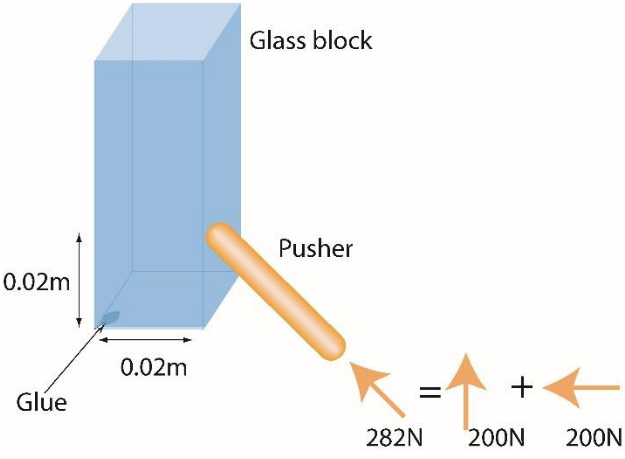


Figure 14 An example of a cross-product calculation

We also use the cross product when we calculate the "Poynting Vector". The rate of energy flow due to an electromagnetic wave.

$$\boldsymbol{S} = \boldsymbol{E} \times \boldsymbol{H}$$

Where $\boldsymbol{E}$ is the electric field and $\boldsymbol{H}$ is the magnetic field. If $\boldsymbol{E} = \boldsymbol{E_o}\cos(\boldsymbol{k}.\boldsymbol{r} - wt)$ and $\boldsymbol{H} = \boldsymbol{H_o}\cos(\boldsymbol{k}.\boldsymbol{r} - wt)$ and $\boldsymbol{E_o}$ and $\boldsymbol{H_o}$ are orthogonal then the Poynting Vector

$$S = \boldsymbol{E_o} \times \boldsymbol{H_o}\cos^2(\boldsymbol{k}.\boldsymbol{r} - \omega t)$$

And the average value of the Poynting Vector

$$\langle S \rangle = \tfrac{1}{2}\boldsymbol{E_o} \times \boldsymbol{H_o} = \tfrac{1}{2}E_oH_o$$

1.6.4. Dot product

A dot product of vectors **a** and **b** has magnitude $ab\ cos(\theta)$ and is a scalar. Dot products are useful when quantities interact only to the extent they are in the same direction

$$(a_x, a_y, a_z) \cdot (b_x, b_y, b_z) = a_x . b_x + a_y . b_y + a_z . b_z$$

For example, the amount of work done by a force is $\boldsymbol{F.d}$ where $\boldsymbol{d}$ is the distance moved and $\boldsymbol{F}$ is the force.

1.6.5. Curl

Imagine a three dimensions fluid. Place a tiny ball at a point in the fluid. The magnitude of the curl of the fluid velocity is simply twice the speed the ball spins (in radians/sec). The direction of the curl is the axis it spins around (an axis has two directions, so we choice the one that it spins around in a right-handed sense). Mathematically this can be written:

$$\nabla \times \boldsymbol{f}(x, y, z) = \left(\frac{\partial f_z}{\partial y} - \frac{\partial f_y}{\partial z}\right)\hat{\boldsymbol{\imath}} + \left(\frac{\partial f_x}{\partial z} - \frac{\partial f_z}{\partial x}\right)\hat{\boldsymbol{\jmath}} + \left(\frac{\partial f_y}{\partial x} - \frac{\partial f_x}{\partial y}\right)\widehat{\boldsymbol{k}}$$

1.6.6. Divergence

Once again, imagine our three dimensions fluid. Imagine a little cube of the fluid. The divergence of the velocity of the liquid is the rate at which the fluid is getting bigger. Mathematically this can be written

$$\nabla \cdot \boldsymbol{f}(x, y, z) = \frac{\partial f_x}{\partial x} + \frac{\partial f_y}{\partial y} + \frac{\partial f_z}{\partial z}$$

1.6.7. Laplacian

The Laplacian is a measure of the curvature of a function at a given point

$$\nabla^2 f(x, y, z) = \left[\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}\right]$$

1.6.8. Vector Laplacian

The “Vector Laplacian” operator ∇^2 is a three-dimensional measure of how much the gradient is changing as we look across our material. It can be written

$$\nabla^2 \boldsymbol{f}(x, y, z) = \left[\frac{\partial^2 f_x}{\partial x^2} + \frac{\partial^2 f_x}{\partial y^2} + \frac{\partial^2 f_x}{\partial z^2}\right]\hat{\boldsymbol{\imath}} + \left[\frac{\partial^2 f_y}{\partial x^2} + \frac{\partial^2 f_y}{\partial y^2} + \frac{\partial^2 f_y}{\partial z^2}\right]\hat{\boldsymbol{\jmath}} + \left[\frac{\partial^2 f_z}{\partial x^2} + \frac{\partial^2 f_z}{\partial y^2} + \frac{\partial^2 f_z}{\partial z^2}\right]\widehat{\boldsymbol{k}}$$

It’s a measure of how much the material has been distorted.

1.7. Questions and Practical

1.7.1. Practical 1

1.7.1.1. Connector Cleaning

1.

    a. Take the connector, check that it's clean using the connector microscope.

    b. Put a fingerprint on the end of the connector.

    c. View it again – make sure you can see the fingerprint.

    d. Clean the connector using the cloth cleaner

    e. View again and check it is clean

1.7.1.2. FC Connectors

2. Use the Calipers to measure the key width on the two connectors. Which of "a" and "b" are narrow or broad key connectors?

    a. ……..

    b. ………

1.7.1.3. Troubleshooting (difficult)

3. Examine Sample 1. No light is coming out of this connector join. Find the two issues (you can use the tools provided)

4. Example sample 2. For fiber 12, no light is being transmitted, Using the red-light illuminator, figure out the source of the problem. Please handle carefully.

1.7.2. Theory

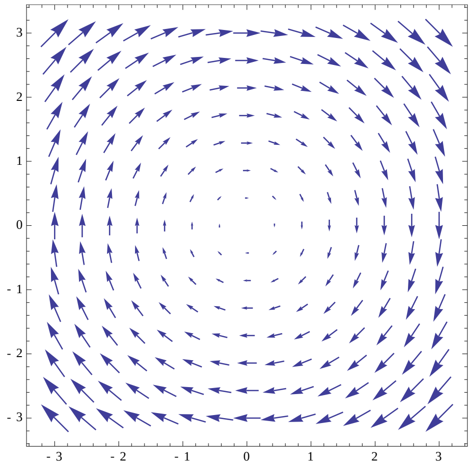
5. The phase of a wave is ***k.r*** (the dot product) where k is the propagation vector (we'll learn about this next week) and r is the position. Consider a 1550nm wave where the propagation vector is $(3.2 \times 10^6, 2.4 \times 10^6, 0)$ radians/metre with zero phase at $r = 0$. What is the phase at $(5 \times 10^{-6}, 5 \times 10^{-6}, 1 \times 10^{-5})$?

6. Maxwell's equations tell us the relationship between electric and magnetic fields. The first equation is

   $\nabla \times \mathbf{E} = -\mu \frac{d\mathbf{H}}{dt}$ in free space. If the electric field is

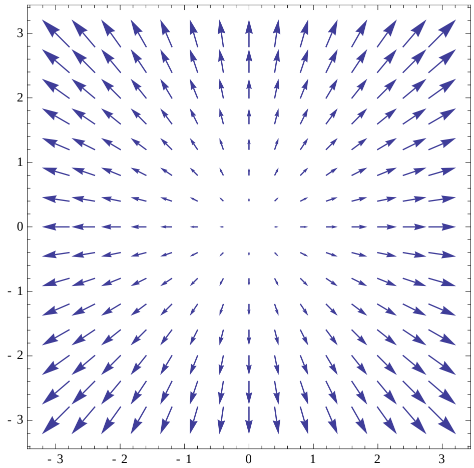
$$\mathbf{E} = (E_o \sin(\omega t - kz), 0, 0)$$

   where E_o, ω and k are just constants, what will **H** be?



7. a. The plot beside this question is the function $\boldsymbol{F}(x, y, z) = (y, -x, 0)$. What is the curl of this function?

   b. What is the divergence of this function?



8.

   i. This plot is the function $\boldsymbol{G}(x, y, z) = (x, y, 0)$. What is the divergence of this function?

ii. What is the curl of $\boldsymbol{G}$?

c. (bonus question) Now, starting with the value of **H** you calculated in question 6, use the next of Maxwell's equations

$$\boldsymbol{\nabla} \times \mathbf{H} = \epsilon \frac{\mathrm{d}\mathbf{E}}{\mathrm{dt}}$$

to work out what electric field **E** will be created by the magnetic field.

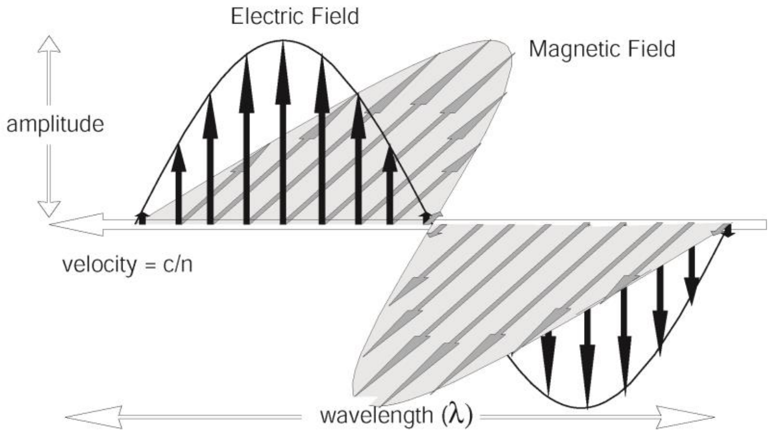
(bonus question) Use the result from question 4 and compare it to the electric field you started with in question 3. Can you use this to figure out a relationship between ω, k, ϵ and μ?

The velocity of a wave is simply $\frac{\omega}{k}$. Can you work out the velocity of the light wave in terms of μ and ϵ?

Chapter 2. Wave Equation, Wavelength and Frequency

We're going to talk a lot about waves in this chapter, so let's revise a few concepts.

The frequency (*f*) of a wave is simply how many waves go past each second. The angular frequency ($\omega = 2\pi f$) is the phase (in radians) that goes past a fixed point in a second.



The wave number (k) is usually how much wave phase fits in a metre ($k = \frac{2\pi}{\lambda}$). This is sometimes called the "angular wave number" because chemists use "wave number" to mean how many waves fit in a centimetre (and measure it in cm^{-1}). In modern telecommunications we mostly use the C-band, consisting of wavelengths from about 1530-1565 nm, although some companies are now using the Super-C band or extended C band (1524.50 – 1571.65 nm). In a vacuum, light travels at $c = 299\ 792\ 458$ m/s. When light travels through a material it will slow down (later we will find out why), by an amount "n" which we call the refractive index. In general the velocity of light

$$v = \frac{c}{n}$$

The velocity $f\lambda = \frac{\omega}{k} = v$. In a vacuum $f\lambda_o = \frac{\omega}{k_o} = c$ (where λ_o is the wavelength in a vacuum).

It's important that you can quickly swap between a wavelength range and a frequency range. If you need to do a very rough approximation, just use the rule of thumb that in the C-band 0.8 nm (or 800 pm) is approximately 100 GHz. So, for example 200 pm is about 25 GHz, or 1 pm is about $\frac{1}{8}$ GHz. Often, however, you need to be precise:

$\frac{df}{d\lambda} = -\frac{c}{\lambda^2}$ so $\delta f \approx -\frac{c}{\lambda^2}\delta\lambda$ or to convert the other way $\delta\lambda \approx -\frac{c}{f^2}\delta f$ so at the short wavelength end of the c-band (at 196 THz) a 50GHz channel is $\frac{2.99792458 \times 10^8\ \mathrm{ms^{-1}}}{(196\times 10^{12})^2\ \mathrm{m^2}} . 50 \times 10^9\ \mathrm{s^{-1}} = 390\ \mathrm{pm}$.

2.1. Wave Equation

Maxwell realized that the various equations describing electric and magnetic fields ("Maxwell's Equations") could be rearranged to make a wave equation.

Maxwell's equations for an isotropic, non-conducting material are

$$\nabla \times \boldsymbol{E} = -\mu.\frac{\partial \boldsymbol{H}}{\partial t} \qquad (2.1)$$

$$\nabla \times \boldsymbol{H} = \epsilon.\frac{\partial \boldsymbol{E}}{\partial t} \qquad (2.2)$$

$$\nabla.\boldsymbol{E} = 0 \qquad (2.3)$$

$$\nabla.\boldsymbol{H} = 0 \qquad (2.4)$$

We can use the vector identity $\nabla \times (\nabla \times \boldsymbol{F}) = \nabla(\nabla.\boldsymbol{F}) - \nabla^2\boldsymbol{F}$ to derive the wave equation

$\nabla^2\boldsymbol{E}(x,y,z) = \mu\epsilon\frac{\partial^2\boldsymbol{E}}{\partial t^2}$ and $\nabla^2\boldsymbol{H}(x,y,z) = \mu\epsilon\frac{\partial^2\boldsymbol{H}}{\partial t^2}$

where $\boldsymbol{E}$ is the electric field and $\boldsymbol{H}$ is the magnetic field.

In one dimension, this reduces to

$\frac{\partial^2\boldsymbol{E}}{\partial z^2} = \mu\epsilon\frac{\partial^2\boldsymbol{E}}{\partial t^2}$

And this has a solution of the form

$\boldsymbol{E}(z,t) = \boldsymbol{E}_o(kz - \omega t)$

Where $\frac{\omega}{k} = \frac{1}{\sqrt{\mu\epsilon}} = \frac{1}{n\sqrt{\mu_o\epsilon_o}} = \frac{c}{n}$

The ratio $\frac{\omega}{k}$ gives us the velocity of the wave (can you explain why?). We can measure the electrical constants μ_o and ϵ_o, to determine the speed of light in a vacuum. Note that many textbook use $\boldsymbol{E}_o(\omega t - kz)$ which will result in different signs in many equations.

Consider this plane wave travelling in z direction (i.e. uniform in the x and y directions).

If it is uniform then $\frac{\partial E_x}{\partial x} = \frac{\partial E_y}{\partial y} = 0$. From Maxwell's equations (assuming there is no charges present) $\nabla.\boldsymbol{E} = 0$, so $\frac{\partial E_z}{\partial z}$=0 and we can ignore any constant field and set $E_z = 0$ and similarly $H_z = 0$. In other words, a plane wave moves in a direction perpendicular to both its electric and magnetic fields.

One solution to the wave equation is a simple sinusoidal plane wave with an arbitrary phase θ.

$\boldsymbol{E}(z,t) = \boldsymbol{E}_o\cos(kz - \omega t + \theta) = Re[\boldsymbol{E}_o e^{i(\mathrm{kz-\omega t+\theta})}]$

This is very useful for two reasons. The first is that we can build any other (continuous) function out of sinusoids. The other reason is that now we can replace all our $\frac{\partial}{\partial t}$ terms with $-i\omega$ if we use the $\boldsymbol{E}_o e^{i(\mathrm{kz}-\omega\mathrm{t}+\theta)}$ form.

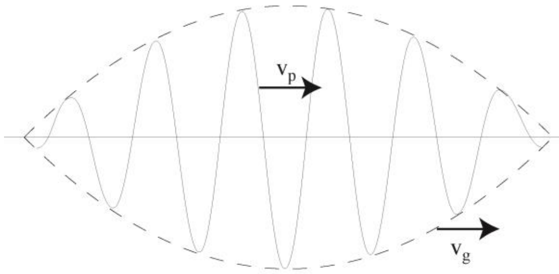
Using a sinusoid, ω is the angular frequency (measured in radians/s), $\omega = 2\pi f$ where f is the frequency of the wave.

"k"is the wavenumber. In terms of the wavelength $k = \frac{2\pi}{\lambda}$.

The wavelength is

$$\lambda = uT = \frac{2\pi}{k}$$

2.2. Group Velocity



We can calculate the phase velocity v_p of the wave very simply. To stay still on the wave ("surf the wave") we need

$$kz - \omega t = 0$$

i.e.

$$u_p = \frac{z}{t} = \frac{\omega}{k}$$

The group velocity is different. Take the case of two beating waves (waves with slightly different frequencies and wavelengths). This is the simplest example of a wave with an envelope.

$$E = \frac{1}{2} E_o e^{i[(k+\delta k)z-(\omega+\delta\omega)t]} + \frac{1}{2} E_o e^{i[(k-\delta k)z-(\omega-\delta\omega)t]}$$
$$E = \frac{1}{2} E_o e^{[ikz-\omega t]} \left(e^{i[\delta k.z-\delta\omega.t]} + e^{-i[\delta k.z-\delta\omega.t]} \right)$$
$$E = E_o e^{[ikz-\omega t]} \cos(\delta k.z - \delta\omega.t)$$

The " $E_o e^{[ikz-\omega t]}$" term is the wave, and the "$\cos(\delta k.z - \delta\omega.t)$ " is the envelope. The envelope has velocity $u_g = \frac{\delta\omega}{\delta k}$ or in the limit $\delta k \to 0$, the group velocity

$$u_g = \frac{d\omega}{dk}$$

This is important for us because all data signals (i.e. information) travel at the speed of the group velocity.

We can rearrange our equations to show

$\frac{1}{u_g} = \frac{dk}{d\omega} = \frac{d}{d\omega}\left(\frac{2\pi}{\lambda}\right) = \frac{d}{d\omega}\left(\frac{2\pi n}{\lambda_o}\right) = \frac{d\lambda_o}{d\omega}.\frac{d}{d\lambda_o}\left(\frac{2\pi n}{\lambda_o}\right) = \frac{d}{d\omega}\left(\frac{c.2\pi}{\omega}\right).\frac{d}{d\lambda_o}\left(\frac{2\pi n}{\lambda_o}\right)$

Where λ_o is the wavelength in a vacuum, and c is the speed of light in a vacuum. Expanding this out ($\frac{d}{d\omega}\left(\frac{c.2\pi}{\omega}\right).\frac{d}{d\lambda_o}\left(\frac{2\pi n}{\lambda_o}\right) = -\frac{c.2\pi}{\omega^2}\left(-\frac{2\pi n}{{\lambda_o}^2} + \frac{2\pi n}{\lambda_o}\frac{dn}{d\lambda_o}\right) = \frac{cn}{f^2{\lambda_o}^2} - \frac{cn}{f^2{\lambda_o}^2}\frac{dn}{d\lambda_o}$) gives (when we know $c = f\lambda_o$)

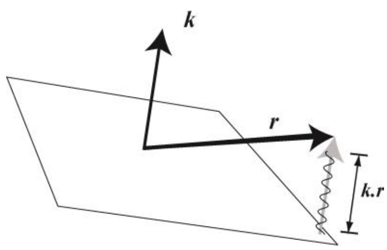
$$\frac{1}{u_g} = \frac{1}{u} - \frac{\lambda_o}{c}.\frac{dn}{d\lambda_o}$$

In words, this means that if the refractive index is changing with wavelength, then the group velocity will be different than the phase velocity. However if the *rate* at which the group velocity changes with wavelength, itself changes then the group velocity will not be constant with wavelength and dispersion will occur (pulses will spread out). This is why we measure group delay ripple in our products.

2.3. Wave Equation in three dimensions

In three dimensions, we can write the magnitude of a plane wave as

$U(x, y, z, t) = U_o\cos(\boldsymbol{k}.\boldsymbol{r} - \omega t) = U_o e^{i(\boldsymbol{k}.\boldsymbol{r} - \omega t)}$



Where $\boldsymbol{k}$ is a vector in the direction of the wave, with magnitude k and $\boldsymbol{r}$ is the position we are considering (this has nothing to do with the unit vector in the z direction: $\widehat{\boldsymbol{k}}$). You can think of $\boldsymbol{k}.\boldsymbol{r} - \omega t$ as a measure of where we are on the wave; how many radians forward you've gone spatial $\boldsymbol{k}.\boldsymbol{r}$ minus how many waves have passed you (ωt). If you are staying still relative to the wave (like a surfer on a water wave) then $\boldsymbol{k}.\boldsymbol{r} - \omega t$ is constant (the "phase" is constant).

In this form, the del operator can be replaced with $\boldsymbol{k}$, and $\frac{\partial}{\partial t}$ can be replaced with ω

$$\boldsymbol{\nabla} = \hat{\boldsymbol{\imath}}\frac{\partial}{\partial x} + \hat{\boldsymbol{\jmath}}\frac{\partial}{\partial y} + \widehat{\boldsymbol{k}}\frac{\partial}{\partial x} \rightarrow \boldsymbol{k}$$

$$\frac{\partial}{\partial t} \rightarrow -i\omega$$

Now Maxwell's equations then become

$$\boldsymbol{k} \times \boldsymbol{E} = -\mu\omega\boldsymbol{H}$$

$$\boldsymbol{k} \times \boldsymbol{H} = -\epsilon\omega\boldsymbol{E}$$

$$\boldsymbol{k}.\boldsymbol{E} = 0$$

$$\boldsymbol{k}.\boldsymbol{H} = 0$$

This tells us that $\boldsymbol{E}, \boldsymbol{H}$ and $\boldsymbol{k}$ are orthogonal in a uniform medium.

The energy flow per unit area is $\boldsymbol{S} = \boldsymbol{E} \times \boldsymbol{H}$. It's fairly easy to show that the magnitude of $\boldsymbol{S}$ (which we call the intensity of the light), $I = \frac{1}{2}E_oH_o = \frac{1}{2}\left(\sqrt{\frac{\mu_o}{\epsilon_o}}\right)nE_o^2$.

2.4. Phase

Where you are on a wave is called the phase (in degrees it's between 0 and 360, in radians, it's between 0 and 2π). If you have a local laser at the same frequency, you can compare the phase of the local laser with a wave coming in. We can then do phase encoding. Rather than turn a laser on and off to send information, we can change the phase of the light . QPS (Quadrature Phase Shifting) means jumping between four possible phase values (0, π/2, π, 3π/2).

This lets us send 2 bits of information on each "symbol". Rather than "on"=1 and "off"=0, we have "zero phase=00, "π/2 phase"=01, "π phase"=10, "3π/2 phase"=11.

2.5. Photons

One of the interesting things to come out of quantum mechanics is that the energy in light comes in lumps, called photons. Each photon has energy

$$E = hf$$

where $h = 6.62607004 \times 10^{-34}$ m^2 kg s^{-1}. For infrared light at telecommunications wavelengths (1550nm) this is enough energy to push one electron through about 1V of electrical potential. Wavelengths used in Finisar – what you need to know

2.6. Measuring Wavelengths and Frequencies

The most common way we use to measure the wavelength of light is to simply connect a fiber to an optical spectrum analyzer (OSA), but it's not very accurate. An Agilent OSA is generally accurate to 1 nm (about one 100 GHz optical channel!). Careful calibration can improve this but they are never very accurate. Also be careful to choose the "wavelength in vacuum" setting, or all the wavelengths will be calculated for "wavelength in air". The great advantage of an OSA is that they generally operate from the visible out to 1800nm or more. For mid-IR light use an FTIR (Fourier Transform Infrared spectroscopy system).

A wavemeter (Bristol is a well known brand) is generally accurate to 0.01 GHz, but can only look at one wavelength at a time and is slow, typically taking a second or two for a measurement. For more complex spectra consider using a (II-VI) WaveAnalyzer. This instrument is accurate to 0.5 GHz and can handle complex spectra in the C-band or L-band (but currently not outside these bands).

The Finisar Iris combines a tunable laser, a polarization controller and a receiver. This is great for measuring the wavelength dependent behavior of passive (non-light emitting) products. More about this in later chapters.

2.7. Characteristics of different wavelength ranges

Different wavelengths are useful for different tasks. UV photons have enough energy to kick electrons out of orbit, breaking bonds. They are particularly good for causing chemical reactions. Their short wavelength also means that they can be focused down to a very small spot size (we'll calculate this in a few weeks when we look at the Rayleigh criteria).

Photons with visible wavelengths (390 to 700nm) can move electrons between orbitals. The particular wavelengths that get absorbed depend on what orbital levels are available (so they are good wavelengths to tell the difference between materials).

Photons with far infra-red wavelengths (2-4 microns) interact strongly with the vibration modes of atomic bonds. They are very useful for identifying different organic chemicals.

Between the visible and far infrared is a fairly quiet region, the near infrared (700-2000nm), where the interactions between materials and light can be quite limited (semiconductors being commonly an exception). Because interactions are limited, amazing levels of transparency can be achieved – perfect for telecommunications.

2.8. ITU Grid and Flexgrid

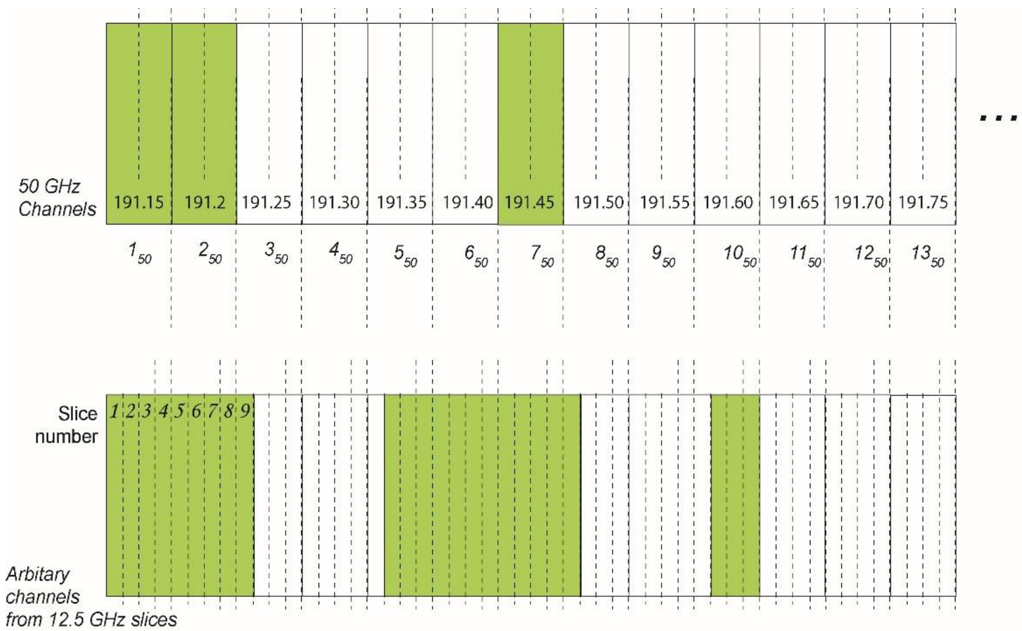


Figure 15 100 GHz and 50 GHz channel plans

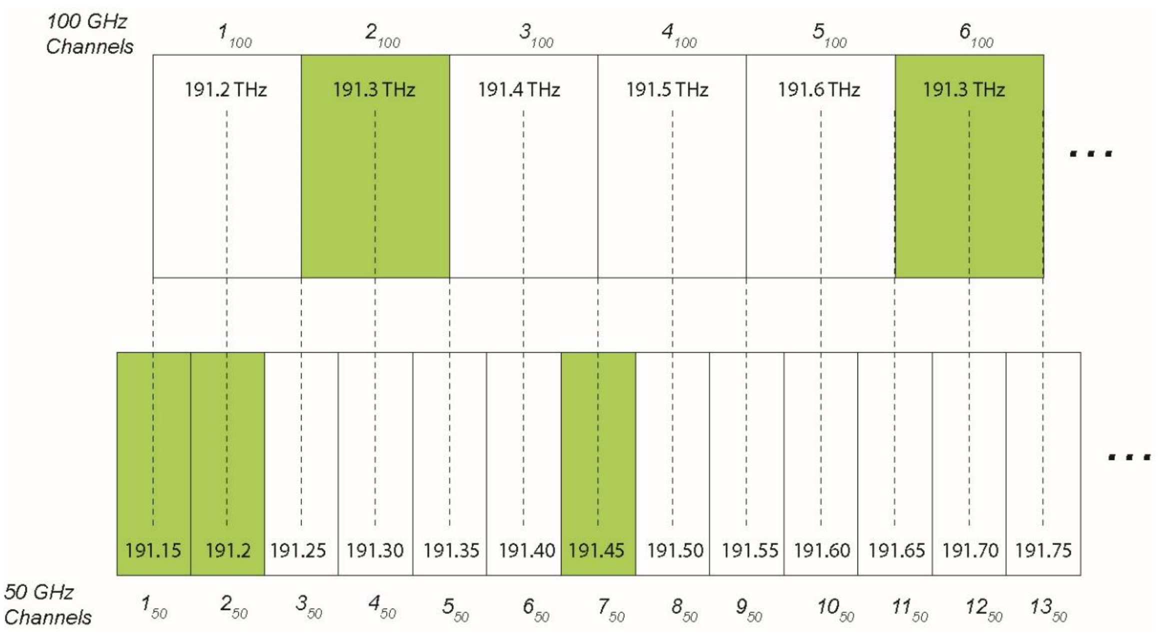
2.8.1.1.1. When people first put multiple signals, each at a different frequency through the same optical fiber, they needed a consensus on which frequencies to use. They borrowed the "ITU Grid" concept from the radio spectrum. The center frequency of each channel would be at a multiple of 100 GHz (0.1 THz), as shown in Figure 15. Usually 100 GHz channels would fit between the first channel (typically 191.2 or 191.3 THz) and the last channel, 48 channels later, close to 196.1 THz. As filter edges became sharper and more accurate, channels were narrowed to 50 GHz. This allowed 96 channels to fit in the C-band.

As new transponders move to data rates of 400 Gb/s or more, it's become difficult to fit these signals in a 50 GHz channel, but no one wants to move back to the spectrally wasteful 100 GHz channels, so instead we're moving to Flexgrid (probably my sole contribution to the telecom vocabulary), where the customer can build any channel out of slices. These were originally 12.5 GHz wide, but more recently, have been 6.25 GHz wide. Note from Figure 16 that the edges of the slices correspond to the edges of the 50 GHz and 100 GHz channels, so standard channel plans can be constructed with Flexgrid.

75 GHz channels are popular for 400 GHz transmission but some new systems still use 50 GHz channels, while being Flexgrid-ready.

2.9. Staying Safe (general remarks- not a substitute for a laser safety course!)

Light in the C and L bands is usually relatively eye-safe. This light can't get through the cornea and so never focuses on the retina. It generally takes real effort (or 10's of milliwatts) to do any real damage.



In contrast light between 700 and 1400nm can be very dangerous. You can't see it, but it still focusses down on the retina at the back of the eye. Any source more than 1 mW should be treated with serious respect (safety glasses, warning signs, good training etc.). Be particularly careful if the beam is collimated since a stray beam can affect someone across the room (or further away if the room has windows!).

Before commissioning a rig that includes lasers, make sure it's reviewed for laser safety. The main idea of such reviews is to figure out how much light could enter the iris should a mistake occur. The relevant standard is AS/NZS IEC 60825.1:2014. Remember too that high power lasers can start fires. In a previous company, we had the fiber output of a laser fall on a stack of papers. The resulting fire set off the sprinkler system which caused more damage. As of June 2022 Finisar Australia's laser safety officer is Philip Hambley.

2.10. The Optical Bands

Table 1 Typical channels for the C-band 100GHz ITU Grid

Band Name	Wavelengths (nm)	Description
O-band	1260 – 1360	Original telecom band. Still used in point-to-point systems and some CATV systems
S-band	1460 – 1530	Supervisory channels and PON (passive optical network) return signals
C-band	1530 – 1565	Almost all DWDM systems, Works well with erbium doped fiber amplifiers, CATV
L-band	1565 – 1625	Very popular in Japan and for expanded DWDM systems
U-band	1625 – 1675	Ultra-long wavelength band – little used at the moment

Table 2 The New Extended Bands

Band Name	Approx. Wavelengths (nm)	Approx. Frequencies (THz)	Description
Super C-band	1524.50 – 1571.65	190.750 – 196.650	The exact wavelength range varies but generally includes ~6THz (120 50GHz channels)
Super L-band	1575 – 1627	184.250-190.350	The exact wavelength range is still not well defined but is expected to include ~6THz (120 50GHz channels)

Table 3 Wavelengths and Frequencies for most common 50GHz channels

Frequency (THz)	Wavelength (nm)	Frequency (THz)	Wavelength (nm)	Frequency (THz)	Wavelength (nm)
191.7	1563.86	193.2	1551.72	194.7	1539.77
191.8	1563.05	193.3	1550.92	194.8	1538.98
191.9	1562.23	193.4	1550.12	194.9	1538.19
192	1561.42	193.5	1549.32	195	1537.40
192.1	1560.61	193.6	1548.51	195.1	1536.61
192.2	1559.79	193.7	1547.72	195.2	1535.82
192.3	1558.98	193.8	1546.92	195.3	1535.04
192.4	1558.17	193.9	1546.12	195.4	1534.25
192.5	1557.36	194	1545.32	195.5	1533.47
192.6	1556.55	194.1	1544.53	195.6	1532.68
192.7	1555.75	194.2	1543.73	195.7	1531.90
192.8	1554.94	194.3	1542.94	195.8	1531.12
192.9	1554.13	194.4	1542.14	195.9	1530.33
193	1553.33	194.5	1541.35	196	1529.55
193.1	1552.52	194.6	1540.56	196.1	1528.77

2.11. Revision and questions

2.12. Revision

We saw a lot of equations in this chapter. Some I want you to have “seen”, but these are the ones that I want you to be able to use automatically:

$c = f\lambda_o = \frac{\omega}{k_o}$ (where λ_o is the wavelength in a vacuum and $c = 299\ 792\ 458$ m/s)

Or rearranged

$$f = \frac{c}{\lambda_o} \qquad (1)$$

Taking the derivative, a small change in frequency

$$\delta f \approx -\frac{c}{\lambda^2}\delta\lambda \qquad (2)$$

You don’t absolutely need this, because you can just calculate the frequency for each of the wavelengths and subtract them, but do remember that in the C-band 50GHz frequency range is very close to 0.4nm wavelength range.

The speed of light in a material is

$$v = \frac{c}{n} \qquad (3)$$

Where *n* is the refractive index.

This means that the wavelength changes as we go from one material to another (the frequency can’t change otherwise there would be a pileup of waves at the interface)

$$\lambda = \frac{\lambda_o}{n} \qquad (3)$$

We also derived the equation

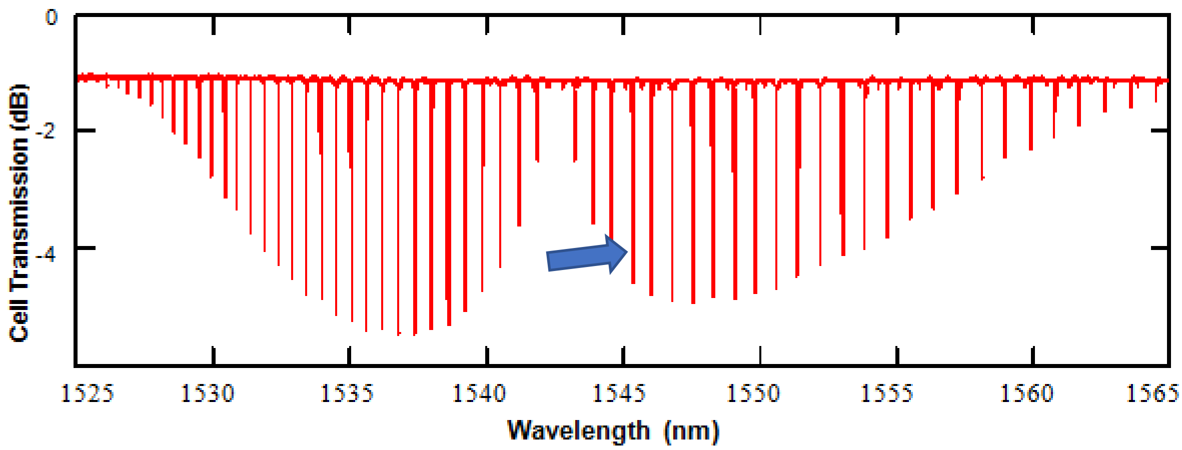
$$\frac{1}{u_g} = \frac{1}{u} - \frac{\lambda_o}{c}.\frac{dn}{d\lambda_o} \qquad (4)$$

You’ll probably never use this, but what I want you to understand is that the group velocity (the speed of the pulse) is different (and almost always slower) than the phase velocity. If the group velocity is changing with wavelength ($\frac{d^2n}{d{\lambda_o}^2} \neq 0$) then the pulse will disperse.

We often use HCN gas cells to calibrate test and measurement systems.

R Branch	Wavelength (nm)	P Branch	Wavelength (nm)
26	1527.63342	1	1543.11423

25	1528.05474	2	1543.80967
24	1528.4857	3	1544.51503
23	1528.92643	4	1545.23033
22	1529.37681	5	1545.95549
21	1529.83688	6	1546.69055
20	1530.30666	7	1547.43558



Graph of cell transmission vs. wavelength for 16.5cm $H^{13}C^{14}N$ cell.

(from www.wavelengthreference.com)

1. The absorption line marked with the arrow has what frequency (use the table to find the wavelength and calculate the frequency)?

2. What is the angular frequency of this light?

3. In a Chemistry Handbook absorption lines are measured in cm^{-1}. What would the wavenumber of this line be in cm^{-1} ? (as Chemists define it – note they don't multiply by 2π, they just change the wavelength to cm and invert it).

4. What is the period of this light?

5. Diamond has a refractive index of 2.4. How much would the wavelength and frequency change in diamond?

6. What is the frequency of light at 1528nm?

7. What is the frequency of light at 1564nm?

8. What is the frequency range across the C-band?

9. How many 50GHz channels can we fit in this range?

10. If a telecommunications channel is 50GHz wide centred at 1575nm, how many picometres does this correspond to (either use equation 1 at each end of the range and subtract them or use equation 2)?

11. 80km of SMF-28 fiber has 18.4 ps/(nm.km) of dispersion. If the signal is 25GHz wide, how much will it have spread after 80km? If it is a 20Gbaud/s signal (20×10^9 symbols/s), will this be a problem?

12. Use the equipment to measure the wavelengths of the sources. Choose the most appropriate instruments and explain your choice.

13. (Extension question) The refractive index of glass for *visible* wavelengths is often modelled by Cauchy's equation:

$$n = A + \frac{B}{\lambda^2}$$

Find the group velocity at 500nm, when A=1.5 and B=30,000 nm^2.

Note: do not use the Cauchy equation for infrared wavelengths. It doesn't work! Instead use the (more complicated) Sellmeier equation which does work. More details later in the course.

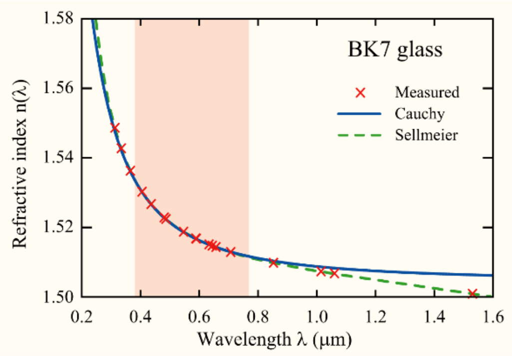
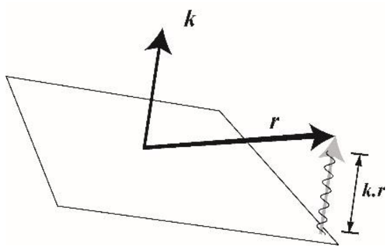


Figure 17 Comparison of the Cauchy and Sellmeier equation (Bob Mellish – from http://en.wikipedia.org/wiki/File:Cauchy-equation-1.svg).

Chapter 3. Power and Polarization –Part 1 Wave Equation

In three dimensions, we can write the magnitude of a plane wave as

$U(x, y, z, t) = U_o \cos(\boldsymbol{k}.\boldsymbol{r} - \omega t) = U_o e^{i(\boldsymbol{k}.\boldsymbol{r} - \omega t)}$



Where $\boldsymbol{k}$ is a vector in the direction of the wave, with magnitude k and $\boldsymbol{r}$ is the position we are considering (this has nothing to do with the unit vector in the z direction: $\hat{\boldsymbol{k}}$).

In this form, the del operator can be replaced with $i\boldsymbol{k}$, and $\frac{\partial}{\partial t}$ can be replaced with $i\omega$

$$\boldsymbol{\nabla} = \hat{\boldsymbol{\imath}}\frac{\partial}{\partial x} + \hat{\boldsymbol{\jmath}}\frac{\partial}{\partial y} + \hat{\boldsymbol{k}}\frac{\partial}{\partial x} \rightarrow i\boldsymbol{k}$$

$$\frac{\partial}{\partial t} \rightarrow -i\omega$$

Now Maxwell's equations then become

$$\boldsymbol{k} \times \boldsymbol{E} = -\mu\omega\boldsymbol{H}$$

$$\boldsymbol{k} \times \boldsymbol{H} = -\epsilon\omega\boldsymbol{E}$$

$$\boldsymbol{k}.\boldsymbol{E} = 0$$

$$\boldsymbol{k}.\boldsymbol{H} = 0$$

This tells us that $\boldsymbol{E}, \boldsymbol{H}$ and $\boldsymbol{k}$ are orthogonal in a uniform medium.

The energy flow per unit area is $\boldsymbol{S} = \boldsymbol{E} \times \boldsymbol{H}$. It's fairly easy to show that the magnitude of $\boldsymbol{S}$ (which we call the intensity of the light), $I = \frac{1}{2}E_oH_o = \frac{1}{2}\left(\sqrt{\frac{\mu_o}{\epsilon_o}}\right)nE_o^2$.

3.1. Polarization – Jones Vectors

If the electric field vector is perpendicular to the direction the light is travelling, then we can describe it with a two-element vector (since the component in the third direction is zero). But since it is a wave, we need a phase term as well. We use the complex argument.

$$\boldsymbol{E}(z,t) = E_{ox}\hat{\boldsymbol{\imath}}.\cos(kz - \omega t + \theta_x) + E_{oy}\hat{\boldsymbol{\jmath}}.\cos\left(kz - \omega t + \theta_y\right)$$

Using complex notation

$$\boldsymbol{E}(z,t) = Re[\mathrm{e}^{\mathrm{i}(kz-\omega t)}\left(\hat{\boldsymbol{\imath}}.E_{ox}e^{i\theta_x} + \hat{\boldsymbol{\jmath}}.E_{oy}e^{i\theta_y}\right)]$$

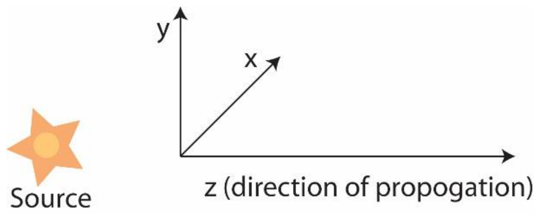
Usually, the fact that we just take the real component is implicit and we write this as

$$\boldsymbol{E}(z,t) = \mathrm{e}^{\mathrm{i}(kz-\omega t)}\left(\hat{\boldsymbol{\imath}}.E_{ox}e^{i\theta_x} + \hat{\boldsymbol{\jmath}}.E_{oy}e^{i\theta_y}\right) = \boldsymbol{E_o}\mathrm{e}^{\mathrm{i}(kz-\omega t)}$$

A very convenient way to write $\boldsymbol{E_o}$ is the Jones Matrix notation

$$\boldsymbol{E_o} = \begin{bmatrix} E_{ox}e^{i\theta_x} \\ E_{oy}e^{i\theta_y} \end{bmatrix}$$

The choice of "x" and "y" co-ordinates here is arbitrary and there are lots of other possible choices. In this representation



$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is linearly polarized light in the x-direction.

$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is linearly polarized light in the y-direction.

$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is ……………………………………………………..

$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is ……………………………………………………..

$\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$ is linearly polarized light, polarized an angle θ

$\frac{1}{\sqrt{2}}\begin{bmatrix} e^{i\frac{\pi}{2}} \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} i \\ 1 \end{bmatrix}$ is right circularly polarized light with respect to the receiver – i.e. the vector rotates clockwise in time looking back to the source.

$\frac{1}{\sqrt{2}}\begin{bmatrix} e^{i\frac{3\pi}{2}} \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} -i \\ 1 \end{bmatrix}$ is left circularly polarized light with respect to the receiver. Note Born & Wolf's "Principles of Optics" define it this way, but the quantum mechanics crowd, IEEE and Charles Tsao (in "Optical Fiber Waveguide Analysis) use the opposite definition. The most recent version of the Wikipedia article on Jones Vectors agrees with the convention used here (as of July 2023) but that article seems to change regularly!)

It's important to be clear what we mean by "right circularly polarized light with respect to the receiver". If you watch the beam in time, looking *back* towards the source, the electric field vector moves in a clockwise direction. Note that most optics texts (including "Introduction to Modern Optics") use this convention. Electrical engineers and radio astronomers use the opposite terminology for left and right handedness.

What do you think happens when an LCP beam bounces off a mirror?

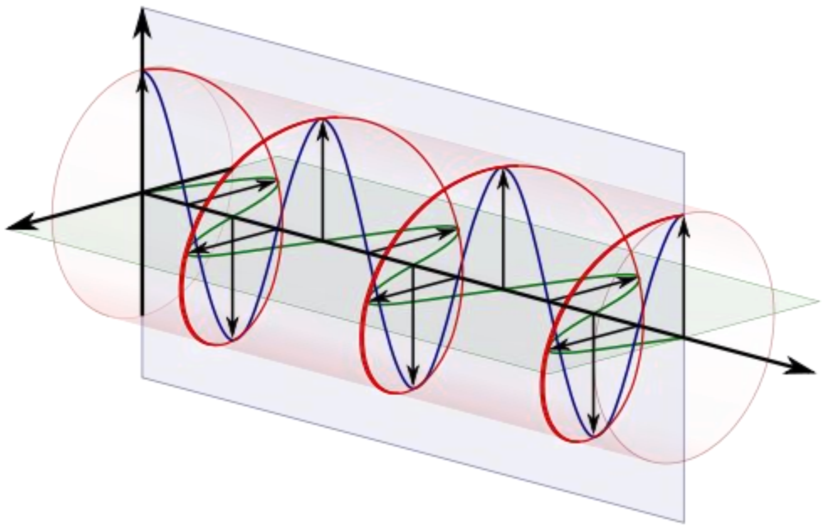


Figure 18 Left circularly polarized light (picture from Wikipedia Commons) looking back towards the source

Left or right handedness (from the point of view of the receiver) is determined by pointing one's left or right thumb *toward* the source, *against* the direction of propagation, and matching the curling of one's fingers to the temporal rotation of the field as the wave passes a given point in space.

There is nothing special about horizontally and vertically polarized light. Sometimes it is much more useful to choose a circular basis set. Then

$\begin{bmatrix}1\\0\end{bmatrix}$ is left circularly polarized light.

$\begin{bmatrix}0\\1\end{bmatrix}$ is right circularly polarized light.

$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}$ is x polarized light using a circular basis set

$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}$ is y polarized light using a circular basis set

An arbitrary polarization using this basis set can be written as

$$\boldsymbol{E_o} = \begin{bmatrix} E_{ol}e^{i\theta_l} \\ E_{or}e^{i\theta_r} \end{bmatrix}$$

For the rest of this chapter, we will assume horizontal and vertical polarizations as our basis set.

The simplest and most common way a material affects a polarization state is that it causes one polarization to travel at a different speed to the orthogonal state. This is called birefringence. For a wave plate, the fast axis is marked with a dot or flat spot, on unmounted optics, or a line on mounted optics. Using a linear basis set, a quarter-wave plate with its fast axis in the vertical direction would look like:

$\begin{bmatrix} e^{i\pi/4} & 0 \\ 0 & e^{-i\pi/4} \end{bmatrix}$ or, keeping the y component phase change zero, $\begin{bmatrix} i & 0 \\ 0 & 1 \end{bmatrix}$

This means that the x component of the light

$$\boldsymbol{E_x}(z,t) = Re\left[\mathrm{e}^{\mathrm{i}(kz-\omega t)}\left(\boldsymbol{\hat{\imath}}.E_{ox}e^{i\frac{\pi}{2}}\right)\right] = \boldsymbol{\hat{\imath}}.E_{ox}Re[\mathrm{e}^{\mathrm{i}\left(kz-\left(\omega t-\frac{\pi}{2}\right)\right)}]$$

So the x component of the electric field arrives a quarter of a wave later than the y component.

How would this affect light polarized at 45°?

$$\begin{bmatrix} i & 0 \\ 0 & 1 \end{bmatrix} . \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} i \\ 1 \end{bmatrix}$$

How would it affect light at -45°? (try the calculation)

A quarter-wave plate with its fast axis in the horizontal direction would look like

$$\begin{bmatrix} -i & 0 \\ 0 & 1 \end{bmatrix}$$

A wave plate with its slow axis in the horizontal plane producing an arbitrary phase shift (γ) relative to the horizonal can be written as

$$\begin{bmatrix} e^{i\gamma} & 0 \\ 0 & 1 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} e^{i\gamma/2} & 0 \\ 0 & e^{-i\gamma/2} \end{bmatrix} \qquad (3.1)$$

In particular, a half wave plate oriented in the X-Y plane will have the form

$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ or $\begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix}$, keeping the average phase change zero.

What does this do to the polarization of linear polarized light, polarized at angle θ? (reminder the Jones vector for linearly polarized light at angle θ to the horizontal is $\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$)

What angle was the light rotated by?

We will show next week that the general form for a birefringent element is

$$\boldsymbol{B} = \begin{bmatrix} \cos(\frac{\phi}{2}) + i\frac{\gamma}{\phi}\sin(\frac{\phi}{2}) & \left(\frac{i\alpha}{\phi} + \frac{\beta}{\phi}\right)\sin(\frac{\phi}{2}) \\ \left(\frac{i\alpha}{\phi} - \frac{\beta}{\phi}\right)\sin(\frac{\phi}{2}) & \cos(\frac{\phi}{2}) - i\frac{\gamma}{\phi}\sin(\frac{\phi}{2}) \end{bmatrix} \qquad (3.2)$$

where $\phi = \sqrt{\alpha^2 + \beta^2 + \gamma^2}$ and (for a linear basis set), γ is how far the horizontal component moves forward relative to the vertical component, α is how far the +45 degree component moves forward relative to the -45 degree component and β is how far the LCP moves forward compared to the RCP component.

If a linear waveplate is at an angle θ to the positive x axis, with retardance ϕ, then $\gamma = \phi\cos 2\theta$ and $\alpha = \phi\sin 2\theta$.

The Jones matrix for a mirror is

$$M = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

Since on reflection (to maintain a right-handed co-ordinate system) we change the sign of the x-axis.

3.2. Polarization in Optical Fibers

3.2.1.1. Bending Fibers creates linear birefringence

Creating birefringence in an optical fiber is very simple. One way is to simply bend the fiber. The resulting birefringence is

$$\gamma = \frac{3.35\times10^{-9}}{\lambda R^2}\ \text{rad/m}$$

where R is the bending radius[1] and the slow axis is in the plane of the bend. If we assume a wavelength of 1550 nm, and multiply by $2\pi R$ to make a loop then we can calculate that

$$\gamma_l = \frac{0.013}{R}\ \text{radians/loop}$$

You should be able to show that a 3.4 cm radius loop will make a $\frac{\pi}{8}$ radian wave-plate. This is the principle behind a fiber loop polarization controller.

From this equation we can see how a polarization maintaining (PM) fiber works (and doesn't). A typical PM optical fiber has a birefringence of around 6000 radians/m (or a "beat length" of around 1mm). If we put light in along one of the "principal axes" (typically the slow axis), it is unaffected by the birefringence. Calculate the matrix from equation 3.1. But now let's add a 5 radian birefringence at +45 degrees to the principle axis distributed over a meter length. Use

[1] for more details see Scott Rashleigh's excellent paper "Origins and control of polarization effects in single-mode fibers", *Lightwave Technology, Journal of* , vol.1, no.2, pp.312,331, Jun 1983.

equation 3.2 to calculate the resulting matrix. You can see that the effect is essentially zero (if you remember how to calculate eigenvectors you can prove it).

Other ways you can create a linear birefringence in an optical fiber are:

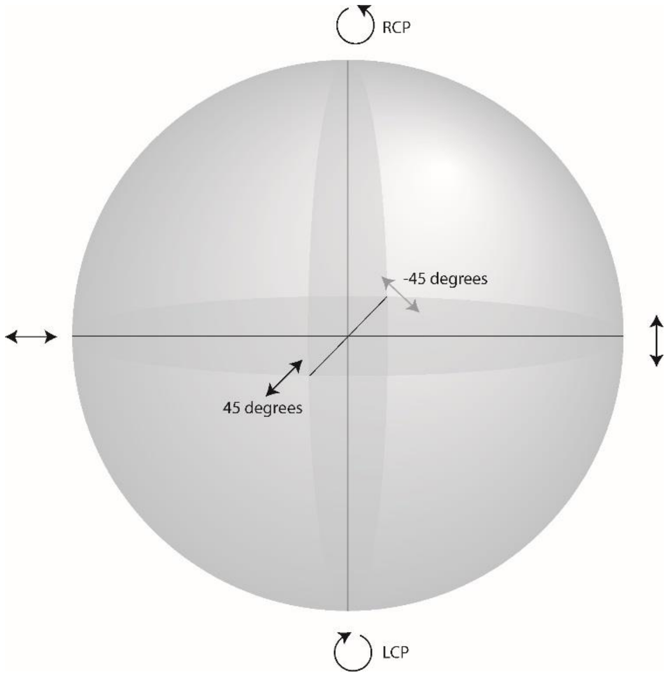
1. An elliptical core
2. A stressed core (generally made by embedding two cylinders of boron doped glass either side of the core – as the glass cools from its liquid state at around 1000°C, these shrink faster, stressing the core. The slow axis of the fiber lines up with this stress elements). Note that this birefringence is temperature sensitive, not surprisingly changing by ~1/1000 per degree change in temperature. A good HiBi (often called polarization preserving or panda) fiber will create a phase delay of around 6000 radians/meter.
3. Applying an external compressive stress.

3.2.1.2. Twisting fibers creates circular birefringence

When you twist an optical fiber, the polarization of the light will move in the direction of the twist, but only about 7% as much as the twist.

3.2.2. Polarization – Poincaré Sphere

We can also represent polarization graphically on the Poincaré Sphere.



Where $\theta = \theta_r - \theta_l$ and $\tan\frac{\psi}{2} = \frac{E_{or}-E_{ol}}{E_{ol}+E_{or}}$. Notice that θ is twice the angle of the linear polarization. “Orthogonal” polarizations are opposite each other on the sphere. A birefringent element creates a right-handed rotation of magnitude ϕ around the slow axis (the component phase-shifted forward) going through the point (α, γ, β). Birefringences that occur at the same point can be added as vectors, but ones that occur in sequence can’t.

Considering our previous example, if we add a simultaneous orthogonal 5 radian birefringence to a 6000 radian birefringence (i.e. at 90 degrees in Poincaré space), how much will that change the direction of the principal axes?

3.3. Polarizers

When you need to polarize infrared light in the telecom bands, ordinary polaroid film won't work. However, you have a lot of other choices:

1. In-line polarizers are made by polishing back the side of a fiber and putting a thin layer of metal close to the core. They only transmit one polarization (the other is absorbed) so don't put more than ~300mW of power through one of these. These typically give about 28 dB of extinction ratio (~600:1) and cost about $150.
2. Polarising couplers. These separate out the two polarizations in a PM fiber, with around 30dB extinction ratio (1000:1) but are more expensive
3. Polarising fiber. Polarizing fiber only works over a narrow wavelength range and this range shifts as the fiber is bent. However in some applications it is extremely useful.
4. Bulk optic polarizing beam splitter cubes: These split light into vertical and horizontal polarizations with a 30 dB extinction ratio, splitting the light at a 90 degree angle. They are typically specified for 1200 – 1600nm wavelengths for telecom applications. A small cube will cost ~$200.
5. Calcite polarizers: These are made from mined, optically perfect calcite and are expensive! However they will give 100,000:1 (50 dB) extinction ratios. We'll talk about how these work later in the course

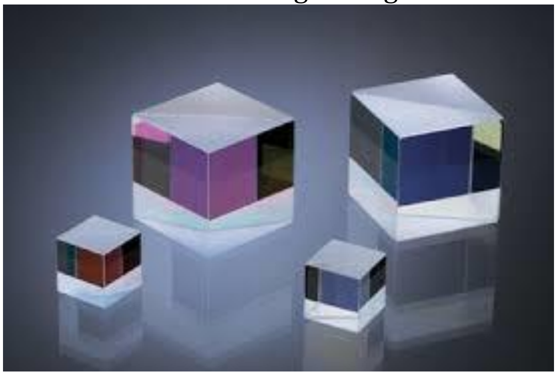


Figure 19 Polarizing beam splitter (from Ealing Catalog)

3.4. Depolarizers

"Depolarizer" can mean a lot of different things.

1. **Spatial depolarizer** It can mean that different points across the beam have different polarizations (spatial depolarising). These are usually made using a quartz wedge. This doesn't work in a single mode fiber because the single mode can only have one polarization for a given time and frequency.
2. **Polarization Scrambler (Temporal depolarizer)** If we change the polarization quickly in time then we can "depolarize" the signal (until someone looks at it with an even faster detector). The FiberPro PS3000 is very popular. General Photonics also makes one (PCD-104) .
3. **Frequency depolarizer** A Lyott depolarizer consists of two or three lengths of HiBi (polarization preserving) fiber spliced at 45 degrees to the previous fiber, resulting in each frequency having a different polarization. This isn't much use if you're using a narrow laser source or looking at a narrow frequency.

Sources

Edge emitting diode lasers typically have around 200:1 polarization extinction. VCSEL (vertical cavity lasers) tend to be somewhat randomly polarized. ASE (amplified spontaneous emission) fiber light sources have a mostly randomly (rapidly time varying) polarization, however they generally have some residual polarization so following them with a polarization scrambler can be helpful.

Chapter 4. Polarization Part 2

The aim of this week's class is to make you comfortable solving problems using Poincaré Space and Jones matrices. Once you can do this, any combination of waveplates, mirrors and polarizers will make sense.

4.1. Understanding Elliptical polarizations

An elliptical polarization with major axes length "a" and minor axis length "b" can be drawn as an ellipse on an X-Y graph or as a point in Poincaré Space.

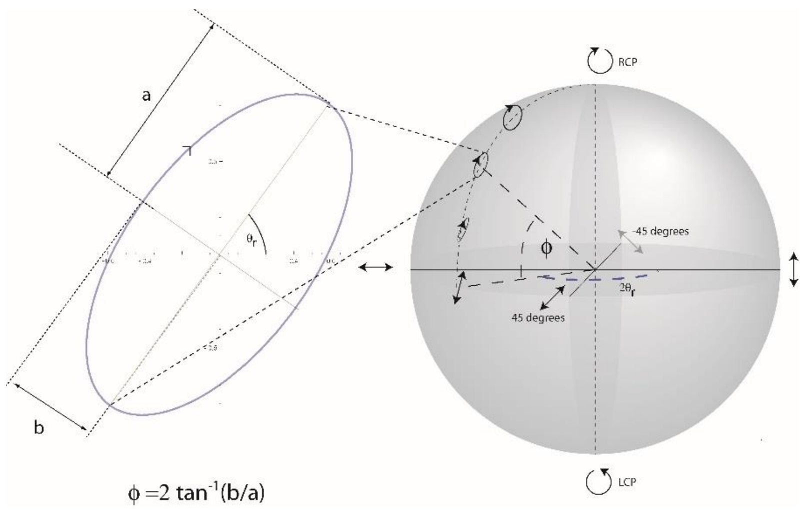


Figure 20 The relationship between representing the electric field vector on an X-Y plot and the representation of polarization in Poincaré Space.

If we consider a position (s_1, s_2, s_3) on the sphere where s_1 is the coordinate in the vertical/horizontal direction, s_2 in the -45/+45 direction, and s_3 in the RCP/LCP direction and derive it from the Jones vector $\begin{bmatrix} a_x e^{i\varepsilon} \\ a_y \end{bmatrix}$

$$s_1 = {a_x}^2 - {a_y}^2$$

$$s_2 = 2a_x a_y \cos\varepsilon$$

$$s_3 = 2a_x a_y \sin\varepsilon$$

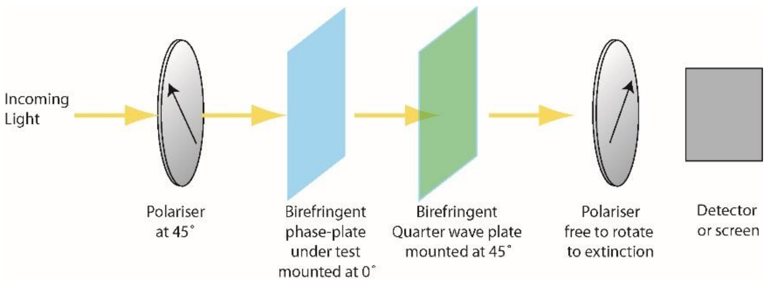
The (x,y,z) coordinates are called the Stokes Parameters. Normally people will add another Stokes parameter

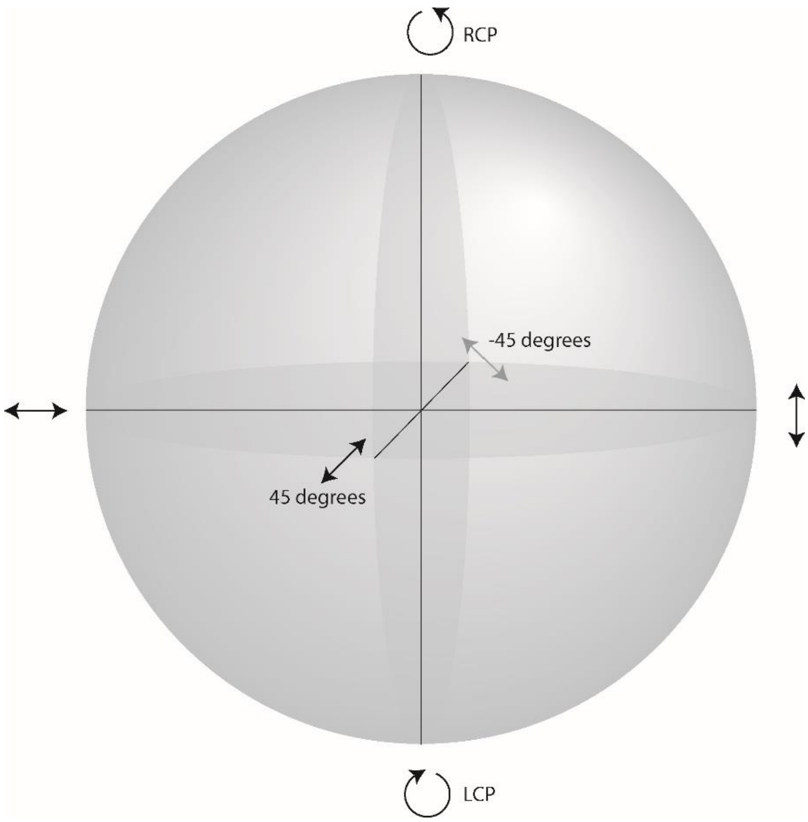
$$s_o = {a_x}^2 + {a_y}^2$$

which gives the intensity of the light (if the light is partly depolarized $s_o > {a_x}^2 + {a_y}^2$).

4.2. More about Poincaré Space

A birefringence can be represented as a rotation in Poincaré Space, where the polarization state rotates in a right-handed sense about the slower "preserved" polarization state (or clockwise when viewed from the side of the faster polarization). Let's try an exercise. A way to measure a quarter waveplate is to place in order a polarizer, our waveplate under test, a reference waveplate and another polarizer (called an "analyzer". We first align the input polarizer so the incoming light is polarized at 45 degrees. We then align the waveplate under test so its fast axis is horizontal. Then we follow it with our reference waveplate with its fast axis at 45 degrees. Then we rotate the analyzer to get extinction of the light. If the waveplate has a retardance of 87 degrees, what angle will we need to rotate the polarizer to, to get extinction?





4.3. Ways of producing birefringence

There are lots of ways of producing birefringence in an optical fiber. We can produce a circular birefringence simply by twisting the fiber through an angle τ. For silica fiber

$$\beta \approx 0.146\tau$$

We can create a linear birefringence by forcing the fiber between two flat plates

$$\gamma = 4C_s \frac{f}{\pi r E}$$

Where $C_s \approx \frac{-3.34\times10^{-11}}{k}$ m^2kg^{-1} and E is Young's modulus ($7.45\text{x}10^9$ $kg.m^{-2}$), r is the fiber radius and f the force in Newtons/meter. Similar equations exist for V-grooves, and they can all be found in the paper by Scott Rashleigh[2].

Many materials are linearly birefringent. Later in the course we will discuss calcite, quartz, lithium niobate. Circularly birefringent materials (often called "optically active" materials) are usually organic and chiral (the left-handed and right-handed versions of the molecules are different). A magnetic field going through a material can also create circular birefringence (the is called the "Faraday Effect" and is what we use to make isolators).

4.4. Describing any waveplate

This week I'd like to make sure you have the tools to solve polarization problems in one dimension. We will solve problems graphically in Poincaré Space, and algebraically using Jones Matrices.

A phase delay of γ between two orthogonal polarization states that form the basis of your Jones matrix representation can be represented by the matrix. We'll keep using the horizontal and vertical linear polarizations.

$$\begin{bmatrix} e^{i\frac{\gamma}{2}} & 0 \\ 0 & e^{-i\frac{\gamma}{2}} \end{bmatrix}$$

A quick check will show you that the $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ vectors are unaffected except for a phase shift. This is equivalent to a rotation about the poles in Poincaré Space. We are adding phase to the x component. This means that if we imagine ourselves to be a surfer trying to stay at the same point on the wave as it travels then we would need to decrease z (move less far than for the other polarization) or take more time to travel (i.e. the wave has travelled more slowly). The x polarization is phase shifted forward, it has travelled more slowly, so the rotation is in a right-handed sense around the horizontal direction (clockwise when looking at the sphere from the vertical polarization side)

[2] Scott Rashleigh "Origins and control of polarization effects in single-mode fibers," *Lightwave Technology, Journal of* , vol.1, no.2, pp.312,331, Jun 1983.

A phase delay (birefringence) of α between the components of +45° $\begin{bmatrix}1\\1\end{bmatrix}$ and -45° $\begin{bmatrix}-1\\1\end{bmatrix}$ polarization states can be written as

$$\begin{bmatrix}\cos\frac{\alpha}{2} & i\sin\frac{\alpha}{2}\\ i\sin\frac{\alpha}{2} & \cos\frac{\alpha}{2}\end{bmatrix}$$

Where +45° is the slow axis for positive α.

Similarly a phase shift of β between RCP and LCP light (RCP going faster) can be written as

$$\begin{bmatrix}\cos\frac{\beta}{2} & \sin\frac{\beta}{2}\\ -\sin\frac{\beta}{2} & \cos\frac{\beta}{2}\end{bmatrix}$$

We can't build up an arbitrary rotation with components (α, β, γ) by applying one after the other. The order in important (for the mathematically inclined, matrix operations don't commute). However infinitesimal rotations do commute and we can describe any of these rotations as the product of infinitesimal rotation (expand the right hand side and you will get the corresponding Taylor expansions of the functions on the left hand side).

$$\begin{bmatrix}e^{i\frac{\gamma}{2}} & 0\\ 0 & e^{-i\frac{\gamma}{2}}\end{bmatrix} = \lim_{n\to\infty}\left(\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \frac{\gamma}{2n}\begin{bmatrix}i & 0\\ 0 & -i\end{bmatrix}\right)^n$$

$$\begin{bmatrix}\cos\frac{\alpha}{2} & i\sin\frac{\alpha}{2}\\ i\sin\frac{\alpha}{2} & \cos\frac{\alpha}{2}\end{bmatrix} = \lim_{n\to\infty}\left(\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \frac{\alpha}{2n}\begin{bmatrix}0 & i\\ i & 0\end{bmatrix}\right)^n$$

$$\begin{bmatrix}\cos\frac{\beta}{2} & \sin\frac{\beta}{2}\\ -\sin\frac{\beta}{2} & \cos\frac{\beta}{2}\end{bmatrix} = \lim_{n\to\infty}\left(\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \frac{\beta}{2n}\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}\right)^n$$

If you're trying this, note that something of the form $\lim_{n\to\infty}\left(1+\frac{a}{n}\right)^n$ can be written as the Taylor Expansion

$$\mathrm{e}^{\mathrm{a}} = \lim_{n\to\infty}\left(1+\frac{a}{n}\right)^n = 1 + \frac{a}{1!} + \frac{a^2}{2!} + \frac{a^3}{3!} + \cdots$$

The four unitary matrices $\boldsymbol{I} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, \boldsymbol{\sigma_x} = \begin{bmatrix}i & 0\\ 0 & -i\end{bmatrix}, \boldsymbol{\sigma_y} = \begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}, \boldsymbol{\sigma_z} = \begin{bmatrix}0 & i\\ i & 0\end{bmatrix}$ form a (closed) "group"; which means that if you multiply any two together, you get another matrix in the group (try it and see that $\boldsymbol{\sigma_x}.\boldsymbol{\sigma_y} = \boldsymbol{\sigma_z}$). If you've studied second year quantum mechanics, then you have probably come across the Pauli matrices describing the interaction of a spin ½ particle in a magnetic field; the mathematics is the same.

We can then combine these infinitesimal rotations. A rotation with components (α, β, γ) has the form

$$\exp\left(\mathrm{i}\frac{\alpha}{2}\boldsymbol{\sigma_x} + \mathrm{i}\frac{\beta}{2}\boldsymbol{\sigma_y} + \mathrm{i}\frac{\gamma}{2}\boldsymbol{\sigma_z}\right) = \lim_{n\to\infty}\left(\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \frac{\alpha}{2n}\begin{bmatrix}i & 0\\ 0 & -i\end{bmatrix} + \frac{\beta}{2n}\begin{bmatrix}0 & i\\ i & 0\end{bmatrix} + \frac{\gamma}{2n}\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}\right)^n$$

$$= \begin{bmatrix}\cos(\frac{\phi}{2}) + i\frac{\gamma}{\phi}\sin(\frac{\phi}{2}) & \left(\frac{i\alpha}{\phi} + \frac{\beta}{\phi}\right)\sin(\frac{\phi}{2})\\ \left(\frac{i\alpha}{\phi} - \frac{\beta}{\phi}\right)\sin(\frac{\phi}{2}) & \cos(\frac{\phi}{2}) - i\frac{\gamma}{\phi}\sin(\frac{\phi}{2})\end{bmatrix}$$

Where $\phi = \sqrt{\alpha^2 + \beta^2 + \gamma^2}$

This lets us describe a uniform birefringent element like a wave-plate or a circularly birefringent organic chemical. Crystals can be a bit more complex, since they can also cause a beam to change direction – but we will discuss that later.

4.5. Describing a Polarizer

A polarizer is not a unitary matrix. It removes light from a path. A linear polarizer (using a linear basis set) transmitting at angle θ (in real space) has the form[3]

$$\begin{pmatrix}\cos^2\theta & \cos\theta\sin\theta\\ \cos\theta\sin\theta & \sin^2\theta\end{pmatrix}$$

Where θ is the angle in real space between the direction of maximum transmission and the x axis.

4.6. Describing a Mirror Orthogonal to the Beam

A mirror adds an interesting degree of conceptual complexity because not only does it swap left handed and right handed components, but also, after a mirror +45° and -45° are reversed. The Jones Matrix (in linear coordinates) for a mirror is $\boldsymbol{M_m} = \begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}$. In Poincaré Space a mirror swaps the left-handed and right-handed values and +45° to -45° values.

4.7. A first example of a compound structure: A Faraday Rotator mirror (orthoconjugate reflector)

A Faraday rotator consists of a circularly birefringent element produced by a magnetic field in the direction of the light's (incoming) travel producing a 45° rotation, followed by a mirror, at which point the light returns through the circularly birefringent element, travelling against the magnetic field (so the phase shift is reversed)

---

[3] Using a circular basis set it is

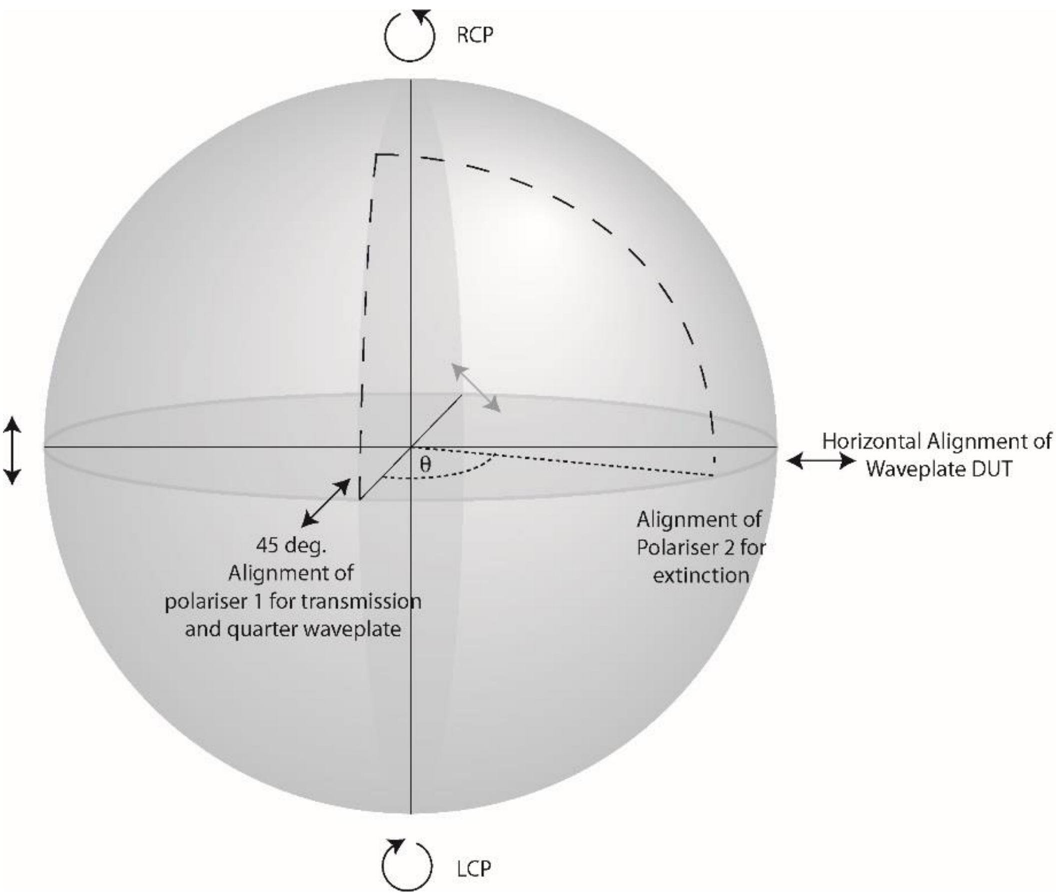
$$\frac{1}{2}\begin{pmatrix}1 & e^{-i2\theta}\\ e^{+i2\theta} & 1\end{pmatrix}$$

$$\boldsymbol{M_{fr}} = \begin{bmatrix} \cos\frac{\pi}{4} & \sin\frac{\pi}{4} \\ -\sin\frac{\pi}{4} & \cos\frac{\pi}{4} \end{bmatrix} \cdot \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos\frac{\pi}{4} & -\sin\frac{\pi}{4} \\ \sin\frac{\pi}{4} & \cos\frac{\pi}{4} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

This device has the remarkable effect of converting any polarization state to its orthogonal state. Note that linearly polarized light at 45° returns as linearly polarized light at 45°, but only because the direction of the x-axis has been reversed on reflection. The electrical field vector is at 90°

4.8. A second example of a compound structure: A measurement system for a phase plate (using Poincaré Space)

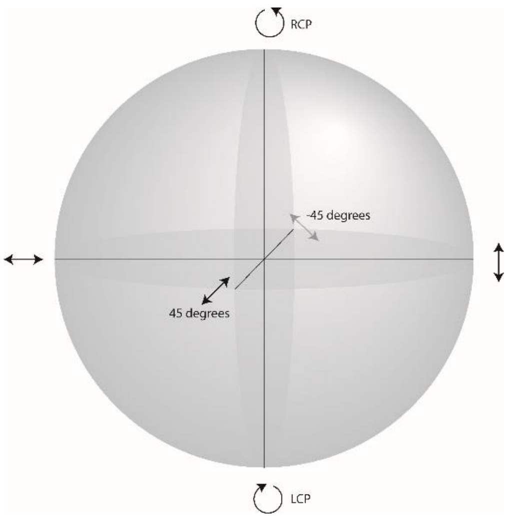
Here's the sketch which I hope you were able to produce earlier. With a pair of crossed polarizers and a quarter wave-plate (of imperfect quality), we can build a very accurate measurement system for measuring the retardance of a wave-plate. This is most easily seen on the Poincaré sphere



We align the optical axis of our reference quarter-wave plate using "crossed polarizers". We then align the quarter wave plate being tested at 45 degrees to the reference quarter wave plate. Finally we measure the polarization of the light that emerges. If the resulting polarization is θ degrees to the horizontal then the "quarter wave plate" actually has a retardance of 90-2θ degrees.

4.9. Questions

1. We send light polarized at +45° into a half-wave plate with its fast axis at 90° (vertical). What is the polarization of the output light? (First solve this with Jones Matrices and then with Poincaré Sphere)



2.
   a. What is the Jones vector for linearly polarized light at +22.5° to the horizontal? (see page 32 if you need the formula)

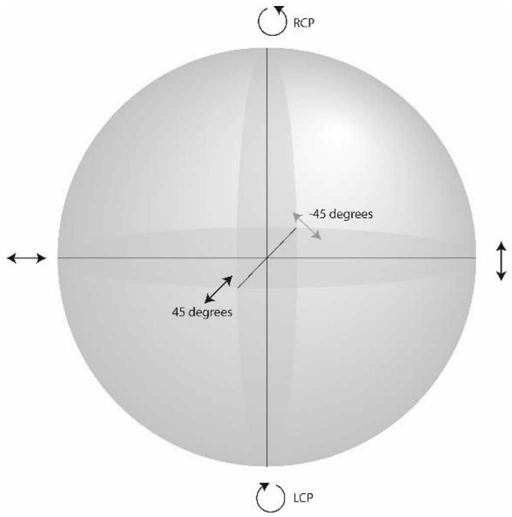
   b. What is the polarization Jones vector for the light from question 1 after it has gone through a half wave plate with its fast axis in the vertical direction? (calculate this with Jones vectors and then using the Poincaré Sphere).

   c. How much has the angle of the light changed by in real space?

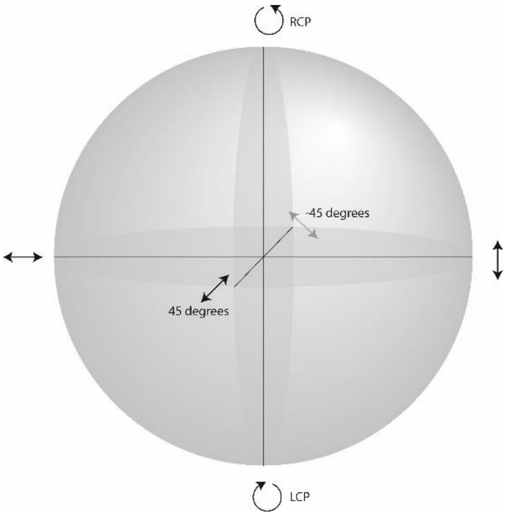
3. A linear polarizer can be written: $\begin{pmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{pmatrix}$

Multiply the matrix by the vector $\begin{bmatrix} a \\ b \end{bmatrix}$ to show that whatever polarization goes in to the polarizer, the same polarization comes out. What polarization is blocked? What is transmitted?

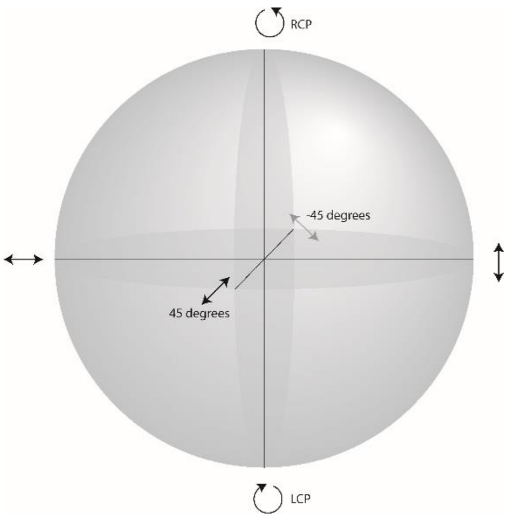
4. You have a half wave plate. What angle would you orient the fast axis of the waveplate to in order to convert horizontally polarized light to light at +66 degrees?



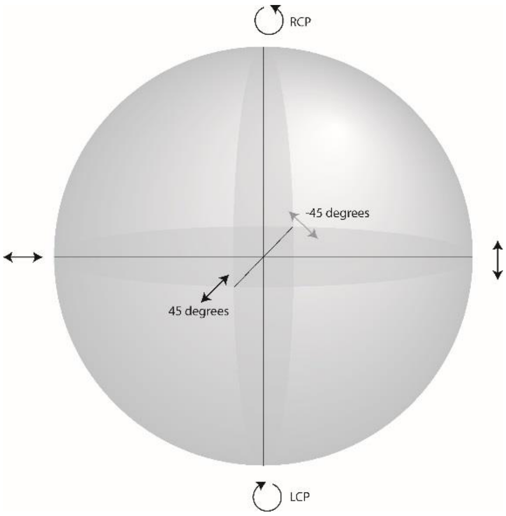
5. You have a quarter waveplate and a light source polarized at 30 degrees to the horizontal. What angle would you orient the fast axis of the waveplate to produce right circularly polarized light?

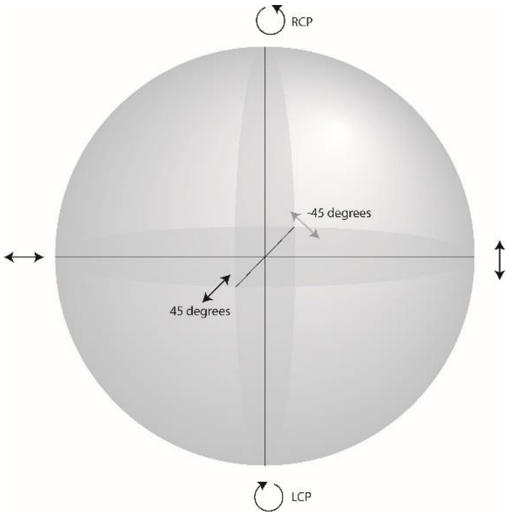


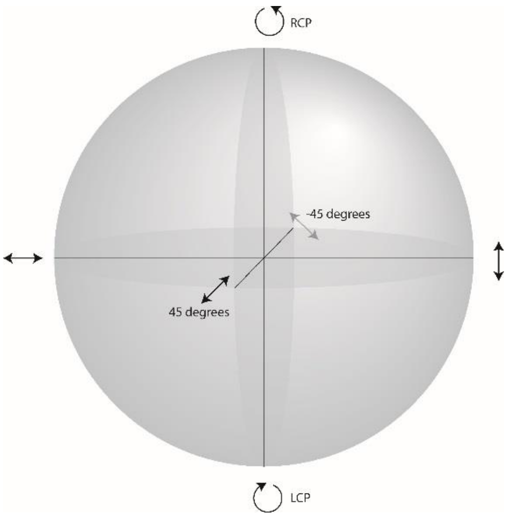
You have a rotatable quarter waveplate, followed by a half wave plate followed by another quarter wave plate. What angles could you set the plates to in order to change light which is right-handed elliptically polarized with an ellipticity such that $\frac{b}{a} = \sqrt{2}$-1, to left hand elliptical with an ellipticity of $\frac{b}{a} = \sqrt{2}$-1 (use the formula $\phi = 2 \arctan \frac{b}{a}$)?

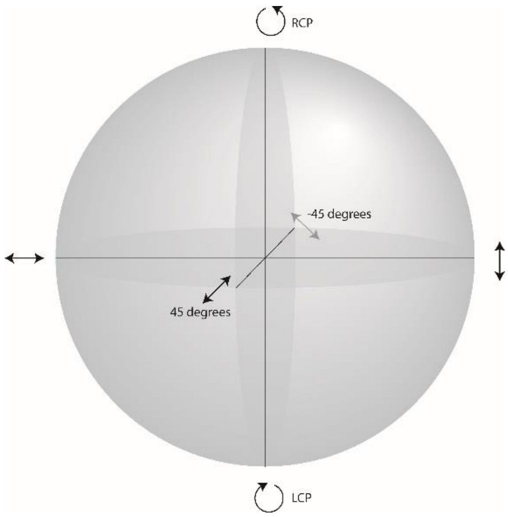


6. You have the same three wave plates followed by a mirror. What angles could you set the plates to in order to change light which is right-handed elliptically polarized with an ellipticity of $\sqrt{2}$-1, to left circularly polarized light?


RCP
-45 degrees
45 degrees
LCP


RCP
-45 degrees
45 degrees
LCP


RCP
-45 degrees
45 degrees
LCP


RCP
-45 degrees
45 degrees
LCP

Chapter 5. Refraction

5.1. Refraction – the mathematics

Two ways of thinking about Snells' Law.

The picture way;

The wavefront is always perpendicular to the direction of travel. In a given time Δt the wavefront travels $\frac{c}{n}\Delta t$. Using triangles we will show that

$$n_1 \sin\theta_1 = n_2 \sin\theta_2$$

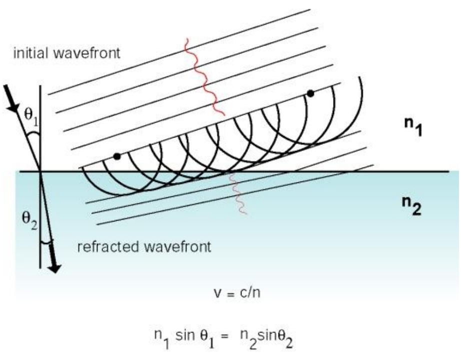


Figure 21 The change in velocity of the wave causes a change in the direction

The mathematical way (I've switched to θ and θ' here to be consistent with your textbook.

In order for the waves to stay synchronised at the surface (the phase is continuous)

$$\boldsymbol{k}.\boldsymbol{r} = \boldsymbol{k}'.\boldsymbol{r} = \boldsymbol{k}''.\boldsymbol{r}$$

Where $\boldsymbol{k}$ is the incident wavevector, $\boldsymbol{k}'$ the reflected wavevector and $\boldsymbol{k}''$ the transmitted wavevector. If we place our origin on the interface, then

$$\boldsymbol{k}.\boldsymbol{r} = kr\cos(\theta - \tfrac{\pi}{2}) = kr\sin\theta$$

and similarly

$$\boldsymbol{k}''.\boldsymbol{r} = k''r\cos\left(\theta' - \frac{\pi}{2}\right) = k''r\sin\theta'$$

So

$$k''r\sin\theta' = kr\sin\theta$$

So

$$\frac{2\pi n_2}{\lambda_o} r\sin\theta' = \frac{2\pi n_1}{\lambda_o} r\sin\theta$$

Thus

$$n_2\sin\theta' = n_1\sin\theta$$

5.2. The reflected wave

But what about the reflected wave? We can use the fact that the electric field and magnetic field are continuous to work out what happens at boundaries. We'll consider the two cases separately where the electric field is parallel to the surface and when the magnetic field is parallel to the surface.

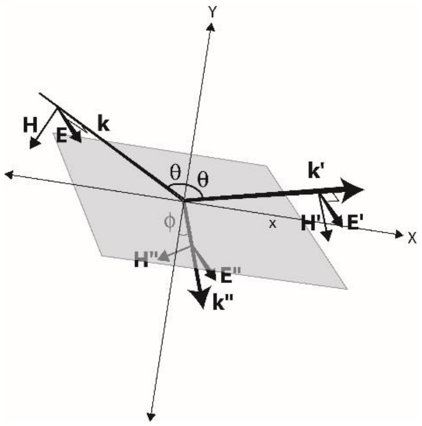


Figure 22 Reflection with a transverse electric field

First consider the TE case (Electric field is parallel). The field on one side of the boundary must match the field on the other side:

$$E + E' = E'' \qquad ...(1)$$

(electric field is continuous over the boundary) and

$$-H\cos\theta + H'\cos\theta = -H''\cos\phi \qquad ...(2)$$

(magnetic field is continuous over the boundary)

i.e. $$-kE\cos\theta + k'E'\cos\theta = -k''E''\cos\phi \qquad ...(3)$$

We can use the fact that the velocity of light $\frac{c}{n} = \frac{\omega}{k}$ to say $k = n\frac{\omega}{c}$ and convert this equation to a form using the refractive index

$$-n_1E\cos\theta + n_1E'\cos\theta = -n_2E''\cos\phi \qquad ...(4)$$

From equation 1 we know that $E'' = E + E'$

So $-n_1E\cos\theta + n_1E'\cos\theta = -n_2(E + E')\cos\phi$

Let $r_s = \frac{E'}{E}$

Then

$$-n_1 \cos\theta + n_1 r \cos\theta = -n_2(1 + r)\cos\phi$$

Getting *r* by itself

$$r_s(n_1 \cos\theta + n_2 r \cos\phi) = n_1 \cos\theta - n_2 \cos\phi$$

$$r_s = \frac{\cos\theta - n\cos\phi}{\cos\theta + n\cos\phi}$$

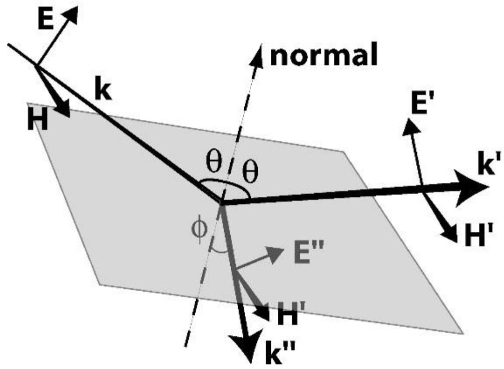


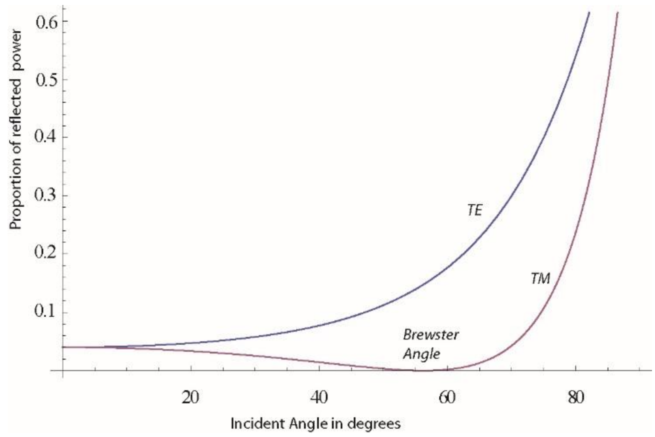
Figure 23 Reflection with transverse magnetic field

Following through the same maths with the transverse magnetic field gives

$$r_p = \left[\frac{E'}{E}\right]_{TM} = \frac{-n\cos\theta + \cos\phi}{n\cos\theta + \cos\phi}$$

The energy is proportional to the square of the electric field

$R_s = r_s^2$ and $R_p = r_p^2$



A Mathematica file with these equations will be stored along with these notes.

All the equations are given on page 44 of your textbook. On page 45 of your textbook you can see graphs for each case. Note that there is a special angle where the transverse magnetic field is not reflected. This is called the Brewster angle. Using equation 2.59 in your textbook we can show that, at the Bewster angle

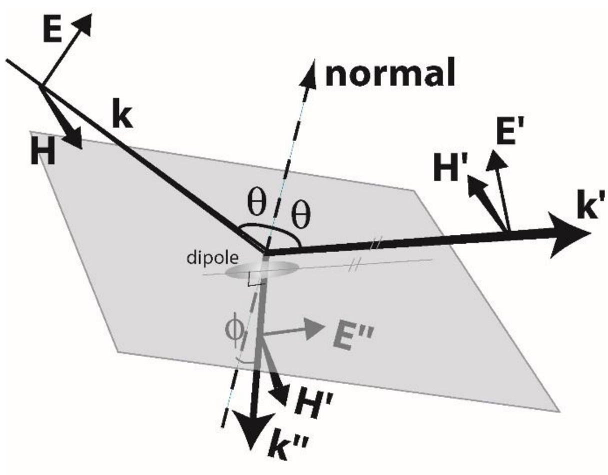
$$\theta = \tan^{-1}\frac{n_2}{n_1}$$

For glass with a refractive index of 1.5, $\theta = 57$degrees. The Bewster angle is extremely useful. You can use it to produce a polarized reflection (How would you do that?) but more often it is used to put an object in a polarized beam without creating a reflection.

For normal incidence the reflected power is

$$R = \left[\frac{n-1}{n+1}\right]^2$$

5.3. Why does Brewster's angle work

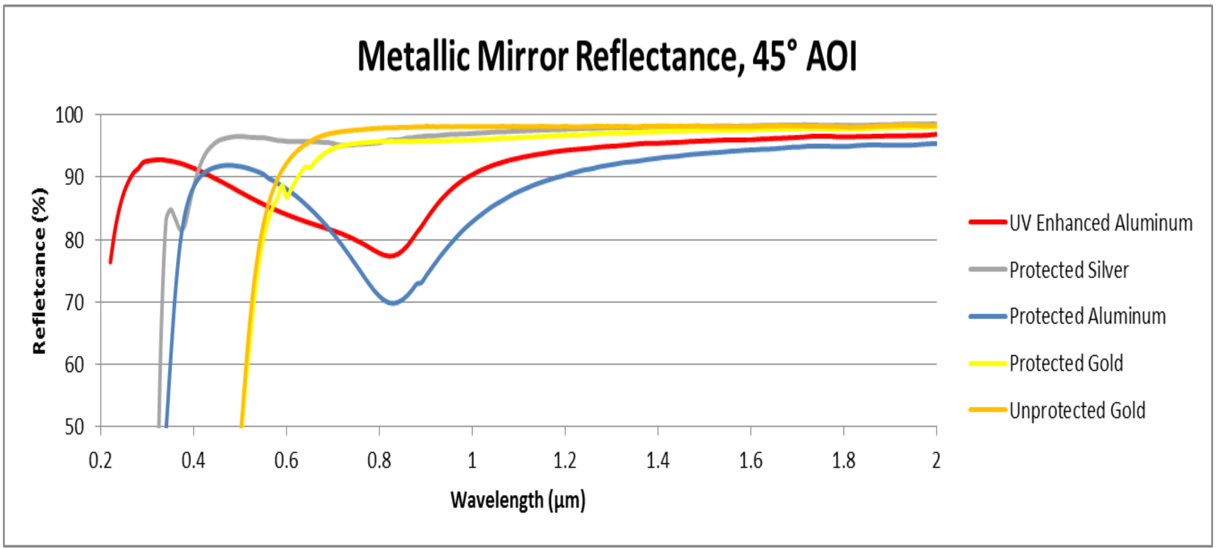


We can think of atoms and molecules as little antennas (or dipoles) where an electron can bounce back and forth, driven by the incoming wave. Since light (and radio waves) are transverse waves we always put our antenna (dipole) perpendicular to the wave (if it is end on then it can't absorb or transmit anything). In the case of a TM wave the molecules on the surface act like dipoles and oscillate perpendicular to the wave in the material. However there is an angle where the dipole lines up with the reflected wave and so can't transmit any power in that direction. That is Brewster's angle.

5.4. Mirrors, beam splitters etc.

5.4.1. Mirror coatings

Later we will study dielectric mirrors (built of stacks of thin films), but for reflecting a broad range of wavelengths it's hard to go past gold in the infra-red.



Adding a thin coating of protective dielectric material generally degrades the reflectivity by about 1% but makes it much easier to clean (bare gold can be cleaned with isopropyl and acetone – just be very gentle) .An "enhanced gold" mirror has a dielectric coating designed to improve the reflectivity over a range of wavelengths. At 1550 nm, better than 99% can be achieved. Gold is a hopeless reflector below 600nm so protected silver or protected aluminium should be used (or a dielectic mirror). Dielectric mirrors can give similar reflectivity but are always somewhat polarization sensitive.

5.4.2. Mirror Substrate

The light that isn't reflected is turned into heat. This can cause the underlying substrate to heat and the beam to wander. If this (and other) thermal issues are not a problem, then a fused silica substrate generally works very well. If it is a problem, then use a "Zerodur" substrate. Zerodur is a semi crystalline glass that has (as you would guess) very close to zero thermal expansion. It's not great for optics since the crystalline properties causes light to scatter, but it makes a great substrate material.

5.4.3. Curved Mirror or Lens

Curved mirrors are often a better choice than a lens for two reasons

1. No wavelength dependence (the refractive index of glass is very wavelength dependent)
2. Easy to support. Big lenses are limited in size because they distort over time, but mirrors can be supported from the rear.

5.4.4. Which type of mirror to use?

Most LCoS are coated with a protected aluminium coating allowing them to work well over a broad wavelength range including all visible wavelengths. For normal mirror however enhanced gold coatings are normally used.

5.5. Week 5 Tutorial Questions

5.5.1. Key Equations

Snell's law (angles to the perpendicular)

$$n_1 \sin\theta = n_2 \sin\phi$$

Transverse Electric:

$$r_s = \frac{\cos\theta - n\cos\phi}{\cos\theta + n\cos\phi} \quad , \quad R_s = r_s^2, \quad T + R = 1$$

Transverse Magnetic

$$r_p = \left[\frac{E'}{E}\right]_{TM} = \frac{-n\cos\theta + \cos\phi}{n\cos\theta + \cos\phi}$$

The Bewster angle

$$\theta = \tan^{-1}\frac{n_2}{n_1}$$

1. (do this in pairs so you have someone to hold the power meter!)
   a. If $\theta = 45°$ Calculate ϕ, the angle of refraction (assume $n_1 = 1, n_2 = 1.5$ and use Snell's law)

   b. Calculate $\cos\phi$ and $\cos\theta$

   c. Calculate the proportion of the power you would expect to be reflected from a vertically polarised beam (TE) hitting a vertical glass surface (RI n=1.5)? (remember you need to use $R_s = r_s^2$)

   d. Now calculate the proportion of the power you would expect for horizontally polarized light (TM).

e. Measure the optical power from your laser pointer (in mW) using the power meter:

f. Measure the power of the reflected light for TE (When held top up the laser is horizontally polarized but you want vertically polarized), remember to subtract off the background power from room lights:

g. What proportion of the power is this?

h. Now measure the power of TM light (in this case horizontally polarized)?

i. What proportion of the light is this?

j. How does this compare to your calculation?

2. A quick revision polarization question. Shine a polarized laser through the 3D movie glasses from the back and rotate the laser (make sure the laser isn't pointing at anyone, but towards a screen). What does the glass material appear to be? Now do the same thing from the front of the glasses. Use a polarizer to check the polarization of this light. What do you think the glasses consist of?

3. SMF-28 optical fiber has a refractive index of 1.5. What proportion of the light will be reflected back along the fiber from a disconnected flat (LC/PC say) connector?

(Use the equations

$$r_s = \frac{\cos\theta - n\cos\phi}{\cos\theta + n\cos\phi}$$

And $R_s = r_s^2$)

4. If I wanted to shine a beam of polarised light into water (n=1.3) without creating a reflection, then we know that the "reflected light" needs to be at 90 degrees to the transmitted light. So

$$\theta' = 90 - \theta$$

$\sin\theta' = \sin(90-\theta) = \cos\theta$ at Brewster's Angle, but $\frac{n_1}{n_2}\sin\theta = \sin\theta'$ from Snell's law

So $\frac{n_1}{n_2}\sin\theta = \cos\theta'$ or $\tan\theta = \frac{n_2}{n_1}$ at Brewster's angle

What angle would I cause the beam to enter the water to get no reflection? Use the fish-tank and see if it works

5. Which way would you orient the polariser in sunglasses if you wanted to block most of the reflected glare off water on the road, or a lake? Test your theory with the laser and some water and a polariser

Chapter 6. Total Internal Reflection

6.1. The Evanescent Wave

When light passes from a medium with higher refractive index to a medium with lower refractive index, there is a critical angle beyond which the light is completely reflected. From Snell's law

$$n_1 \sin\theta_1 = n_2 \sin\theta_2$$

When $\theta_2 = 90°$

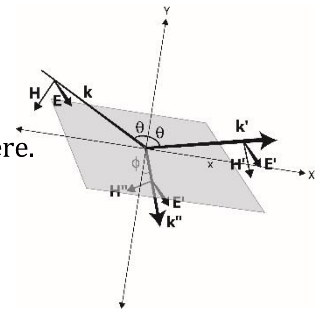
$$n_1 \sin\theta_1 = n_2$$

$$\therefore \qquad \theta_1 = \sin^{-1}\frac{n_2}{n_1}$$

But from last week's discussion, the field at the boundary must be continuous. So what's happening…

Consider the transmitted wave

$$\boldsymbol{E}_{trans} = \boldsymbol{E}'' e^{i(\boldsymbol{k}''.\boldsymbol{r}-\omega t)}$$



For simplicity consider the reflection in the XY plane as shown here.

The $\boldsymbol{k}''.\boldsymbol{r}$ term can be expanded (letting $\boldsymbol{r} = (x, y, z)$and $n = \frac{n_2}{n_1}$)

$$\boldsymbol{k}''.\boldsymbol{r} = k''x\sin\phi + k''|y|\cos\phi$$

$$= k''x\sin\phi + k''|y|\sqrt{1-\frac{\sin^2\theta}{n^2}}$$

$$= \frac{k''}{n}x\sin\theta + ik''|y|\sqrt{\frac{\sin^2\theta}{n^2}-1}$$

So now the component of the transmitted field along the surface is sinusoidal but the component perpendicular is exponentially decreasing.

$$\boldsymbol{E}_{trans} = \boldsymbol{E}'' e^{-k''|y|\sqrt{\frac{\sin^2\theta}{n}-1}} e^{i(\frac{k''}{n}x\sin\theta-\omega t)}$$

6.2. The Phase Shift on TIR

Since the wave doesn't stop at the boundary but penetrates slightly into it, it's not surprising that total internal reflection creates a different phase shift in the TE and TM waves. We can rewrite our reflection equations

$$r_p = \left[\frac{E'}{E}\right]_{TM} = \frac{-n\cos\theta + \cos\phi}{n\cos\theta + \cos\phi}$$

Using Snell's law this becomes

$$= \frac{-n\cos\theta + \sqrt{1 - \frac{\sin^2\theta}{n^2}}}{n\cos\theta + \sqrt{1 - \frac{\sin^2\theta}{n^2}}}$$

$$= \frac{-n^2\cos\theta + \sqrt{n^2 - \sin^2\theta}}{n^2\cos\theta + \sqrt{n^2 - \sin^2\theta}}$$

$$= \frac{-n^2\cos\theta + i\sqrt{\sin^2\theta - n^2}}{n^2\cos\theta + i\sqrt{\sin^2\theta - n^2}}$$

And similarly

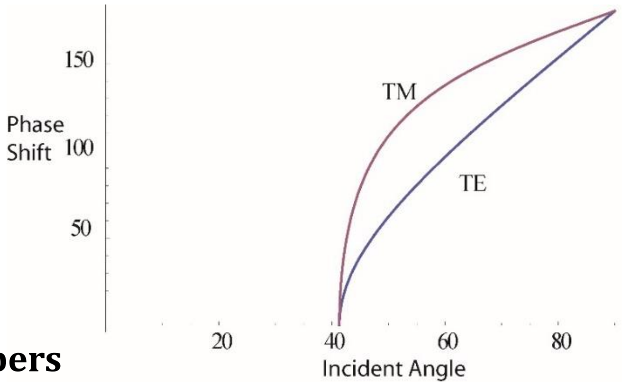
$$r_s = \frac{\cos\theta - n\cos\phi}{\cos\theta + n\cos\phi}$$

$$= \frac{\cos\theta - \sqrt{n^2 - \sin^2\theta}}{\cos\theta + \sqrt{n^2 - \sin^2\theta}}$$

$$= \frac{\cos\theta - i\sqrt{\sin^2\theta - n^2}}{\cos\theta + i\sqrt{\sin^2\theta - n^2}}$$

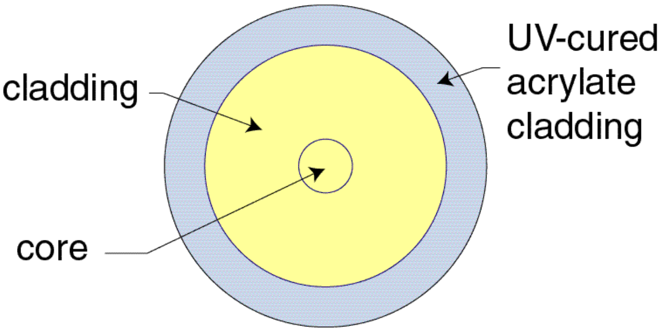
The numerator and the denominator are simply complex conjugates of each other so these are both numbers with magnitude 1 and an associated phase shift. It takes a page of algebra (and the trig identity $\tan(a-b) = {(\tan a - \tan b)}/{(1 - \tan a \tan b)}$) to calculate the relative phase shift between TM and TE components as

$$\Delta = 2\tan^{-1}\left(\frac{\cos\theta\sqrt{\sin^2\theta - n^2}}{\sin^2\theta}\right)$$



6.3. Optical Fibers

We often use total internal reflection to change the direction of light using a prism, but the most common use is in an optical fiber. In an optical fiber



An optical fiber consists of a core of glass, a cladding of slightly lower refractive index glass and a plastic coating. Normally the core is nearly pure SiO_2 (silica glass) with a small amount of added germanium to increase the refractive index. The cladding is typically also nearly silica but sometimes includes a small amount of fluorine to lower the refractive index. Fluorine is a highly mobile ion and during fusion splicing will tend to migrate into the core. This can be used to make fibers with adjustable core sizes in the splicing region (rearcing the splice increases the effective core size). UNA-1 from Nufern (now part of Coherent) is an example of such a fiber, but there are plenty of others.

The most common fiber in the world is SMF-28 which has a core size of about 9 microns. This is a "single mode" fiber which means that the shape of the electric field distribution across the core is fixed (and other energy leaks away – more about this in a later talk).

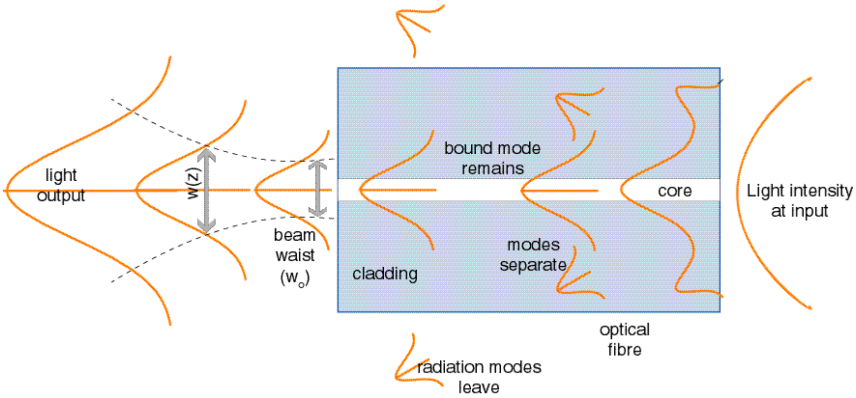


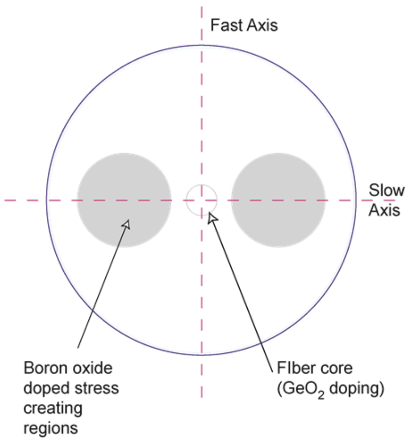
Table 4 Typical Data for SMF-28 (or equivalent) optical fiber

Single mode wavelength range	1260 - ~1640 nm

Mode field diameter @ 1550nm	10. 4 μm
Dispersion slope	0.086 ps/(nm^2.km) *zero @ 1313nm*
Typical loss	0.19 dB/km (2%/km)
Bending loss	<0.1 dB, 100 turns 25 mm radius
PMD	0.1 ps/√km

However, for small devices a low-bending loss fiber is often preferable. A fiber such as Corning Clear-Curve fiber can be bent to extremely small radii (e.g around a pencil) with almost no loss. This is generally done by creating a "W" refractive index profile, effectively creating a second total internal reflection, folding the light back into the core.

6.4. Polarization Persevering and Polarizing Optical Fibers

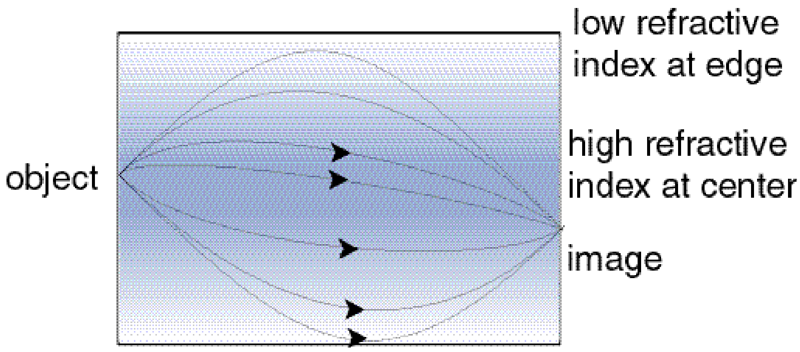


PM fibers are highly birefringent. This birefringence is typically created by embedding two cylinders of boron doped glass in the fiber cladding. As the glass cools from around 1000°C, this boron-doped glass shrinks more than the surrounding glass creating large stresses across the fiber core. The resulting stress birefringence creates a full waveplate approximately every millimetre around the fiber (the fiber is said to have a "beat length" of 1mm).

The direction aligned with the stress rods is the slow axis of the fiber. While you can use either axis, if the electric field of the light is aligned along this axis it will be better confined in the fiber core. In fact there is a range of wavelengths for a PM fiber where the light is only confined in along the slow axis. Light in the other axis leaks out. When the fiber is used in this way it's referred to as "polarising fiber" (there are other sorts of polarising fibers that involve putting metal films near the fiber core).

6.5. Selfoc or GRIN or "G" lens

We can further expand this concept by reducing the refractive index gradually as we get further from the center of the fiber. This is done in "multi-mode" fibers to reduce mode-dispersion. It's also done in GRIN lens (often called "G lens") to make a lens that has flat ends. These can be cut to any length to provide a focused, collimated or diverging beam.



Tutorial For Week 6

6.6. Key Equations

Critical Angle

$$\therefore \qquad \theta_1 = \sin^{-1}\frac{n_2}{n_1}$$

Relative Phase Shift on TIR (TM is slowed), so TM component is multiplied by $e^{-\Delta}$

$$\Delta = 2\tan^{-1}\left(\frac{\cos\theta\sqrt{\sin^2\theta - n^2}}{\sin^2\theta}\right)$$

(the effective penetration depth (1/e value) is about half a wavelength at 45 degrees n=1.5 – this also results in a tiny beam offset. The TM component penetrates further)

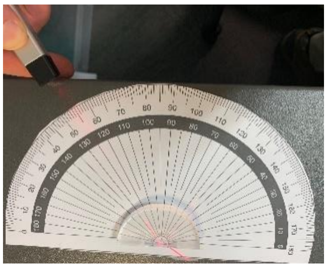
6.7. Questions

1.
    a. Use the fish tank to measure the angle for total internal reflection for a water – air interface (you'll need to angle the beam up so it hits the air/water boundary from below).

    b. Calculate the angle assuming n=1.33.

2.
    a. Use the laser pointer and plastic half circle to find the critical angle.

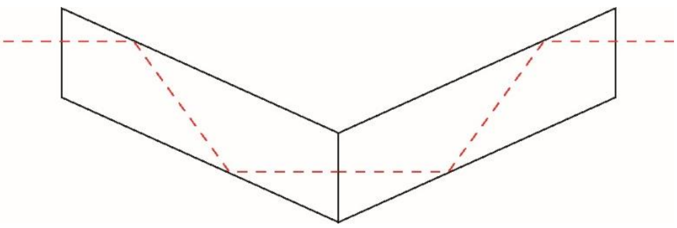
    b. Use this to estimate the refractive index of the plastic.



3. A waveplate with low wavelength sensitivity, and very little angular sensitivity, can be made using a piece of glass cut as shown. If we use a glass with refractive index of 1.5 (i.e. n=0.66) and the light hits the surfaces with an angle of 74.74 degrees, what will the total phase shift be? (this is called a double Fresnel Rhomb)

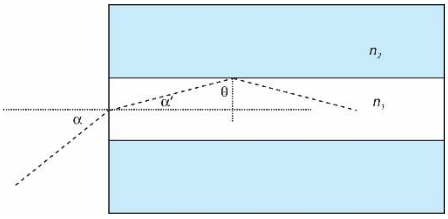
    Use the equation

$$\Delta = 2\tan^{-1}\left(\frac{\cos\theta\sqrt{\sin^2\theta - n^2}}{\sin^2\theta}\right)$$



4. Show that the acceptance angle for an optical fiber is given by

$$\alpha = \sin^{-1}\sqrt{(n_1^2 - n_2^2)}$$



Start with Snell's law, at TIR $\sin\theta = \frac{n_2}{n_1}$, then work back to figure out the angle the light refracts to when it enters the fiber (assume that the refractive index of air is 1 for simplicity). Note that $\cos\alpha' = \sqrt{1 - \sin\alpha'^2}$.

Chapter 7. Superposition

7.1. Concept

The principle of superposition is very simple. When light waves interact, it is the electric (and magnetic) fields that add, not the intensity. So if we add two waves

$$\boldsymbol{E}_{(1)} = \boldsymbol{E_1} e^{i(\boldsymbol{k}''.\boldsymbol{r}-\omega t)} \text{ and } \boldsymbol{E}_{(2)} = \boldsymbol{E_2} e^{i(\boldsymbol{k}''.\boldsymbol{r}-\omega t)}$$

Then the resulting intensity is

$$\begin{aligned} I = \boldsymbol{E}.\boldsymbol{E}^* &= \left(\boldsymbol{E}_{(1)} + \boldsymbol{E}_{(2)}\right).\left(\boldsymbol{E}_{(1)} + \boldsymbol{E}_{(2)}\right) \\ &= \boldsymbol{E}_{(1)}\boldsymbol{E}_{(1)}{}^* + \boldsymbol{E}_{(2)}\boldsymbol{E}_{(2)}{}^* + \boldsymbol{E}_{(1)}\boldsymbol{E}_{(2)}{}^* + \boldsymbol{E}_{(1)}{}^*\boldsymbol{E}_{(2)} \\ &= I_1 + I_2 + 2\,|\boldsymbol{E_1}.\boldsymbol{E_2}|\cos\theta \end{aligned}$$

Notice that $\boldsymbol{E_1}.\boldsymbol{E_2}$ is a dot product. In particular, if the polarisations are orthogonal there is no interference term.

Figure 5 shows Young's 2-Slit experiment. Assuming the screen is far away, calculate the angle of the first order fringe.

This is exactly the same principle that we use to form a grating, which we will discuss in a few weeks' time.

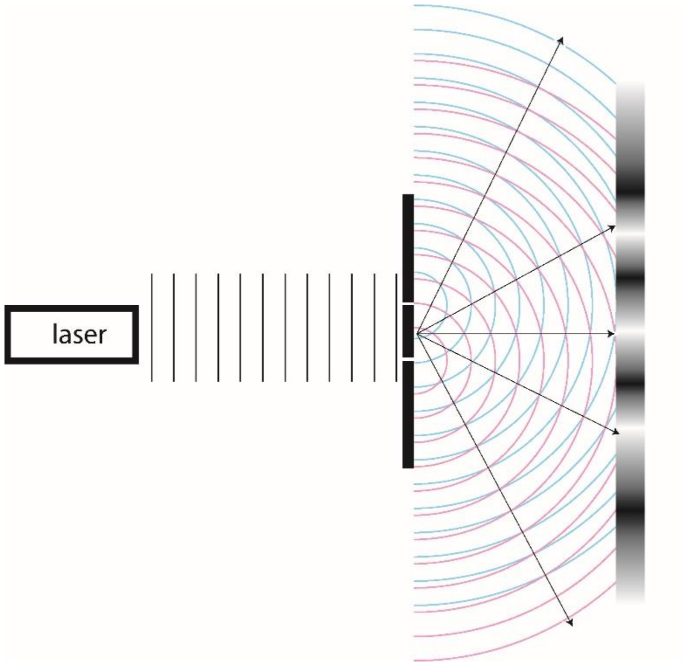


Figure 24 Young's 2 slit experiment

If the screen is far away and the angle is small then $\sin\boldsymbol{\theta} \approx \boldsymbol{\theta}$ (in radians) and we can show that we'll get bright fringes at

$$\boldsymbol{y} = \boldsymbol{0}, \pm\frac{\boldsymbol{\lambda x}}{\boldsymbol{h}}, \pm\frac{\boldsymbol{2\lambda x}}{\boldsymbol{h}}, \pm\frac{\boldsymbol{3\lambda x}}{\boldsymbol{h}}, \ldots$$

Where ***x*** is the distance to the screen and ***h*** is the distance between the slits.

7.2. Michelson Interferometer

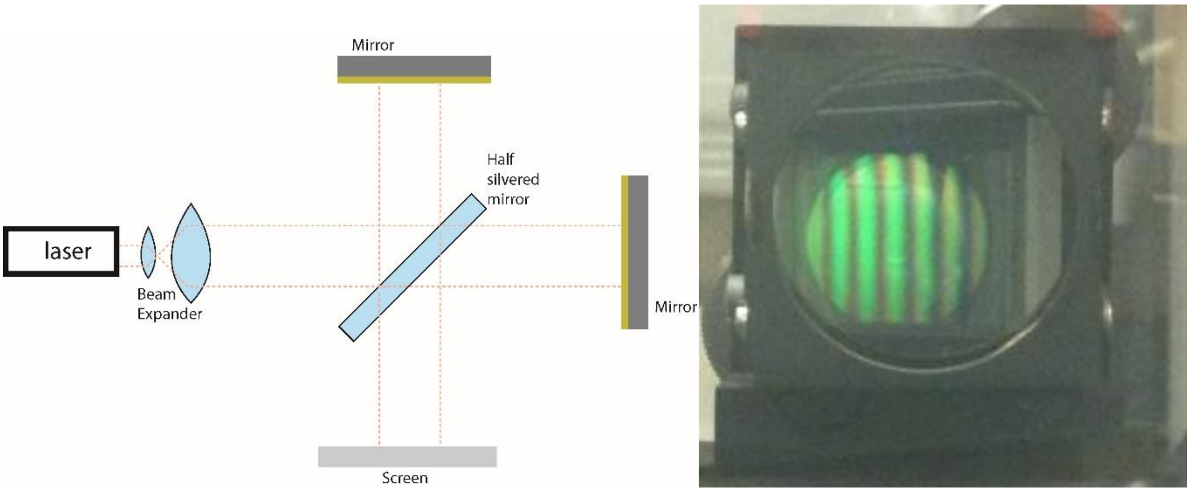
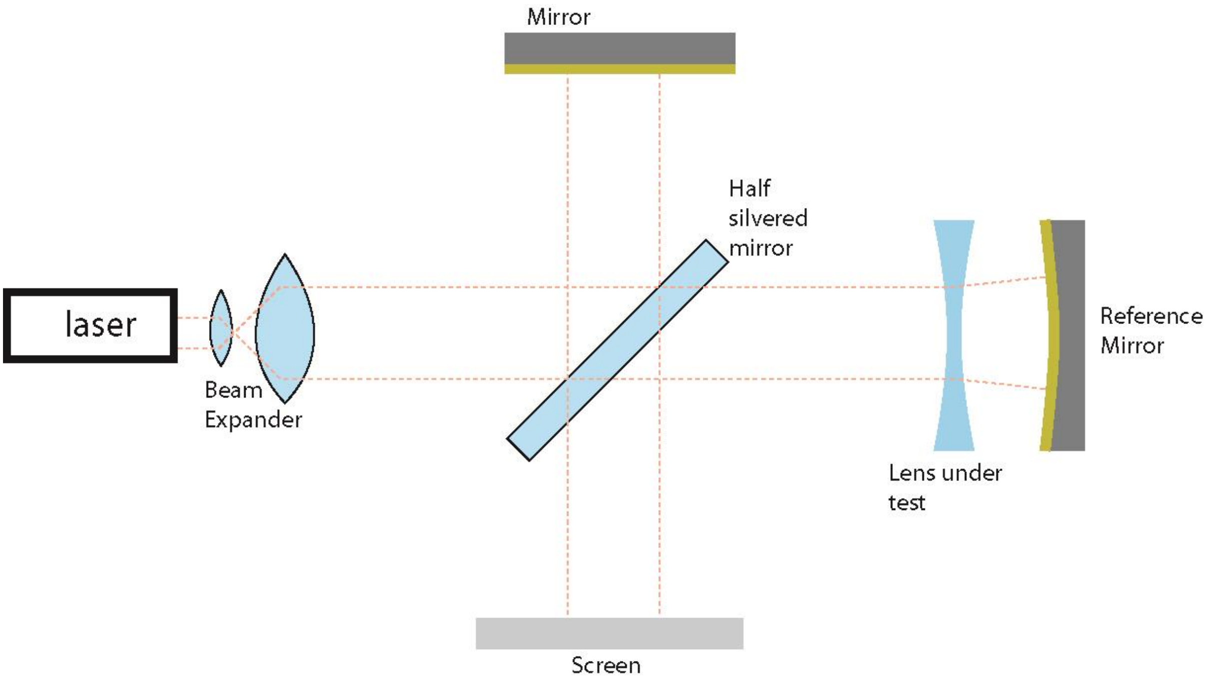


Figure 25 Michelson Interferometer and a typical image (image from David Van Buren CSULA)

Figure 6 shows a typical Michelson interferometer. Adding an optical element into the path, typically a compensating mirror (the Twyman-Green modification) allows us to map out imperfections in the optical element under test.



7.3. Coherent Channel Monitors and Superposition

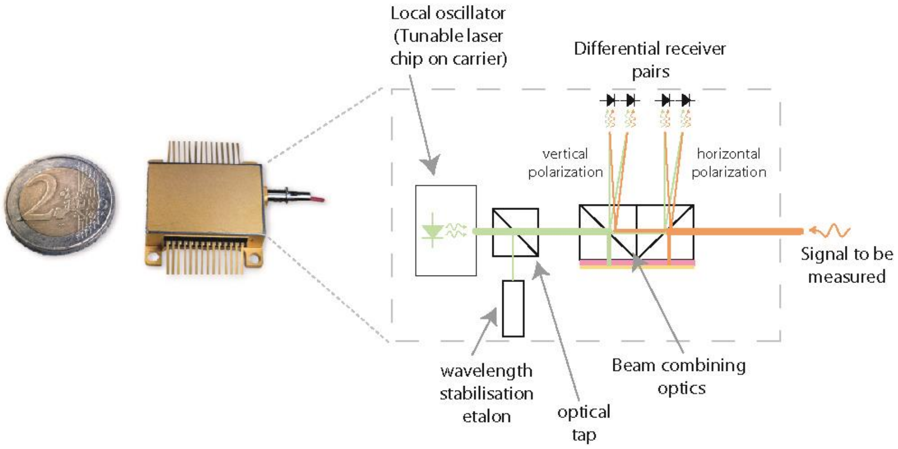


Figure 26 How a coherent channel monitor works (from 2014 OFC paper W4D.7)

The idea behind a coherent channel monitor is that we superimpose the signal from a tunable laser and the signal we want to measure. The two signals will beat together. If the difference in frequency (and hence the beat signal) is within the bandwidth of the receiver (30MHz) then rather an seeing a simple DC current from the photodiodes we'll see an AC signal corresponding to the frequency difference between the signal and the local laser (normally referred to as the local oscillator). This is very similar to how FM radio signals are detected – but at a much higher frequency. By subtracting the signal from the same signal where the local oscillator has a half wave delay, and then only looking at the AC component of the signal, we can remove everything but the beat pattern. By looking at the amplitude of the beat pattern (envelope demodulation) we can read off the power of the signal at that frequency. In practice the local oscillator (laser) covers a range of frequencies about 250MHz wide so the $f_o \pm (B_e + B_o/2)$, where $\mathrm{B_e}$ is the electrical bandwidth of the receiver (30 MHz), and $\mathrm{B_o}$ the optical bandwidth of the local oscillator (~250 MHz).

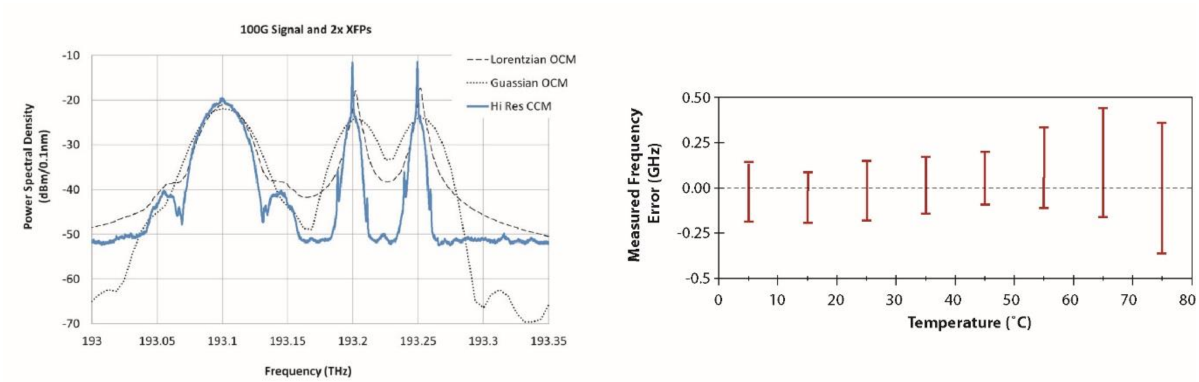


Figure 27 A coherence channel monitor gives a really sharp spectrum (~1GHz resolution) and sub GHz accuracy, but the noise floor is higher than for an OSA (from figures 2&3 OFC paper W4D.7)

7.4. Partial Coherence, coherence time and coherence length

We previously write

$$I = \boldsymbol{E}.\boldsymbol{E}^* = \left(\boldsymbol{E}_{(1)} + \boldsymbol{E}_{(2)}\right).\left(\boldsymbol{E}_{(1)} + \boldsymbol{E}_{(2)}\right)$$

But in-fact this is a time averaged value.

$$I = \langle \boldsymbol{E}.\boldsymbol{E}^{*} \rangle = \langle (\boldsymbol{E}_{(1)} + \boldsymbol{E}_{(2)}).(\boldsymbol{E}_{(1)} + \boldsymbol{E}_{(2)}) \rangle$$
$$= I_1 + I_2 + 2\, Re\langle \boldsymbol{E}_{(1)}.\boldsymbol{E}_{(2)} \rangle$$

The time average is just the function

$$\langle f \rangle = \lim_{T\to\infty} \frac{1}{T} \int_0^T f(t)dt$$

When we interfere a beam with itself, the time over which the time displaced beams have a phase relationship with each other $\tau_o = {}^{1}/_{\Delta f}$. So the distance (in a vacuum) over which they have a relationship is $l_c = {}^{c}/_{\Delta f}$. In practice the fringe amplitude drops gradually and depends on the shape of the frequency distribution, but for a simple rectangular distribution, the coherence length corresponds to the point at which the fringe visibility first becomes zero.

7.5. Spatial Coherence Length

When we have a broad beam, it's sometimes necessary to know how far across the beam you can go before the wavefront becomes incoherent. A point source, such as light from a single mode fibre is always coherent across the beam. The light from a wider source, such as the sun (angular width of 0.00925 radians) is coherent over a less than 0.1mm. The general formula is

$$l_t = \frac{1.22\lambda}{\theta_s}$$

Where θ_s is the angular width of the source in radians.

7.6. A quick diversion on Fourier Series

Data communications happens in time, but most aspects of our Wavelength Selective Switch (WSS) are best understood in terms of frequency so it's very useful to be able to move between these domains. The simplest translation, we've already done. A sin wave in time is a point function in frequency space. It's probably easiest to go next to a Fourier series. The idea here is that essentially any periodic function over an interval 2π can be represented as

$$f(x) = \frac{a_0}{2} + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + a_3 \cos 3x + b_3 \sin 3x + \dots$$

Where

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)dx$$

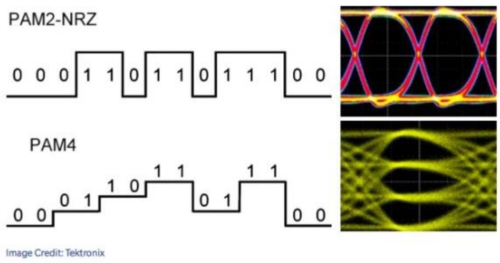
$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx\, dx$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx\, dx$$

The data we transmit looks something like a square wave, so it's worth looking at the Fourier series for a square wave.

7.7. Sending Data through a fibre PAM and QAM

First some definitions – a 10 Gbaud/s is 10 billion "symbols" per second. The baud rate determines the minimum frequency space the signal can occupy. Each "symbol" can carry several "bits" of data. A PAM-4 symbol has four levels defined which correspond to 00,01,10 and 11. Data centers like PAM-4 or PAM-6 signals because they are cheap to detect.



However we aren't limited to just using the amplitude or brightness of the signal. We can also use the polarisation to double the bits per symbol. We can also use the phase. If we graph phase as an angle and amplitude as distance from the origin then we can define a 16QAM signal for example as shown in Figure 5.

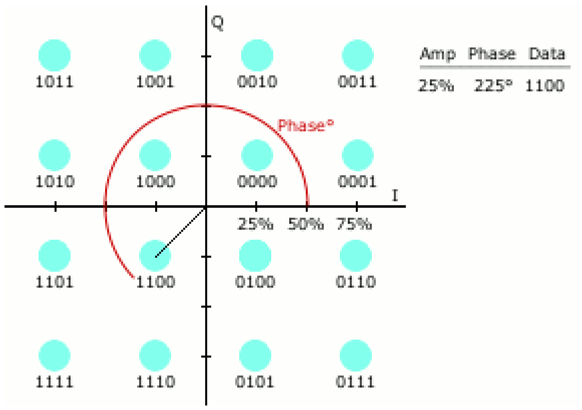


Figure 28 16 QAM

By Chris Watts - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=15781908

7.8. Datacom

Data centres next to move huge amounts of data between racks of computers that are never more than ~2km apart. The data rate in each fiber is governed by how much data the switch on the top of the rack can handle. Up to 200,000 fibers can be going between different buildings in a data-center complex. No surprisingly, cost is the key consideration.

Figure 29 Picture of 10Gb/s XFP transmitters from the (old) Finisar website

At the moment the common speeds are 10 Gb/s, 40 Gb/s and 100 Gb/s. The next generation will use 200 Gb/s and 400 Gb/s. To do this, they will use PAM-4 transmitters and receivers, and probably multiple wavelengths. At the moment all the transmitters are little plug-in devices

7.9. Questions for Week 7

1. Consider Young's two slit experiment. If the screen is 2m away from the slits a wavelength of 600nm is used and a 1 mm fringe spacing is required, what slit separation should be used?

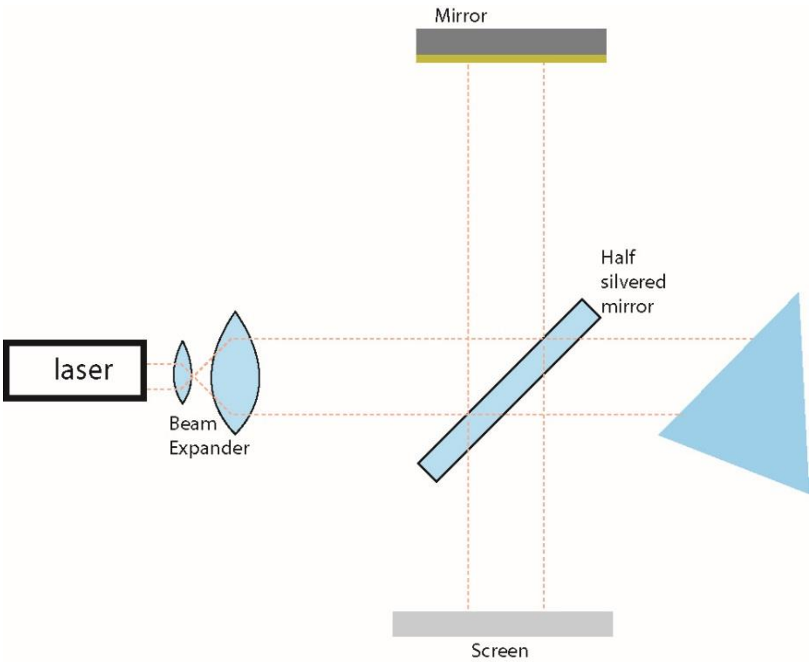
2. Young didn't really use two slits in his famous experiment; he simply placed a very thin piece of card in a beam of sunlight. We can do the same with a laser. Let's use a 630nm laser to measure the diameter of a hair. Measure the spacing of a number of fringes (5 or 6 at least) and use that measurement to calculate the diameter of the hair. Then use a pair of Vernier callipers to measure the diameter of the hair and compare your results. Measuring multiple fringes makes the reading more accurate (think about why you might get different answers from the two measurement methods, and which one will be right).

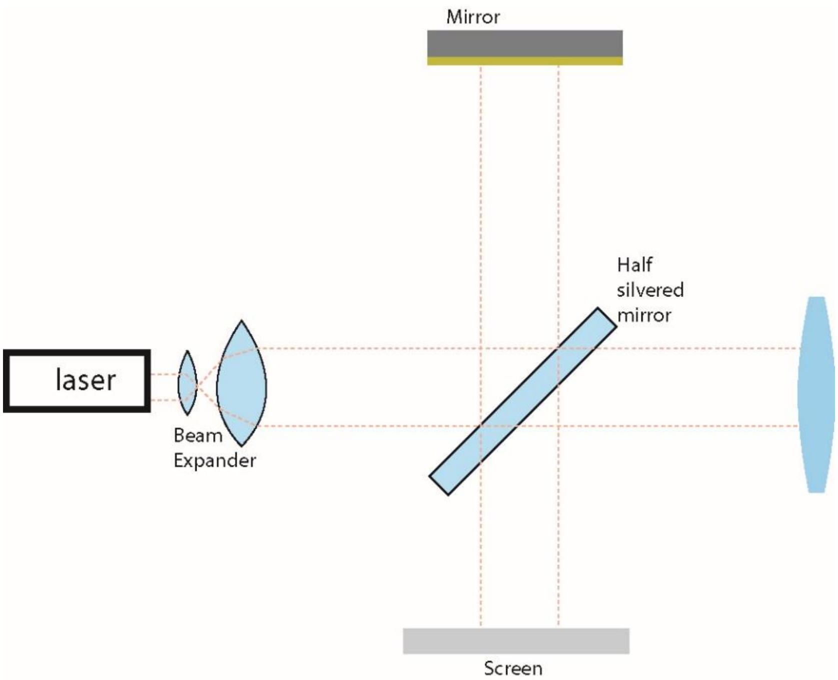
3. We can measure the refractive index of a gas by placing an evacuated gas cell in a Michelson interferometer and then counting fringes as the gas in slowly introduced. If we have a 10cm gas cell and SF_6 gas (refractive index 1.013), how many fringes would we count (assume a 522nm laser)?

4. Consider this image from a Michelson Interferometer. If the green laser has a wavelength of 522nm, what is the path difference between the left and right sides of the image? If the fringes weren't present before a plate under test was put in the beam, what is the variation in thickness across the plate assuming a refractive index of 1.5?

5. If a 1550nm laser beam has a wavelength spread of 100MHz, what is its coherence length?

6. How would you make a Twyman-Green interferometer to check the quality of a prism used in transmission? Add the mirror to the diagram…
   Then do the same for a convex lens using the second diagram.


Mirror
Half
silvered
mirror
laser
Beam
Expander
Screen


Mirror
Half
silvered
mirror
laser
Beam
Expander
Screen

Chapter 8. Optics Communications Specifications – Amplitude, Polarization and Phase

At this point it's worth reviewing how the optics we have learnt relates to the specifications of some of the products we make, and some of the different ways that telecommunications engineers describe the same quantities.

8.1. Power and Loss

Rather than talk about power in Watts, most telecommunications engineers speak in dBm (decibels relative to a mW). The relationship is simply

$$P_{dBm} = 10 * \log_{10}[P_{mW}]$$

Or

$$P_{mW} = 10^{P_{dBm}/10}$$

So using this information, fill in the following table

Power in mW	**Power in dBm**
1mW	
2mW	
4mW	
	10dBm
	20dBm
	23dBm
	-10dBm
	-20dBm
	-23dBm

Rather than talk about loss as a percentage, telecommunications engineers speak in dB (decibels). The relationship is simply

$$L_{dB} = 10 * \log_{10}\left[\frac{P_{in}}{P_{out}}\right]$$

Or in linear units

$$P_{out} = P_{in} * 10^{-L_{dB}/10}$$

This lets us simply subtract the losses in dB from the power in dBm to calculate the output power. For example a 10 dBm transmitter going through a link with a loss of 15dB will give -5dBm at the receiver. Power and loss calculations become very simple. Note that -3dB is very close to 50% and 1dB is close to 20%.

8.2. Group Velocity and Group Delay

In a WSS, the critical parameter in frequency space for us is the “clear channel pass-band”; the frequency range over which our customers can depend on our specifications. In the time domain, the critical parameter is the symbol spacing. For a 10 Gb/s On-Off keyed signal, the symbol spacing is $\frac{1}{10\times10^9} = 100$ ps.

So far we have talked about “group velocity”. In telecommunications people generally talk about “group delay ripple” and “chromatic dispersion”. They are closely related. Group velocity is the speed of the pulse ($v_g = \frac{\partial\omega}{\partial k}$). Group delay ($\tau$) is the relative time that the pulse arrives (the distance divided by the group velocity) . Chromatic dispersion is the rate at which the group delay changes as a function of wavelength $D = \frac{d\tau}{d\lambda}$ and is measured in ps/nm.

Table 5 Excerpts from Typical WSS Specifications

<table>
<tr><th>DESCRIPTION</th><th>UNIT</th><th colspan="3">SPECIFICATION</th><th colspan="2">NOTES</th></tr>
<tr><td>GENERAL</td><td></td><td colspan="3"></td><td colspan="2"></td></tr>
<tr><td>Channel Spacing</td><td>GHz</td><td colspan="3">50</td><td colspan="2">As per ITU G.694.1 specification</td></tr>
<tr><td>No. of channels</td><td></td><td colspan="3">88</td><td colspan="2"></td></tr>
<tr><td>FREQUENCY RANGE</td><td></td><td>MIN</td><td>TYP</td><td>MAX</td><td colspan="2"></td></tr>
<tr><td>Frequency Window</td><td>THz</td><td>191.725</td><td></td><td>196.125</td><td colspan="2"></td></tr>
<tr><td>BANDWIDTH</td><td></td><td>BOL</td><td>Δ</td><td>EOL</td><td colspan="2"></td></tr>
<tr><td>Clear Passband*</td><td>GHz</td><td>± 13.2</td><td>0.7</td><td>± 12.5</td><td colspan="2">Centered on ITU grid</td></tr>
<tr><td rowspan="2">0.5 dB Bandwidth*</td><td>GHz</td><td>± 14.7</td><td>0.7</td><td>± 14.0</td><td>Attn range: 0-15 dB</td><td rowspan="2">Centered on ITU grid;</td></tr>
<tr><td></td><td>± 13.7</td><td>0.7</td><td>± 13.0</td><td>Attn range: 15.1–20.0 dB</td></tr>
</table>

<table>
<tr><th colspan="2">Description</th><th>Unit</th><th colspan="3">Specification</th><th colspan="2">Notes</th></tr>
<tr><td colspan="2">3.0 dB Bandwidth*</td><td>GHz</td><td>± 18.2</td><td>0.7</td><td>± 17.5</td><td colspan="2">Centered on ITU grid</td></tr>
<tr><td colspan="2">Extinction Clear Bandwidth</td><td>GHz</td><td>± 13.2</td><td>0.7</td><td>± 12.5</td><td colspan="2">centered on ITU grid</td></tr>
<tr><th colspan="2">Loss</th><th></th><th>BOL</th><th>Δ</th><th>EOL</th><th colspan="2"></th></tr>
<tr><td colspan="2">Max. Insertion Loss</td><td>dB</td><td>6.2</td><td>0.3</td><td>6.5</td><td colspan="2">Including connectors</td></tr>
<tr><td colspan="2">Insertion Loss Uniformity*</td><td>dB</td><td>0.95</td><td>0.05</td><td>1.0</td><td colspan="2"></td></tr>
<tr><td colspan="2" rowspan="3">Polarization Dependent Loss*</td><td rowspan="3">dB</td><td>0.75</td><td>0.1</td><td>0.85</td><td>Attenuations 0-15 dB</td><td rowspan="3">Over Clear Passband</td></tr>
<tr><td>0.8</td><td>0.1</td><td>0.9</td><td>Attenuations 15-20 dB</td></tr>
<tr><td></td><td></td><td>0.3</td><td>Typical value</td></tr>
<tr><th colspan="2">Loss</th><th></th><th>Min</th><th>Typ</th><th>Max</th><th colspan="2"></th></tr>
<tr><td colspan="2">Return Loss</td><td>dB</td><td>30</td><td></td><td></td><td colspan="2"></td></tr>
<tr><th colspan="2">Group Delay Ripple</th><th></th><th>Min</th><th>Typ</th><th>Max</th><th colspan="2"></th></tr>
<tr><td rowspan="3">Group Delay Ripple</td><td>Clear Passband*</td><td>ps</td><td>-0.75</td><td></td><td>+0.75</td><td colspan="2" rowspan="3"></td></tr>
<tr><td>± 14 GHz*</td><td>ps</td><td>-1.5</td><td></td><td>+1.5</td></tr>
<tr><td>± 16 GHz*</td><td>ps</td><td>-2.5</td><td></td><td>+2.5</td></tr>
<tr><th colspan="2">Dispersion</th><th></th><th>BOL</th><th>Δ</th><th>EOL</th><th colspan="2"></th></tr>
<tr><td colspan="2">First Order PMD (DGD)*</td><td>ps</td><td>0.5</td><td>0.05</td><td>0.55</td><td colspan="2"></td></tr>
<tr><td colspan="2">Chromatic Dispersion*</td><td>ps/nm</td><td>±7</td><td>3</td><td>±10</td><td colspan="2"></td></tr>
</table>

A typical chromatic dispersion specification is shown at the end of Table 1.

Question: For a 100GHz wide channel, what delay does a chromatic dispersion of 7ps/nm correspond to across the band? What fraction of a 30 Gb/s pulse separation would this correspond to?

The “group delay ripple” is how far the group delay diverges from the simple linear change captured by the chromatic dispersion.

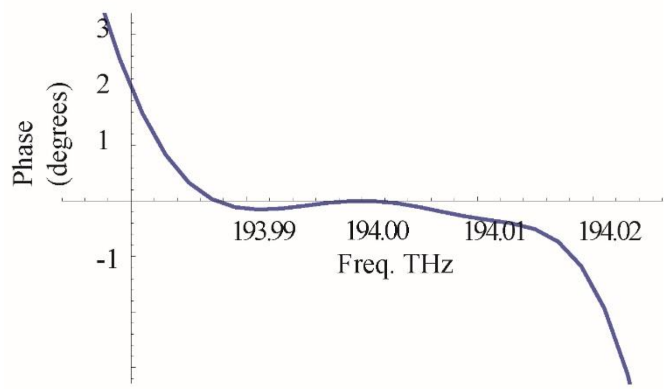


Figure 30 As phase modulation becomes important, phase errors need to be considered. Here's a typical phase error across a WSS channel. The phase error is found by integrating the group delay ripple. Notice that it is quite flat in the middle of the channel but becomes large at the sides

The group delay is polarisation dependent. The time separation of the group delays for different polarisations is called the "differential group delay" or the "first order Polarisation Mode Dispersion". What fraction of the pulse separation for a 50 GBaud signal does a PMD of 0.5 ps use up?

Modern modulation techniques modulate both amplitude and phase to fit in more bits per symbol. A QPSK signal just modulates phase, while a 16-QAM signal modulates amplitude to three distinct amplitudes and 12 possible phases. Most 16-QAM systems also modulate both polarisations separately giving a total of 8 bits/symbol.

If we allow a phase budget of ±10°, what proportion of this budget will the WSS use up?

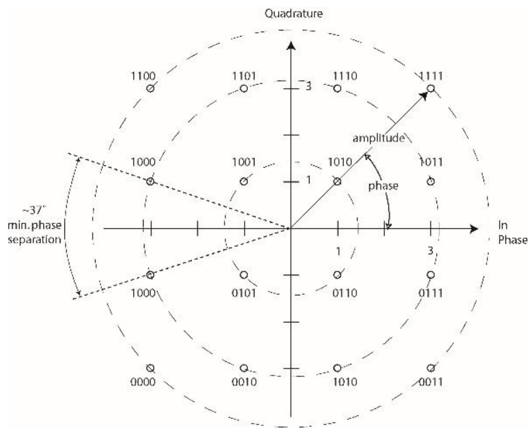


Figure 31 A 16-QAM transmitter has an amplitude separation of distinct values of 66% or a phase separation of at least 25°

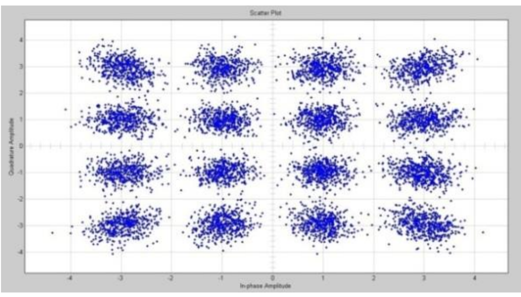


Figure 32 What the output of a 16 QAM system looks like (from "All-Optical Signal Processing for High Spectral Efficiency (SE) Optical Communication" Y.Ben Ezra et. Al. Book chapter in "Optical Communications", Y. Ben Ezra et al. ISBN: 978-953-51-0784-2

8.3. Polarisation Dependent Loss

So far we have talked about polarisers and birefringent elements such as waveplates. Complex devices such as WSS unfortunately combine the properties of both. The PDL is the maximum loss for any polarisation state minus the minimum loss for any state. These two states will always be orthogonal (but they can be any polarisations). If there are separate signals being transmitted on two orthogonal polarisation states, then PDL will have the effect of *mixing* these signals.

A 0.5 to 0.8dB PDL specification is typical for a WSS but this specification is under a lot of pressure from customers. Being a property of the optics alone (except for an EWP4), it's also an expensive specification to tighten.

8.3.1. Some notes on measuring PDL

Polarization dependent loss is simply when one orthogonal state is attenuated relative to another. The simplest way to measure this is to measure the power transmitted through the device after setting the polarization to each of the six states on each of the axes (vertical, horizontal, +45°,-45°, RCP and LCP). Let's call them P_A, P_a, P_B, P_b, and P_C, P_c respectively.

The average power is

$$P_{ave} = \frac{P_A + P_a}{2} = \frac{P_B + P_b}{2} = \frac{P_C + P_c}{2}$$

The relative power lost is

$$P_{A:PDL} = |P_A - P_a| / P_{ave}$$

$$P_{B:PDL} = |P_B - P_b| / P_{ave} = 2|P_B - P_{ave}| / P_{ave}$$

$$P_{C:PDL} = |P_C - P_c| / P_{ave} = 2|P_C - P_{ave}| / P_{ave}$$

The total PDL (in linear units) is simply

$$P_{PDL} = \sqrt{{P_{A:PDL}}^2 + {P_{B:PDL}}^2 + {P_{C:PDL}}^2}$$

Converting this to logarithmic units

$$PDL_{in\ dB} = -10\log_{10}(P_{PDL})$$

Note that waveplate based polarization controllers have different losses in each polarization state so a reference reading needs to be taken regularly for each generated polarization state, and the measured values normalised. Notice that you don't actually need to take all 6 measurements, you could just take 4, but any error P_{ave} gets repeated in each term. The 6-point method gives about half the error at the cost of a longer measurement time.

8.4. Optical Channel Monitors

Here's a specification from our coherent optical channel monitor. The most critical parameters here are frequency and power accuracy and noise floor. In an OSA the power is typically the power in a 0.1nm region. In the WaveAnalyzer it's calculated over 150 MHz, and for the coherent OCM it's given for a 12.5 GHz slot.

ID	DESCRIPTION	UNIT	EOL VALUE		NOTES
2.1.1.1	General		Min	Max	
2.1.1.2	Number of ports		1		
2.1.1.3	Channel Spacing	GHz	flexible grid		As per ITU G.694.1 specification
	Frequency Range		Min	Max	
2.1.1.4	Frequency Window	THz	191.250	196.125	
2.1.1.5	Slice Bandwidth	GHz	1.25GHz		
2.1.1.6	Slices/channel		1	3900	Min channel width: 1.25GHz Max channel width: 4.875THz
	Optical Power (50GHz channel)		Min	Max	
2.1.1.7	Input channel power (Pch)	dBm	-40	0	
2.1.1.8	Total input power	dBm		10	Over the operating frequency range
2.1.1.9	Absolute Channel Power Accuracy	dB	-0.5	+0.5	Optical power >= -30dBm
			-1.0	+1.0	Optical power <-30dBm
2.1.1.10	Relative Power Accuracy	dB		0.5	
2.1.1.11	Power Repeatability	dB		0.2	
2.1.1.12	Power Readout Resolution	dB		0.01	
	Misc		Min	Max	
2.1.1.13	Polarization Dependence	dB		0.5	
2.1.1.14	Optical Return Loss	dB	35		
2.1.1.15	Frequency Accuracy	GHz	-1.0	+1.0	
2.1.1.16	Relative Frequency Accuracy	GHz		0.5	

2.1.1.17	Frequency Readout Resolution	GHz		1.25	
	Scanning mode		Min	Max	
2.1.1.18	Sweep Time	ms		200	
	Static mode		Min	Max	
2.1.1.19	Bandwidth (BW)	GHz	1.25	4850	
2.1.1.20	Sampling Rate	Hz	5	2×10^6	n=1 slice
2.1.1.21	Sampling Interval	ns	n·500	200,000,000	Where "n" is the number of slices comprising the Bandwidth to be sampled
2.1.1.22	Number of samples			5000	

Table 6: Optical Specifications.

Power accuracy and frequency accuracy are tangled since if the frequency measurement is wrong then the power measured in a 12.5 GHz slot will also be wrong. The broader the slot and the more even the power distribution the less this matters. Almost as critical is sweep time since this is the heart-beat of the network control. All the power control "loops" within loops in the network start with the OCM reading. Since one OCM may be monitoring 40 ports, even a 200ms sweep time can really slow down network responses.

8.5. Questions

1. What is the minimum possible loss for a 1x4 splitter (assuming it splits evenly)?

2. A telecommunications link has a transmitter with an output power of +3 dBm. The receiver has a 1 in 10^9 bit error rate at a receive power of -23 dBm. The WSS linecard has a loss of 7 dB. The connector losses in the system are expected to be 2 dB. The system margin is 4 dB. The fiber loss is 0.25 dB/km. What is the maximum distance between the transmitter and the receiver without using an optical amplifier?

3. A 100 Gb/s transmitted signal is 25 GHz wide. If a device has 6ps/nm of dispersion, how much will the high frequencies in the signal be spread relative to the low frequencies?

(you can use 0.8 nm ≈ 100 GHz for the C-band). If there are 4 bits/symbol, what is the symbol separation?

4.

a. Consider the two polarisation states $\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}$ and $\frac{1}{\sqrt{2}}\begin{bmatrix}-1\\1\end{bmatrix}$. Use a dot product to show that they are orthogonal.

b. Now let's introduce a PDL of 3dB to the x component. What will the two states be now?

$$\frac{1}{\sqrt{2}}\begin{bmatrix} \\ 1\end{bmatrix} \text{ and } \frac{1}{\sqrt{2}}\begin{bmatrix} \\ 1\end{bmatrix}$$

c. Again calculate the dot product to see how much of the electric field from the first signal will be present when we try to measure the other signal.

d. Take the dot product of the second signal with itself to work out its amplitude and so calculate the relative amplitudes of the signal and the introduced "noise". This will give us the relative amplitudes of the electric fields.

Chapter 9. Fourier Transforms and what we can do with them…

9.1. Fourier Series

Data communications happens in time, but most aspects of our Wavelength Selective Switch (WSS) are best understood in terms of frequency so it's very useful to be able to move between these domains. The simplest translation, we've already done. A sin wave in time is a point function in frequency space. It's probably easiest to go next to a Fourier series. The idea here is that essentially any periodic function over an interval 2π can be represented as

$$f(x) = \frac{a_0}{2} + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + a_3 \cos 3x + b_3 \sin 3x + \dots$$

Where

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)dx$$

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \cos nx \, dx$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \sin nx \, dx$$

The data we transmit looks something like a square wave, so it's worth looking at the Fourier series for a square wave. We can write a square wave of length 2 as

$$f(x) = \cos \pi x - \frac{1}{3}\cos 3\pi x + \frac{1}{5}\cos 5\pi x - \frac{1}{7}\cos 7\pi x + \frac{1}{9}\cos 9\pi x - \cdots$$

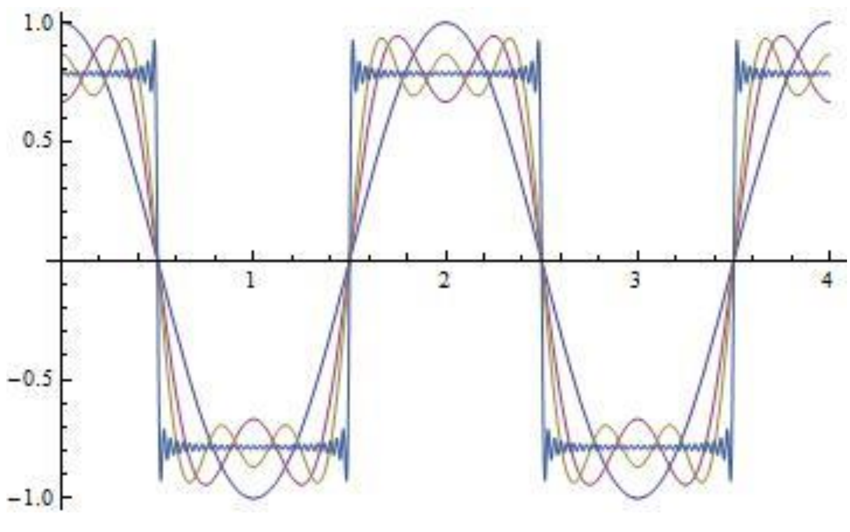


Figure 33 As we add terms the shape converges towards a square wave. This shows 1,2,3,4 and 30 terms. Note that when we don't have a very large number of terms we get a characteristic overshoot near the edge.

9.2. Spectral Broadening

Let's start with our square wave example. If we take

$$E(t) = e^{i\omega t}\cos\omega_m t = \tfrac{1}{2}(e^{i\omega t}(e^{i\omega_m t} + e^{-i\omega_m t})) = \tfrac{1}{2}(e^{i(\omega-\omega_m)t} + e^{i(\omega+\omega_m)t})$$

So in frequency space we have split the signal into two frequencies at $\omega \pm \omega_m$. If we add in the higher order terms then we get components at $\omega \pm 3\omega_m$, $\omega \pm 5\omega_m$ and so on.

When we modulate a monochromatic source we broaden the spectrum. If the signal is non-periodic then we can't use a Fourier Series but we can consider what the series looks like in the limit where the period goes to infinity. At this point, rather than a sum, we get an integral.

A Fourier series of period L can be written

$$f(x) = \frac{a_0}{2} + a_1\cos\frac{\pi x}{L} + b_1\sin x + a_2\cos 2\frac{\pi x}{L} + b_2\sin 2\frac{\pi x}{L} + a_3\cos 3\frac{\pi x}{L} + b_3\sin 3\frac{\pi x}{L} + \dots$$

Or, using complex numbers

$$f(x) = \sum_{n=-\infty}^{n=+\infty} c_n\, e^{in\frac{\pi x}{L}} \tag{9.1}$$

Where $c_o = \frac{a_0}{2}$, $c_n = \frac{1}{2}(a_n - b_n)$, $c_{-n} = \frac{1}{2}(a_n + b_n)$,

$$c_n = \frac{1}{2}\int_{-L}^{L} e^{-in\frac{\pi x}{L}} f(x)dx \tag{9.2}$$

In the limit as L goes to infinity (i.e. the function becomes non-periodic), the sum shown in (9.1) becomes an integral.

This allows us to convert between frequency space and time. The resulting functions we call the Fourier Transforms

$$f(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} g(\omega)e^{-i\omega t}d\omega$$

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

(we've taken equation 9.2 and let $g(\omega) = \sqrt{2\pi}\, c(\omega)$, let $L \to \infty$ and replaced x with $-t$)

So if we start with a wave

$$f(t) = e^{-i\omega t}$$

And then truncate it to a time interval τ_o then the frequency spectrum becomes

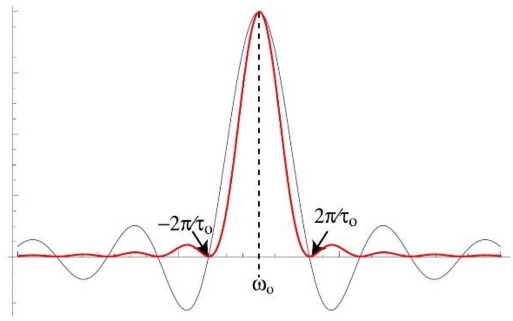
$$g(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\frac{\tau_o}{2}}^{\frac{\tau_o}{2}} e^{-i\omega_o t} e^{i\omega t} dt$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\frac{\tau_o}{2}}^{\frac{\tau_o}{2}} e^{i(\omega-\omega_o)t} dt$$

$$= \frac{2\sin[\frac{\tau_o(\omega-\omega_o)}{2}]}{\sqrt{2\pi}(\omega-\omega_o)}$$

To get the power spectrum we need to square $g(\omega)$. This function has its first zeros at

$\omega - \omega_o = \pm {}^{2\pi}/_{\tau_o}$



If we take a 1550nm source and turn it on and off to send a 40Gb/s signal, how much do we broaden the frequency spectrum?

How could we send the same amount of data, but create less broadening?

9.3. Fourier Transform Spectroscopy

An understanding of Fourier Transforms is essential to understand several of the new products we are currently developing. However, a good place to start understanding how Fourier transforms are so useful is in Fourier Transform Spectroscopy.

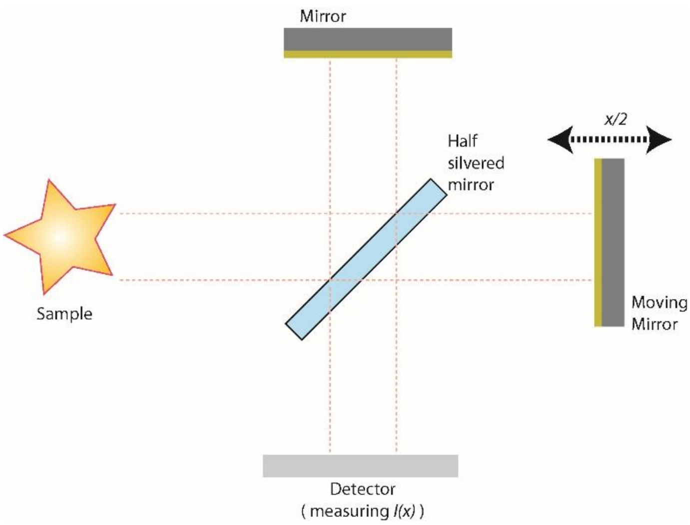


Figure 34 Optical Setup for FT Spectroscopy

If the sample produces a single wavelength (and hence wavenumber), then the moving mirror will cause a sinusoidal intensity variation at the detector with amplitude $(1 + \cos kx)$.

If the source is a mix of wave numbers $G(k)$then (as described in your textbook on page 81)

$$I(x) = \int_0^\infty (1 + \cos kx) G(k) dk$$
$$= \frac{1}{2} I(0) + \frac{1}{2} \int_{-\infty}^{\infty} G(k) e^{ikx}\, dk$$

So using the inverse Fourier transform

$$G(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \big(2\, I(x) - I(0)\big) e^{-ikx}\, dx$$

In words, if we map out the fringes we can read back the spectrum. This is often the best way of measuring a spectrum with very low light levels.

9.4. Optical Coherence Tomography

An understanding of Fourier Transforms is also necessary to understand wavelength scanning Optical Coherence Tomography.

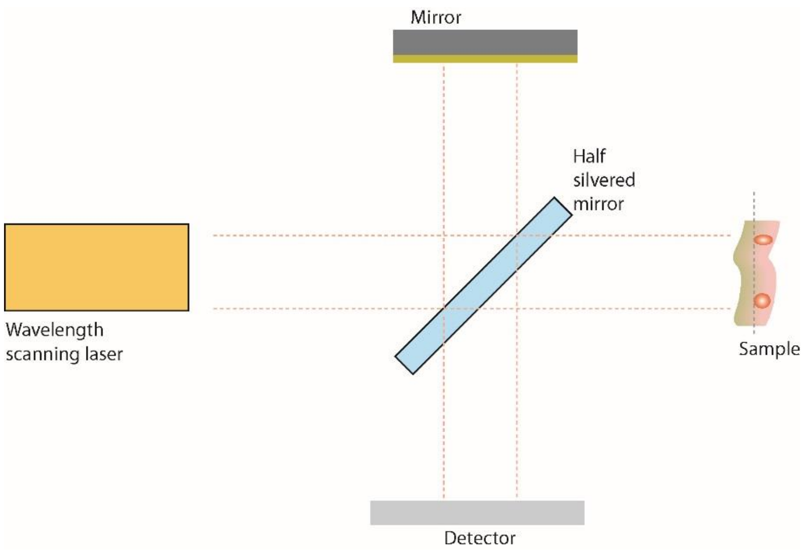


Figure 35 Wavelength Scanning OCT

In a wavelength scanning OCT system a reflection I_r from the sample of depth x/2 (compared to the reference depth "0" corresponding to the other optical path) will produce a sinusoidal response as the wavelength scans.

$$I(k) = I_r \cos^2(k(x_o - x)/2)\,)$$

$$= \frac{I_r}{2}(1 + \cos(k(x_o - x))$$

So for varying reflections at different depths

$$I(k) = \int_0^\infty (1 + \cos kx) I_r(x) dx$$

$$= \frac{1}{2} I_o(0) + \frac{1}{2}\int_{-\infty}^{\infty} I_r(x) e^{ikx}\, dk$$

We can again do the inverse transform and read off the reflections at various depths.

$$I_r(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \big(2\, I(k) - I(0)\big) e^{-ikx}\, dx$$

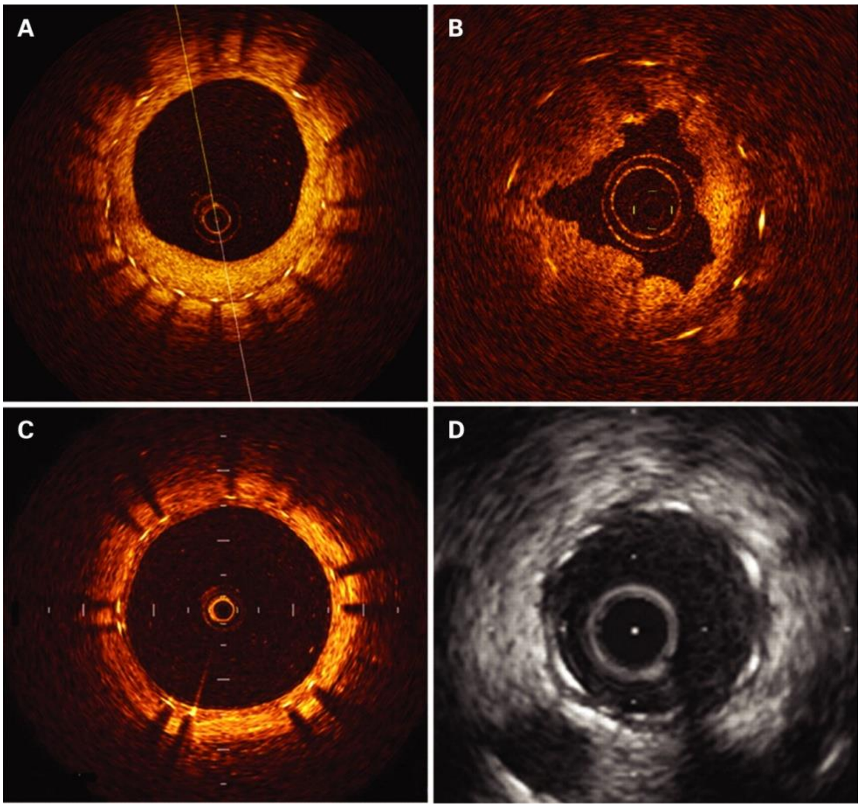


Figure 36 OCT images of a stents placed in one of the major coronary arteries (A,B and C are OCT images, D is an ultrasound image (from *Heart* 2008;issue **94**:pg 1200-1210, Cardiac optical coherence tomography O C Raffel et al.)

9.5. Questions

7. Show that the Fourier transform of e^{-t^2} is $e^{-\omega^2}$ (i.e. that a Gaussian shaped pulse in time corresponds to a Gaussian shaped pulse in frequency)

8. Find the Fourier transform of the sharp peaked function

   $f(t) = e^{-a|t|}, a > 0$

Chapter 10. Etalons

In week 7 we talked about the result of two sources interfering. However, many of the devices we actually use the interference of many beams to achieve high wavelength and/or spatial resolution. Etalons are used as a wavelength reference in virtually every laser Finisar sells, as well as the HR-OCM and the wavelength reference system in the low profile dual WSS. Etalon-like effects also happen accidentally when two partly reflective, parallel surfaces are present in a system (a delaminated glue bond is a typical example).

10.1. Etalons – the Mathematics

First consider two parallel plates as sketched below.

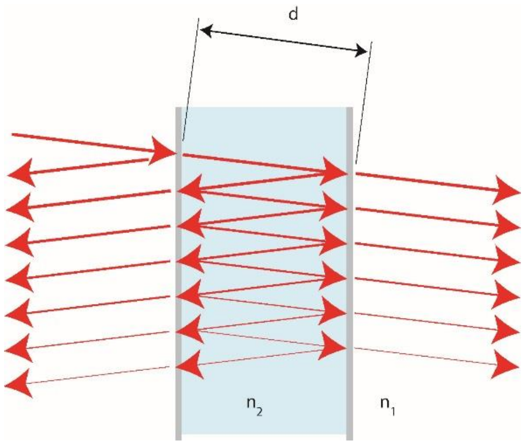


Figure 37 A simple etalon

Each time the light reflects, it is reduced in amplitude and each reflection is phase shifted by Δ, say, where

$$\Delta = 2kd + \delta_r = 2\frac{2\pi}{\lambda}d + \delta_r \qquad (1)$$

The quantity δ_r is the phase shift on reflection. You may remember from the end of week 5 that the electric field strength of light reflected from a dielectric surface (when the beam hits at normal incidence) is

$$r = \frac{n_1 - n_2}{n_1 + n_2}$$

Where the light reflects from a material with refractive index n_2 , coming from a material with refractive index n_1. Notice that when $n_1 \geq n_2$, then the sign of r is positive and there is no phase shift. When $n_1 \leq n_2$ there is a sign change (i.e. a 180° phase shift).

The transmitted electric field amplitude of the wave will be

Equation 5-1

$$E_t = E_o\left(t^2 + t^2r^2e^{i\Delta} + t^2r^4e^{i2\Delta} + t^2r^6e^{i4\Delta} + \cdots\right)$$

$$= E_o t^2 \left(1 + (r^2 e^{i\Delta}) + (r^2 e^{i\Delta})^2 + (r^2 e^{i\Delta})^3 + \cdots\right)$$
$$= \frac{E_o t^2}{1 - r^2 e^{i\Delta}}$$

When the total phase shift is a multiple of 2π then all the field strength (and light intensity) is transmitted, since $t^2 = 1 - r^2$. This is the resonance condition. The larger "r" the sharper the function.

The total transmitted power is

$$I_t = |E_t|^2$$

It takes a few lines of maths to rearrange this to the Airy Function

$$I_t = \left|\frac{E_o t^2}{1 - r^2 e^{i\Delta}}\right|^2 = \left|\frac{E_o T}{1 - R e^{i\Delta}}\right|^2$$

and

$$\left|1 - Re^{i\Delta}\right|^2 = \left(1 - Re^{i\Delta}\right)\left(1 - Re^{-i\Delta}\right) = 1 + R^2 - 2R\cos\Delta = 1 + R^2 - 2R(1 - 2\sin^2 {}^{\Delta}\!/_{2})$$

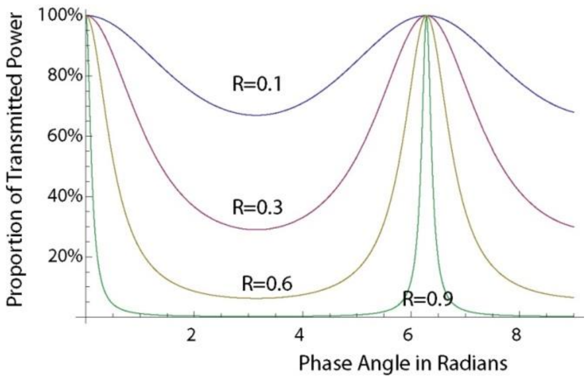
$$= 1 - 2R + R^2 + 4R\ \sin^2 {}^{\Delta}\!/_{2} = (1-R)^2 + 4R\ \sin^2 {}^{\Delta}\!/_{2} = (1-R)^2(1 + \frac{4R}{(1-R)^2}\sin^2 {}^{\Delta}\!/_{2}$$

so

$$I_t = |E_t|^2 = \frac{I_o}{1 + F\sin^2\frac{\Delta}{2}}$$

Where F is the "finesse" of the cavity

$$F = \frac{4R}{(1-R)^2}$$



Wavelength Spacing of a Resonator

Etalons are normally used to select wavelengths of certain spacing. The wavelength change needed to move from one resonance and the next is simply the change needed to fit in an extra wavelength.

It's quite easy to show (and we'll do so) that by differentiating equation 1

$$\Delta\lambda = \frac{\lambda^2}{2mL}$$

Where L is the length of the cavity.

In terms of frequency

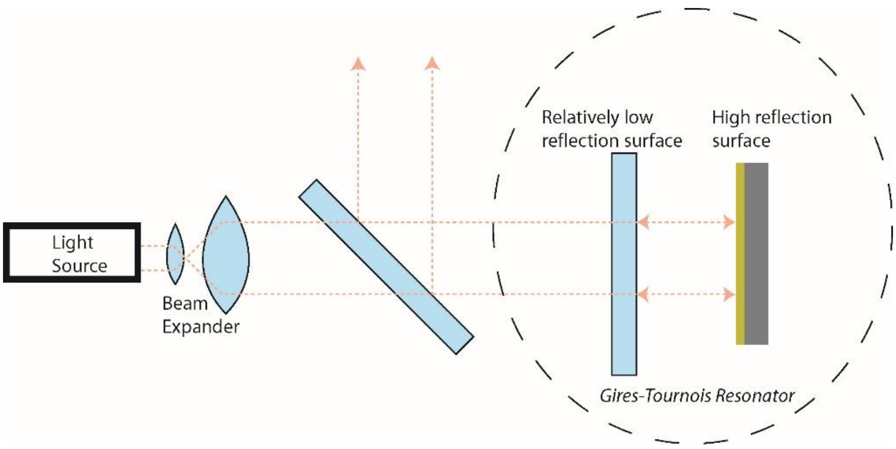
$$\Delta f = \frac{c}{2mL}$$

This makes sense if you think on the etalon as a resonating cavity with a fundamental period of $2L/c$ (the round trip time) and so a fundamental frequency of $c/2L$ in vacuum or $c/2nL$ in a material with refractive index n.

What frequency spacing would a 3mm air etalon indicate?

10.2. Gires–Tournois etalon

The Gires–Tournois etalon has only recently been popular and so doesn't show up in your textbook. It's an immensely useful resonator (and used in some rather creative ways by Steve in various designs). It doesn't produce large amplitude changes, instead, it produces a rapid phase change at resonance.



In the same way we generated equation 5.1, we can get the reflected wave amplitude from a Gires-Tournois Etalon (assuming the second surface reflectivity is 1)

$$E_r = E_o \frac{r_1 - e^{-i\Delta}}{1 - r_1 e^{-i\Delta}}$$

At resonance, when Δ is a multiple of 2π, the phase flips.

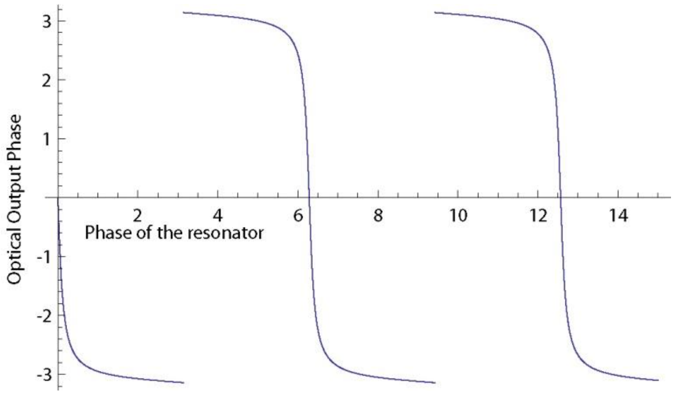


Figure 38 The phase of light front a GT Etalon where the front mirror reflectivity is set to 0.9

10.3. How Etalons Go Wrong

An etalon is a remarkably simple piece of optics (just a rectangular block of glass) and yet they can give unexpected results. Here are several ways they can go wrong…

1. Surface flatness variations.
    a. Small variations in the surface flatness mean that you're actually looking at many similar etalons at once. This has several effects. Obviously, the etalon curve become smeared, as sketched below.
    b. In addition, if the beam moves around on the surface due to temperature or stress changes then the positions of the peaks will also move if the surface isn't flat.
    c.

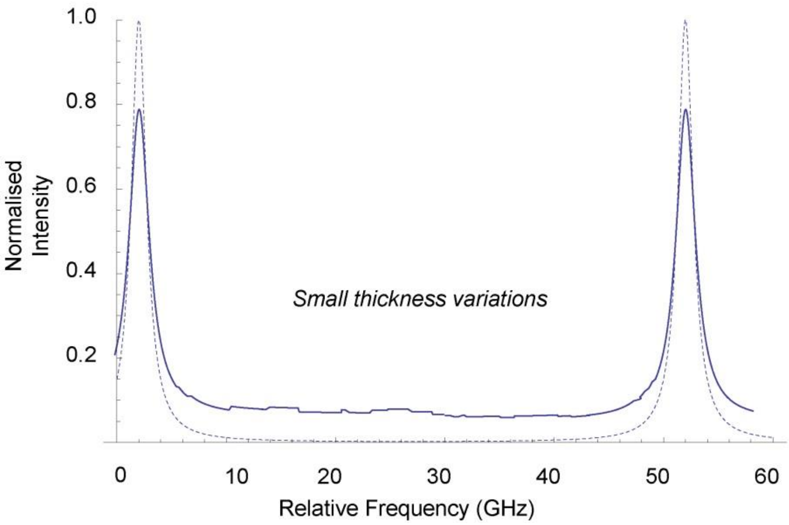


Figure 39 Small variations in etalon thickness cause a characteristic rise in the transmitted intensity between the peaks

2. Temperature errors:

Thermistors are usually used to measure the etalon temperature, but they are not perfect. If they are not very close to the etalon, and there are temperature gradients then they can measure a different temperature. If the thermistor is attached with a silver-filled epoxy then the resistance of the epoxy can vary with temperature. This is particularly common in the last part of an epoxy tube that has been stored vertically. The majority of the silver sinks to the bottom and is dispensed first. The last part of the epoxy has very few conductive paths that can open and close with temperature changes, causing large changes in the resistance.

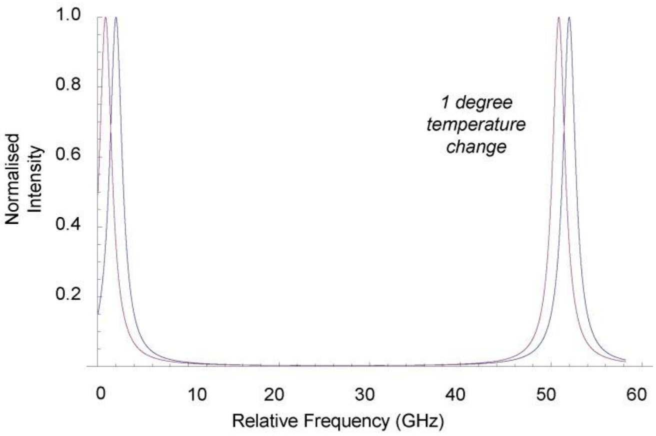


Figure 40 A 50GHz fused silica etalon has approximately 1 GHz/C° temperature drift

10.4. Questions

1. Consider at etalon of refractive index 1.5. What length will give a resonance of 50GHz?

2. How much will the resonance frequency spacing change if the beam enters at 10 degrees?

3. How many fringes will etalon pass through as the beam rotates 10 degrees?

4. Consider an etalon made from a block of silicon with no coatings (refractive index of 3.48), length 3mm. What will be the depth of the fringes (as a percentage) and what will be the separation of the fringes in GHz? (You'll need the equation for reflections for a perpendicular beam (which I hope you know):

$$R = \left[\frac{n-1}{n+1}\right]^2$$

Chapter 11. Etalon Resolution and Multilayer Coatings ...

In week 10 we talked about etalons. This week we'll describe the resolving power of an etalon and then extend the concept of etalons to anti-reflection, high reflection and band-pass coatings.

The Resolving Power of an Etalon

Last week we described the transmitted intensity of an etalon:

$$I_t = |E_t|^2 = \frac{I_o}{1 + F\sin^2\frac{\Delta}{2}}$$

where F is the "finesse" of the cavity

$$F = \frac{4R}{(1-R)^2}$$

R is the reflectivity of the mirrors and

$$\Delta = 2kd + \delta_r$$

This results in a spectrally sharp series of transmissions where the etalon resonates.

In many wavelength selective devices, we need to define a resolving power for the device. The exact point at which we say two frequencies are resolved is somewhat arbitrary, but here we will use the separation at which the sum of the two functions has a dip equal to the amplitude of either one of the functions (called the Taylor criterion)

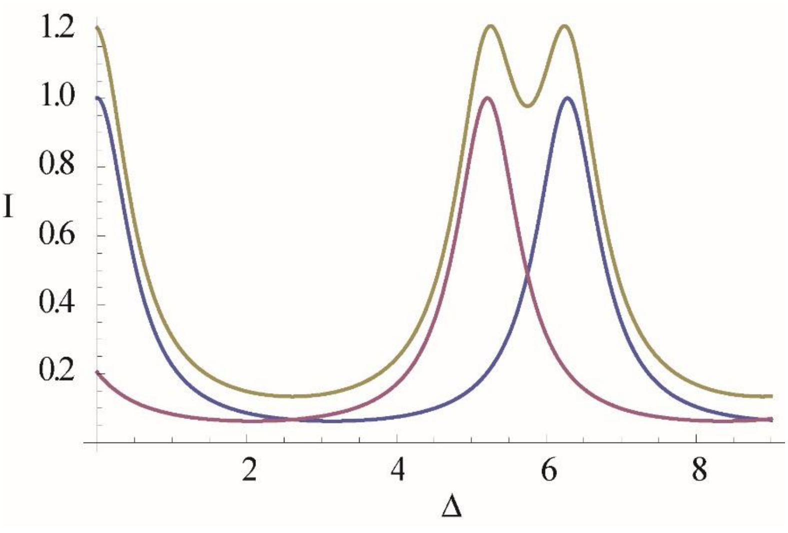


Figure 41 Two Airy functions are said to be resolvable when the sum of the two functions creates a saddle with amplitude equal to the amplitude of one of the contributing functions.

The combined amplitude is

$$I_{total} = I_t(\Delta) + I_t(\Delta') = |E_t|^2 = \frac{I_o}{1 + F\sin^2\frac{\Delta}{2}} + \frac{I_o}{1 + F\sin^2\frac{\Delta'}{2}}$$

At the saddle $\frac{\Delta'}{2} = 2\pi n + \frac{\Delta-\Delta'}{4}$ (since we are halfway to $\frac{\Delta}{2}$) and $I_{total} = I_o$

So

$$I_{total} = I_o = \frac{2I_o}{1 + F\sin^2\frac{\Delta-\Delta'}{4}}$$

And

$$F\sin^2\left(\frac{\Delta-\Delta'}{4}\right) = 1$$

If $\frac{\Delta-\Delta'}{4}$ is small, then we can drop the "sin" and we get

$$|\Delta - \Delta'| = \frac{4}{\sqrt{F}} = 2\left(\frac{1-R}{\sqrt{R}}\right)$$

Or using the third equation from this week

$\Delta = 2kd + \delta_r$, so

$$\Delta - \Delta' = 2(k-k')d = 2\left(\frac{2\pi f}{c} - \frac{2\pi f'}{c}\right)d = \frac{4\pi d}{c}\Delta f$$

And so combining these last two equations

$$\Delta f = \frac{c}{d\pi\sqrt{F}} = \frac{c}{2d\pi}\left(\frac{1-R}{\sqrt{R}}\right)$$

So if we have an etalon spacing d with reflection R from each of the surfaces, we can now calculate its frequency resolving power, which is exactly what we need to do in an etalon based optical channel monitor.

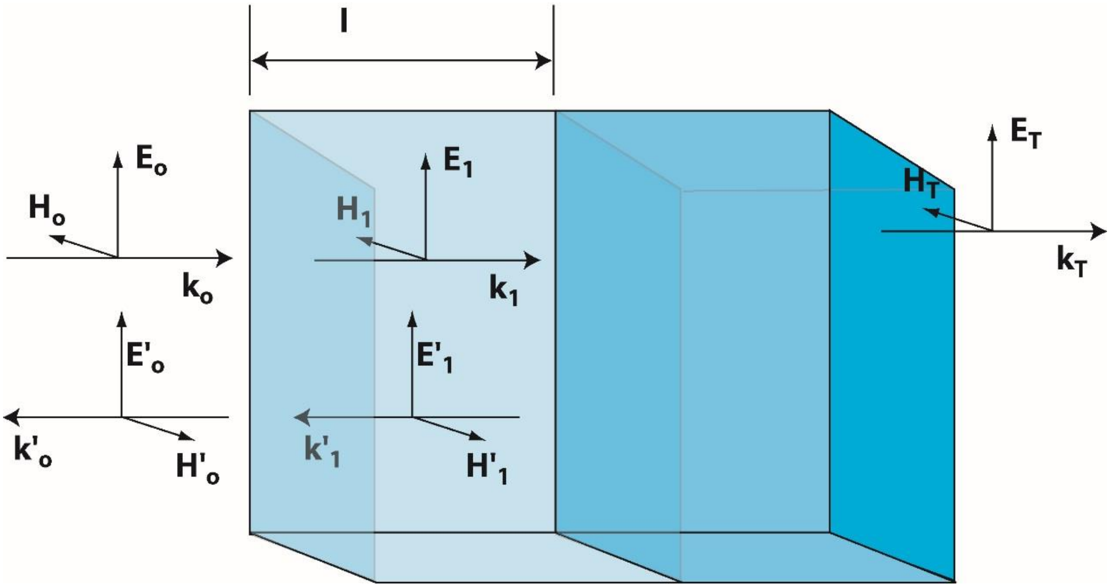
Multi-Layer Coating

There is rarely a surface in a modern optical design that doesn't have some type of dielectric coating over it. To understand multi-layer coatings, we'll use the same maths as we did for etalons, but first the cheats guide:

- Simplest simple layer anti-reflection coating – ¼ wavelength thick magnesium fluoride (n=1.34)
- High reflection coating – many alternate high and low reflection layers ¼ wavelength thick.
- Fabry-Perot interference filter – 2 high reflection coatings separated by a half wavelength thick layer.

Now for the messy mathematics…

Consider two perpendicular surfaces and an incident wave as sketched below



At the first face the electric field must be continuous across the surface (divergence of **E** is zero), so

$$E_o + E'_o = E_1 + E'_1$$

Similarly (but with a sign change so we can keep the handedness correct)

$$H_o - H'_o = H_1 - H'_1$$

Or

$$n_o E_o - n_o E'_o = n_1 E_1 - n_1 E'_1$$

At the second interface we need to worry about the relative phase of the beams

$$E_1 e^{ikl} + E'_1 e^{-ikl} = E_T$$

And

$$n_1 E_1 e^{ikl} - n_1 E'_1 e^{-ikl} = n_T E_T$$

Half a page of algebra (described on pages 97 and 98 of your textbook) lets us eliminate the intermediate fields and finish with a pair of equations that can be written

$$\begin{bmatrix} 1 \\ n_o \end{bmatrix} + \begin{bmatrix} 1 \\ -n_o \end{bmatrix} r = M \begin{bmatrix} 1 \\ n_T \end{bmatrix} t$$

Where $r = \frac{E_o{}'}{E_o}$ (the proportion of reflected light) and $t = \frac{E_T}{E_o}$, the proportion of transmitted light, and

$$\boldsymbol{M} = \begin{bmatrix} \cos kl & -\dfrac{i}{n_1} \sin kl \\ -i n_1 \sin kl & \cos kl \end{bmatrix}$$

$\boldsymbol{M}$ is called the transfer matrix, and it can be shown that if we have multiple layers then this equation stays in the same form, we just multiply the transfer matrices together.

11.1.1. Antireflection Coating

If the film is a quarter wavelength thick then $kl = \pi/2$. Calculate *r.* What happens when $n_1 = \sqrt{n_T}$ and $n_o = 1$?

In reality finding a coating that has $n_1 = \sqrt{1.5}$ =1.22 say, combined with all the other properties we require of a coat (sticks to glass, hard, transparent, doesn't dissolve in water, won't be attacked by fungus or bacteria etc.) is really difficult. Also it will only work for a narrow range of wavelengths. Typically, for single layer coatings, magnesium fluoride is used, which has an index of 1.33. This reduces the reflection to 1%. Figure 4.8 of your textbook shows how three coatings can be combined to give less than 0.4.

11.1.2. High Reflectance Films

A high reflectance film is made from alternating layers of high and low refractive index material. Another commonly used material for low refractive index coating is cerium fluoride (n=1.63). The common high refractive index materials are zirconium dioxide, n = 2.1; zinc sulfide, n = 2.32; titanium dioxide, n= 2.4. You may remember in week 5 we derived the reflection equation (page 14 of your notes)

$$r_s = \frac{\cos\theta - n\cos\phi}{\cos\theta + n\cos\phi}$$

Notice that for $\theta = \phi = 90°$ that r_s is positive if $n < 1$ and negative if $n > 1$ (i.e. there is a 180° or π radian phase change). To get constructive interference of the reflected waves we simply need another or π radian phase change, so the thickness of each layer needs to be π/4, so

$$(\boldsymbol{M_1}.\boldsymbol{M_2})^N = \left(\begin{bmatrix} 0 & -\dfrac{i}{n_L} \\ -in_L & 0 \end{bmatrix}\begin{bmatrix} 0 & -\dfrac{i}{n_H} \\ -in_H & 0 \end{bmatrix}\right)^N = \begin{bmatrix} \left(-\dfrac{n_H}{n_L}\right)^N & 0 \\ 0 & \left(-\dfrac{n_L}{n_H}\right)^N \end{bmatrix}$$

For large N this approaches $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Satisfy yourself that this results in a reflective surface. Another typical high refractive index materials is Ta_2O_5 (n=2.08, with good stability over temperature). Typically 21, 23 or 25 layers will give better than 99.9% reflectivity (An odd number allows you to finish with a high refractive index layer).

As a rule-of-thumb, most coating will work reasonably well up to an incidence angle of about 30°. At higher angles, the angle must be taken into account.

11.1.3. A good place to start looking for design software:

https://lightmachinery.com/optical-design-center/thin-film-cloud/

try an etolon “(1H 1L)5 1H 100M (1H 1L)5 1H”

11.2. Questions

1. Use the coating design software at https://lightmachinery.com/optical-design-center/thin-film-cloud/ to design a high reflection coating (if you need a more advanced, and expensive, program then "RP Coating" is one option, Open Coating. To do this first set the design target wavelengths (choose a range of about 100nm) and the property you are trying to optimise (reflection). Then set up a basic mirror coating profile (e.g. "(1H 1L)4 1H"). Make sure you put in the spaces correctly or it will get very confused. Choose your materials - TiO_2/MgF_2 gives a hard outer coating and good transmission in the visible, but there are many options. Don't forget to set up your plotting range and your layer limit. Now optimize and see how good a mirror you can get with 12 layers.

2. Now select "Etalon simulations", choose the substrate material to be fused silica (silicon dioxide). Place 500M in the middle of two copies of your mirrors (500 quarter wavelengths of a medium refractive index material) to build an etalon.

3. Now move the beam to 20 degrees. How does this affect your etalon? Look at each of the polarisations.

4. a. A good metal coating can achieve a reflectivity of 0.9. If the etalon is spaced at 1mm in air, what is its resolving power? (you'll need the equation $\Delta f = \frac{c}{d\pi\sqrt{F}} = \frac{c}{2d\pi}\left(\frac{1-R}{\sqrt{R}}\right)$)

   b. A dielectric coating can achieve a reflectivity of 0.99. What is the resolving power of a 1mm etalon with such a coating?

5. Gallium Arsenide (GaAs) solar cells have a refractive index of 3.5.
   a. How much of the sun's light is reflected from the surface - use the equation $R = \left[\frac{n-1}{n+1}\right]^2$?

b. How would you design a single layer anti-reflection coating for GaAs (choose a material and calculate the thickness)? Assume your wavelength of maximum interest is green (500nm). You can assume you can use any of the materials we've described (zirconium dioxide, n = 2.1; zinc sulfide, n = 2.32; titanium dioxide, n= 2.4, cerium fluoride n=1.63) although in reality titanium oxide adheres best

c. How much have you improved the performance of the cell (at 500nm)? You can use the equation for a single layer quarter wave thick layer we derived last week

$$R = |r^2| = \frac{(n_T - n_1)^2}{(n_T + n_1)^2}$$

Chapter 12. Week 12 Diffraction

Huygen's principle tells us that every point on a wave acts as a source of new (spherical) waves. We are going to skip over the 5 pages of 3D calculus at the start of chapter 5 in your textbook (Introduction to Modern Optics). There are a few important ideas here (but it's not worth spending 2 weeks to go through it). The key ideas are

1. When you are within a few wavelengths of an aperture there are other terms you need to consider in calculating the power distribution (this is called the "near field")
2. All the mathematics we are doing here is treating the electric field as a scalar. A full vector treatment of any of these diffraction patterns is a vast amount of work.
3. An aperture and its inverse give the same diffraction pattern ($U_p = U_{p1} + U_{p2}$, so where $U_p = 0$ then $U_{p1} = -U_{p2} = e^{-i\pi}U_{p2}$)
4. A full (scalar) analysis allows us to derive a refinement to Huygen's principle: the amplitude of the resulting wave has a directional dependence of $1 + \cos(n, r)$, where θ is the angle between the new wave and the normal to the aperture (assuming that the wave was initially coming straight at the aperture, otherwise the dependence is $\cos(n, r') + \cos(n, r)$, where r' is the direction back to the wave source, from the aperture. This is called the obliquity factor.

12.1. Fraunhofer Diffraction Patterns

In calculating Fraunhofer diffraction patterns we assume that we have a plane wave coming into the aperture and we are interested in plane waves coming from the aperture at a distance from the aperture much greater than the wavelength. This is generally the case (and greatly simplifies the maths). We're also going to assume the angles are small enough that we can ignore the obliquity factor, which can be easily added in if you need to.

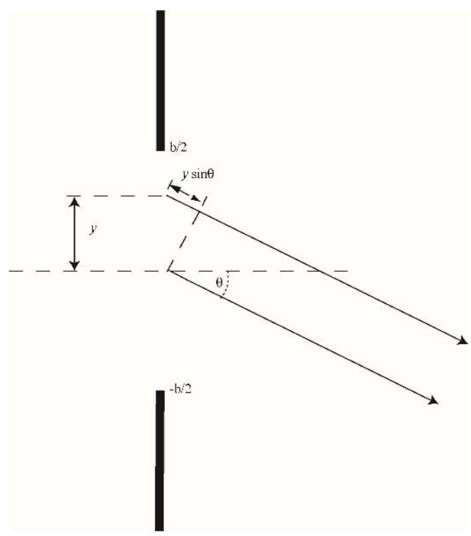
To work out the resulting electric field at position "***r***" relative to the aperture we simply add up the waves coming from every point on the aperture, i.e. the amplitude of the field at ***r*** is

$$U_p = C \iint e^{ikr} dA$$

where we integrate over the area of the aperture.

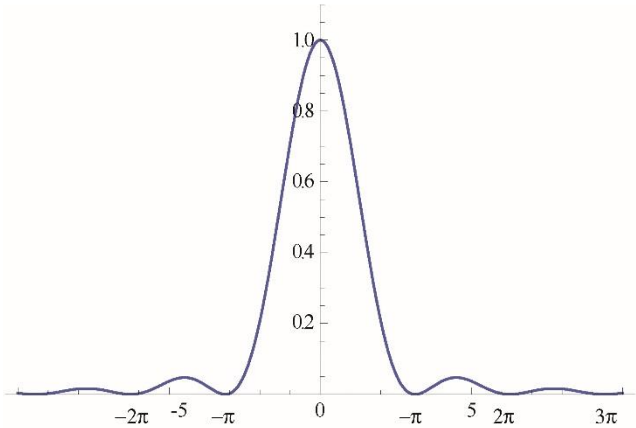
12.2. Single-slit Diffraction Patterns

For a single slit



$U_p = Ce^{ikr_o}\int e^{iky\sin\theta}Ldy$

$$= CLe^{ikr_o}\frac{\sin\beta}{\beta}$$



Where $\beta = \frac{1}{2}kb\sin\theta$ and L is the length of the slit (assumed to be long).

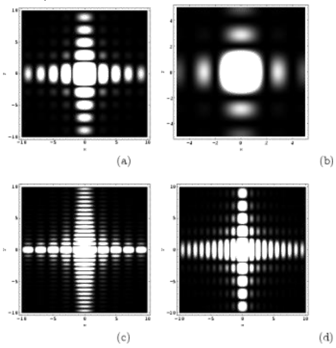
The intensity is the square of this function graphed beside.

Note that the central maxima is twice the width of the other maxima. The height of the second peak is 5% of the height of the centre.

A rectangular aperture can be derived in exactly the same way giving

$$I = I_o\left(\frac{\sin\alpha}{\alpha}\right)^2\left(\frac{\sin\beta}{\beta}\right)^2$$

For a circular aperture we take a slice *dy* across the circle which will be $2\sqrt{R^2-y^2}$ wide.



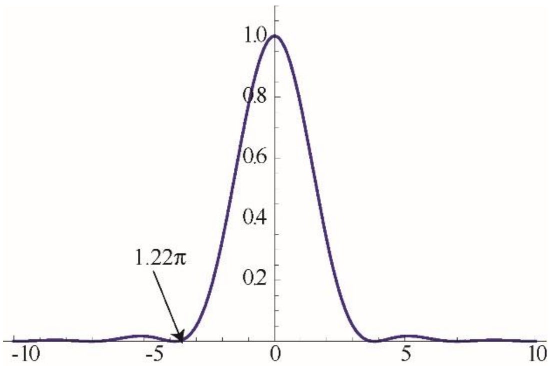
The integral is then

$$U = \int_{-R}^{R} e^{iky\sin\theta}\,2\sqrt{R^2-y^2}dy$$

$$= \int_{-1}^{1} e^{iu\rho}\,2\sqrt{1-u^2}du$$

If we let *u=y/R* and $\rho = kR\sin\theta$ then this is a standard Bessel function integral. Squaring it to give the intensity

$$I = I_o\left(\frac{2J_1(\rho)}{\rho}\right)^2$$



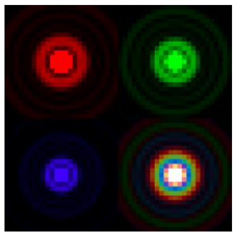


Figure 42 Airy Disk Picture from http://www.flickr.com/photos/jmzawodny/2179210322/sizes/m/ Airy disk

The Airy disk includes 84% of the total light. The minimum has diameter D where

$$\sin\theta = \frac{1.22\pi}{kR} = \frac{1.22\lambda}{2R} = \frac{1.22\lambda}{D}$$

In other words, once a wave has gone through an aperture it has an angular size proportional to the wavelength and inversely proportional to the aperture diameter.

12.3. Rayleigh Criteria

To resolve two distinct objects (located at infinity), we need to be able to resolve their Airy disks. The Rayleigh criteria states that two points can just be resolved when the minimum of one corresponds to the maximum of the other. I.e.

$$\Delta\theta = \frac{1.22\pi}{kR}$$

This gives us a 19% dip between the maxima.

12.4. Young's Double Slit

Consider two slits, separation h, widths b.

$$\begin{aligned}
&\int^{Area} e^{iky\sin\theta}dy = \int_0^b e^{iky\sin\theta}dy + \int_h^{h+b} e^{iky\sin\theta}dy \\
&= \frac{1}{k\sin\theta}\left(e^{ikb\sin\theta} - 1 + e^{ik(h+b)\sin\theta} - e^{ikh\sin\theta}\right) \\
&= \left(\frac{e^{ikb\sin\theta} - 1}{ik\sin\theta}\right)\left(1 + e^{ikh\sin\theta}\right) \\
&= 2be^{i\beta}e^{i\gamma}\frac{\sin\beta}{\beta}\cos\gamma
\end{aligned}$$

Where $\beta = \frac{1}{2}kb\sin\theta$ and $\gamma = \frac{1}{2}kh\sin\theta$. Notice that this is a sinusoid superimposed on a single slit pattern.

The intensity of a double slit pattern is given by the square of the electric field strength

$$I = I_o\left(\frac{\sin\beta}{\beta}\right)^2 .\cos\gamma^2$$

12.5. Diffraction Grating

For a diffraction grating

$$\int^{Area} e^{iky\sin\theta}dy = \int_0^b e^{iky\sin\theta}dy + \int_h^{h+b} e^{iky\sin\theta}dy + \int_{2h}^{2h+b} e^{iky\sin\theta}dy + \cdots\int_{Nh}^{Nh+b} e^{iky\sin\theta}dy$$

$$= 2be^{i\beta}e^{i\gamma}\frac{\sin\beta}{\beta}\frac{\sin N\gamma}{\gamma}$$

And the corresponding intensity is

$$I = I_o\left(\frac{\sin\beta}{\beta}\right)^2\left(\frac{\sin N\gamma}{N\sin\gamma}\right)^2$$

12.5.1. Apodisation

We can remove the Airy disc from an image by using a soft edge. The Fourier transform of a Gaussian function is another Gaussian, so if we use an aperture with soft edges we can remove the Airy rings at the cost of making the spot larger.

12.5.2. Focusing to a point…

We can turn around the equation

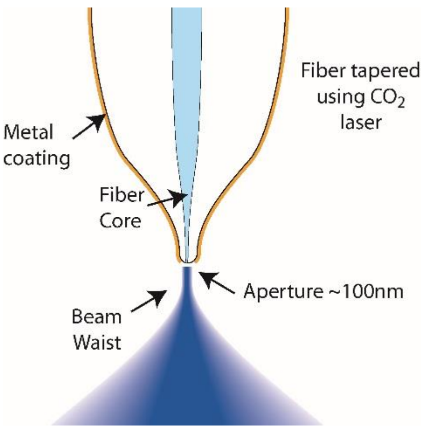
$$\sin\theta = \frac{1.22\lambda}{D}$$

To

$$D = \frac{1.22\lambda}{\sin\theta}$$

This gives us the limit to which a point can be focussed. This is very important in microscopy. If we want to look at (or focus light to) a very small spot we can either increase the angle (use a lens with a very short focal length) or use a shorter wavelength (which is why DVD players use a blue laser) and high end chip makers use UV lasers. This is also why electron microscopes are so good (wavelengths of 2.5 pm at 200 keV – although the lens are so bad they can't get much beyond 0.1 nm).

Cheating the diffraction limit



In microscopy there are a few ways of getting to even higher resolutions. One way is to operate in the "near field". By moving a "pin-hole" over the surface, made with a tapered optical fiber, we can sample the objects reflection at the aperture, which can be significantly smaller than the wavelength of the light. Of course this gives us nearly zero "depth of field" and often results in a broken fiber tip!

Another way is to use a "STED" (Stimulated emission depletion) system. We create a donut shaped beam with a phase that increases from 0 to 360° As we go around the beam. At the centre we get near perfect destructive interference. We use a normal shaped beam at one wavelength to excite certain molecules and then use the donut shaped beam to de-excite them. The only excited molecules left are at the middle of the donut. Unfortunately, the intensities get very high for the donut – it's easy to cook the sample.

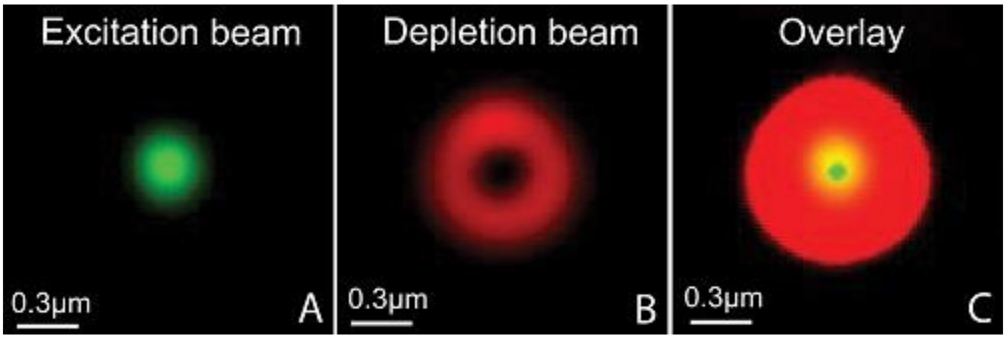


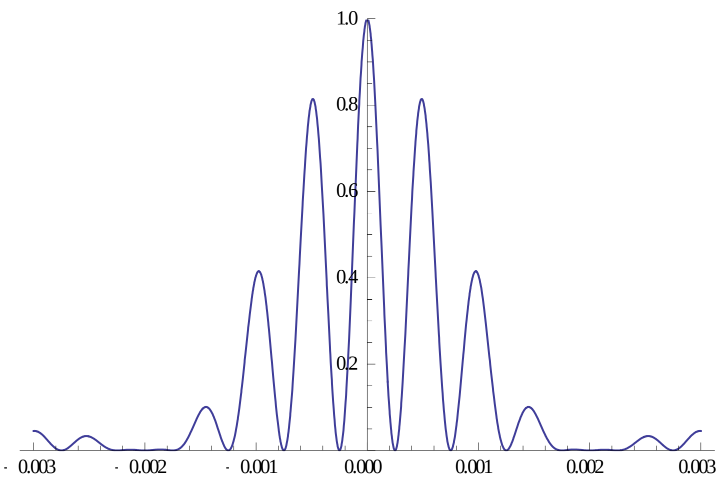
Figure 43 Donut shaped depletion beams (from http://www.anes.ucla.edu/sted/principle.html)

12.6. Questions

1. Measure the single slit diffraction pattern and estimate the width of the slit (assume a 1m distance to the screen and a wavelength of 650nm for the red laser).

2. To read a numberplate on a car from low earth orbit you need to resolve approximately 2cm. What size telescope (diameter of the telescope mirror) would be needed to resolve a numberplate from low earth orbit (assume 500nm light, 100km orbit)?

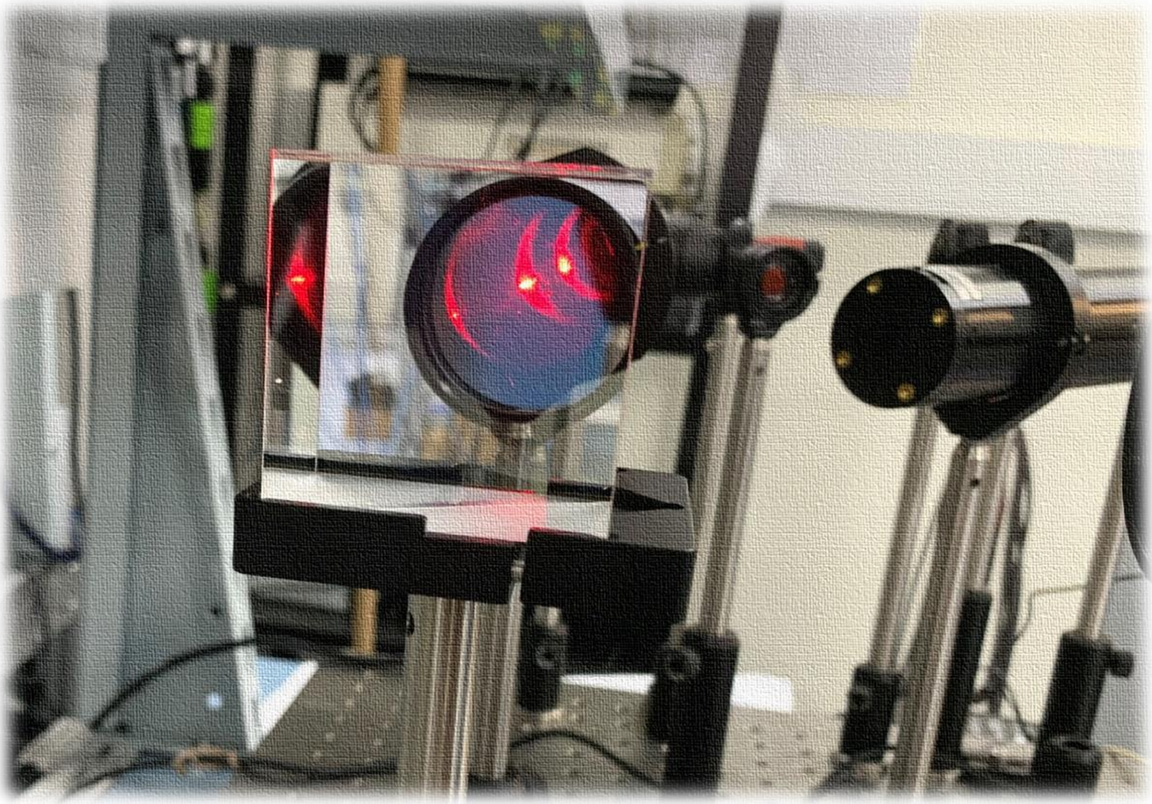
3. The fourth fringe is missing from a two slit interference pattern. What's the ratio between the slit width and the slit separation?



Optics for Telecommunications

Part 2: Theory and Engineering

Ian G. Clarke



This is a human-readable summary of (and not a substitute for) the license. Disclaimer.

Contact Details:

Email: iclarke@ieee.org

Mobile: 0451 849 125

Contents

Chapter 13. More on Diffraction Gratings

Last week we briefly looked at the simplest diffraction grating; a series of slits in the presence of a plane wave. Let's start with what a real grating looks like.

There are two broad categories of bulk optic grating; reflection grating and transmission gratings. Let's consider a typical *reflection grating* from the (Finisar) company Lightsmyth. It could have 1700 lines/mm written into silicon, be gold coated, and have an efficiency of about 95% (i.e. 95% of the light goes where you expect). These gratings can be made polarisation insensitive (about 2% polarisation dependence). Most interestingly, because the lines are created by etching the surface of the silicon, the lines don't need to be straight. In many of our DWP products we still use a polymer on glass grating. This is make by pressing the polymer against a master grating, curing it and coating it in gold. The master grating is created either by ruling the lines (scratching them into the glass) or holographically by interfering two plane waves (essentially the same as we did using the two-slit experiment. For 1550nm gratings we typically use around 1600 lines per mm. Ruled gratings tend to be more efficient, while holographic gratings tend to not produce ghost images from errors in the lines. These are called "replicated holographic" gratings. Blazing a grating preferentially directs the light into one of the "orders" – usually the 1st order.

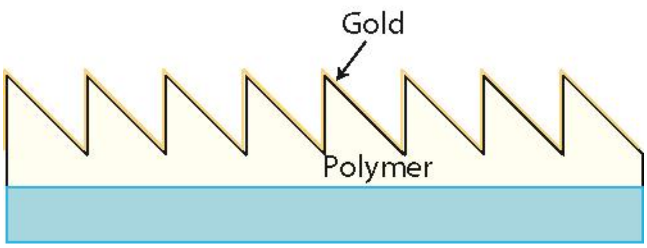


Figure 2 A blazed replicated holographic grating

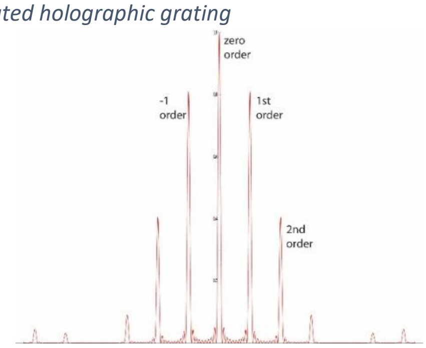


Figure 1 Orders of a diffraction grating

Another approach is to write the grating in the same way as a silicon chip is made. In this case a lithographic mask is made using an e-beam writer. The glass is coated in photoresist and is then exposed to UV light going through the mask. The surface is then etched, with the unexposed (positive) photoresist protecting the exposed surfaces. The results can be seen below:

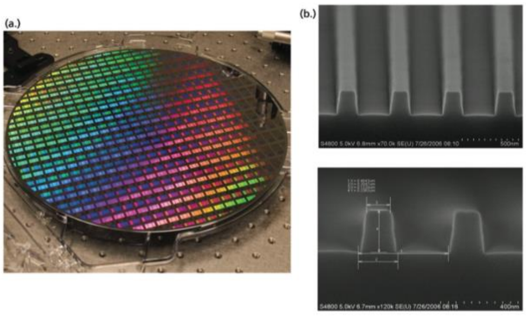


Figure 3 Gratings producted by hotolithography (from https://www.techbriefs.com/component/content/article/tb/features/articles/11645
by Christoph M. Greiner, Ph.D., Senior Scientist, LightSmyth Technologies

The resulting glass surface can be gold coated or simply used from the inside at an angle sufficient to produce total internal reflection. The depth of the grooves can have the same effect as blazing if the depths are chosen correctly (think about what depth that should be).

Transmission gratings are made in the same way, but without the gold coating and operation in transmission rather than reflection. They tend to have lower polarisation dependence.

Handling of gratings is very tricky. It's almost impossible to remove finger prints or dirt and cleaning can make things worse.

Let's look again at the maths behind this:

$$\int^{Area} e^{iky\sin\theta} dy = \int_0^b e^{iky\sin\theta} dy + \int_h^{h+b} e^{iky\sin\theta} dy + \int_{2h}^{2h+b} e^{iky\sin\theta} dy + \cdots \int_{Nh}^{Nh+b} e^{iky\sin\theta} dy$$

$$= \frac{e^{ikb\sin\theta} - 1}{ik\sin\theta}\left[1 + e^{ikh\sin\theta} + \cdots + e^{ik(N-1)h\,\sin\theta}\right]$$

$$= b\frac{e^{2i\beta} - 1}{2i\beta}\left[1 + e^{2i\gamma} + \cdots + e^{i(N-1)2\gamma}\right]$$

$$= b\frac{e^{i\beta}\left(e^{i\beta} - e^{-i\beta}\right)}{2i\beta}\frac{\left(1 - e^{in2\gamma}\right)}{\left(1 - e^{i2\gamma}\right)} = \frac{e^{i\beta}\left(e^{i\beta} - e^{-i\beta}\right)}{2i\beta}\frac{e^{in\gamma}\left(e^{-in\gamma} + e^{in\gamma}\right)}{e^{i\gamma}\left(e^{-i\gamma} - e^{i\gamma}\right)}$$

where $\beta = \frac{1}{2}kb\sin\theta$ and $\gamma = \frac{1}{2}kh\sin\theta$

$$= e^{i\beta}e^{in\gamma}e^{-i\gamma}\frac{2i\sin\beta}{2i\beta}\frac{2i\sin N\gamma}{2i\sin\gamma}$$

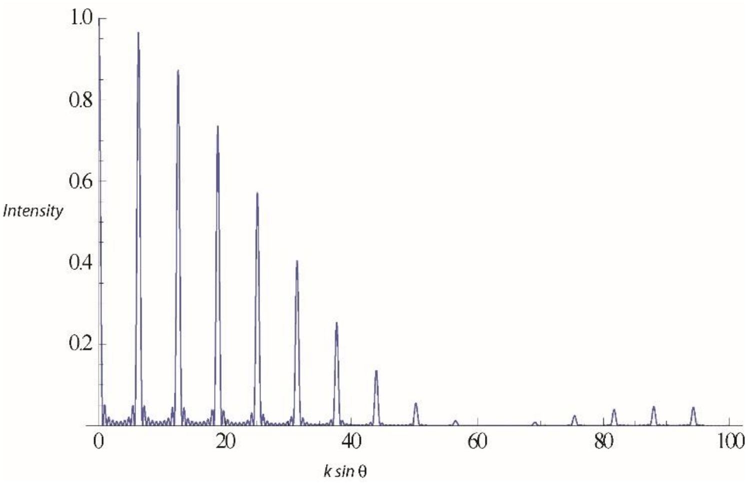
$$= e^{i\beta}e^{i(n-1)\gamma}\left(\frac{\sin\beta}{\beta}\right)\left(\frac{\sin N\gamma}{\sin\gamma}\right)$$

And the corresponding intensity is

$$I = I_o \left(\frac{\sin \beta}{\beta}\right)^2 \left(\frac{\sin N\gamma}{N \sin \gamma}\right)^2$$

where $\beta = \frac{1}{2}kb \sin \theta$ and $\gamma = \frac{1}{2}kh \sin \theta$ (and N just comes from normalising the expression)

This gives an optical output of the form



It's worth pausing here and asking what parameters determine what characteristics of the grating.

The graph is the product of the single slit diffraction pattern with etalon-like series of peaks. The number of peaks between zero degrees and the first minimum is simply the ratio of the γ to β or h to b (the slit width to the slit separation). The sharpness of the peaks is a function of N. If you look closely, you will see that secondary maximum occur at $kh \sin \theta = \frac{3\pi}{N}, \frac{5\pi}{N} \ldots$

13.1. The Resolving Power of a Grating

The angular width of a fringe is simply the first time $\sin(N\gamma) = 0$ when $\sin(\gamma) \neq 0$ i.e. $N\Delta\gamma = \pi$.

If $\gamma = \frac{1}{2}kh \sin \theta$ then

$$\frac{d\gamma}{d\theta} = \frac{1}{2}kh \cos \theta$$

so

$$\Delta\gamma = \frac{1}{2}kh \cos \theta \, \Delta\theta_1$$

And when $\Delta\gamma = \pi/N$

$$\Delta\theta_1 = \frac{2\pi}{kNh \cos \theta} = \frac{\lambda}{Nh \cos \theta}$$

Similarly the maxima occur at $n\lambda = h \sin \theta$ so

$$n \frac{d\lambda}{d\theta} = h \cos \theta$$

So the rate at which the peak of the nth order moves out as a function of wavelength is

$$\Delta\lambda = \frac{h \cos \theta \, \Delta\theta_2}{n}$$

So

$$\Delta\theta_2 = \frac{n\Delta\lambda}{h\cos\theta}$$

So to move the fringe peak over the minimum of the fringe peak of the previously measured wavelength

$$\Delta\theta_1 = \Delta\theta_2$$

i.e.

$$\frac{\lambda}{\Delta\lambda} = Nn$$

The resolving power of a grating equals the number of groves illuminated by the beam multiplied by the order of the diffraction. Note that the wavelength here is the wavelength in the material (not in free space). Gratings that use a high order of diffraction are called Echelle gratings. Let's do an example. Consider a grating with 3200 lines/mm, used in the first order. What is the resolution in GHz of this grating in the C-band (assuming a beam diameter of 1cm)?

What if the grating was internal to a material with refractive index of 3.2?

The height of the secondary maxima is given by

$$I \propto \left(\frac{\sin N\gamma}{N\sin\gamma}\right)^2$$

when $N\gamma = 3\pi/2$ which is simply $1/N^2$

When specifying a grating the key parameters are

1. Lines/mm (typical values are 1000 to 1600 but can be as high as 3000 or more)
2. Efficiency "in Littrow at 1050nm" (a typical value would be 94%)
3. Coating (usually gold)
4. Material (typically silicon, silica)

An off-the-shelf grating would typically cost a few hundred dollars.

The expression "in Littrow" means that the grating is tilted until the preferred order is being directed straight back to the source. This gives a reference efficiency value.

Questions

1. A grating is used to separate the two yellow sodium frequencies (589.0nm and 589.6nm). How many lines would you need to use?

2. A grating has 1700 lines/mm. How width would the beam need to be to resolve wavelengths 0.08nm (10 GHz) apart?

3. Why use a prism and a grating together (a GRISM) when a grating alone is so effective?

4. If you need to improve your optical resolution in a grating-based system, what choices to you have (see if you can list at least five approaches)?

Chapter 14. Fresnel Diffraction, Image Planes and An Introduction to Fiber Cables

14.1. A Brief Introduction to Fiber Cables

For most environments, an optical fiber is far too delicate to be left unprotected, so we cable it. The simplest thing we can do is put it in a furcation tube, as shown in Figure 1.

Figure 4 Furcation tube for a 12 fiber ribbon

While the tube provides protection against abrasion, it's useless against stretching. For that we need a simplex (or duplex for two fibers) cable, as shown in Figure 2.

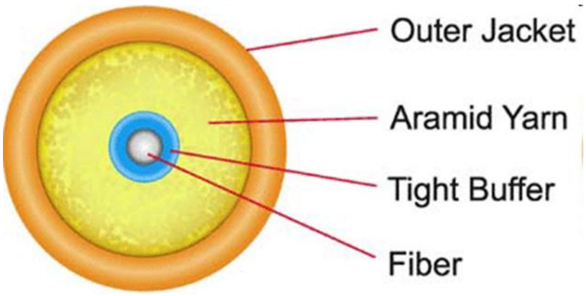


Figure 5 Simplex cable

The Aramid (Kevlar) yarn stretches less than the glass and so takes the load. This is particularly important if you're going up multiple stories in a building, so these are often referred to as riser cables. Figure 3 shows a similar riser cable with 12 fibers and a ripcord. Note that the aramid yarn is useless unless it's secured to something. Too often mechanical designers cut the yarn and leave the load on the fiber with disastrous results. Connectors will always crimp the yarn onto the connector. Make sure the yarn is also connected at the other end where it enters your device.

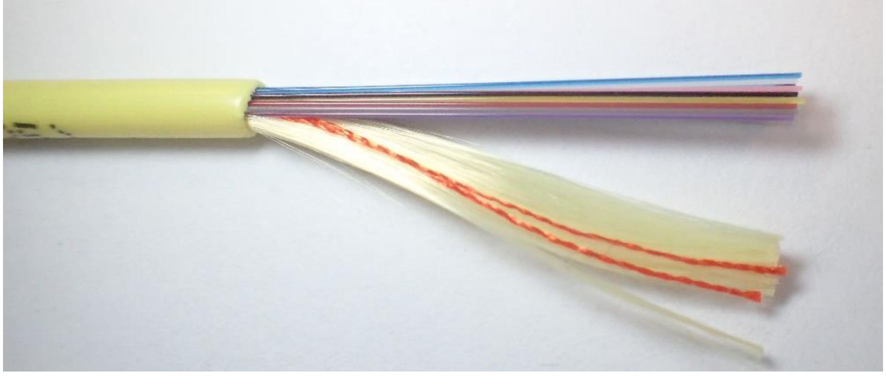


Figure 6 12 fiber riser cable

A plenum cable looks very similar, but it is designed to go through air conditioning ducts and similar spaces without acting as a fuse in a fire or create lots of fumes. This is really important since air conditioning ducts include few if any fire isolation walls. Do not use riser cable in air-conditioning ducts. Plenum cables are always labelled with the word “plenum”.

Figure 7 Plenum mini distribution cable with fiberglass strength member and rip cord. Note the two 12 fiber distribution cables (so you don’t need a splicing box to split them up.

None of these cables can be used outside. Moisture will get in and (if hot and wet enough) can swell the fiber coating and pop it off. Outdoors cables are often gel filled or filled with a powder than forms a blocking gel when exposed to moisture. The cable shown in Figure 5 has both gel inside and Aluminium Interlocking Armour (AIA). Other cables use a fiberglass armour. Both are suitable for aerial instalments. The fiberglass armour is particularly good at discouraging rats eating the cable in market environments.

Figure 8 Armoured single mode

14.2. Fresnel diffraction and Zone Plates

So far we have assumed that the light source is perpendicular to the object causing the diffraction, and non-diverging. This is often not the case. This week I want to look at a grab bag of other diffraction effects to wrap up our study of diffraction.

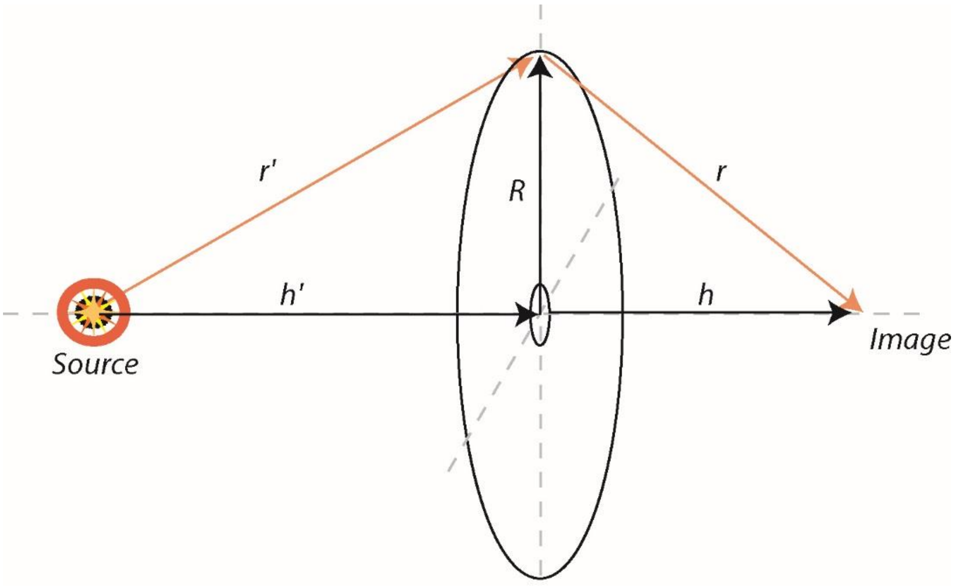


Figure 9 Fresnel Diffraction

If we consider each point on the wavefront a source of a new wave, then clearly some of these waves are going to travel different distances. Consider the wave that travels *r′* and then *r*. If the angles are small (exaggerated in the above diagram) then

$$|\boldsymbol{r}' + \boldsymbol{r}| = \sqrt{h'^2 + R^2} + \sqrt{h^2 + R^2} \approx h' + \frac{1}{2h'}R^2 + h + \frac{1}{2h}R^2 \qquad (1)$$

The disk marked in Figure 15 can be considered in a series of rings, each corresponding to a half wavelength increase in the path length. I.e.

$$\frac{n\lambda}{2} = R^2\left(\frac{1}{2h'} + \frac{1}{2h}\right) \qquad (2)$$

Alternate zones contribute positively or negatively to the overall intensity. The rest of the argument will be somewhat hand-wavy because if we evaluate the resulting integrals, we get Fresnel Integrals that can only be evaluated numerically, but we can still learn quite a lot.

Let's call the intensity contribution from each zone U_1, U_2, U_3... Then we can write the overall intensity as

$$I = |U_1| - |U_2| + |U_3| - |U_4| + |U_5| \ldots$$

$$= \frac{1}{2}U_1 + \left(\frac{1}{2}|U_1| - |U_2| + \frac{1}{2}|U_3|\right) + \left(\frac{1}{2}|U_3| - |U_4| + \frac{1}{2}|U_5|\right)$$

The terms in brackets largely (but not completely) cancel out. If we consider an area the same size as U_1 at the screen we can see that about half the intensity comes from the first Fresnel Zone (this is incidentally, the principle behind a pin-hole camera).

Now we can play a clever trick. Let's block alternate Fresnel zones with black rings. Then no cancelling occurs and all the remaining terms accumulate.

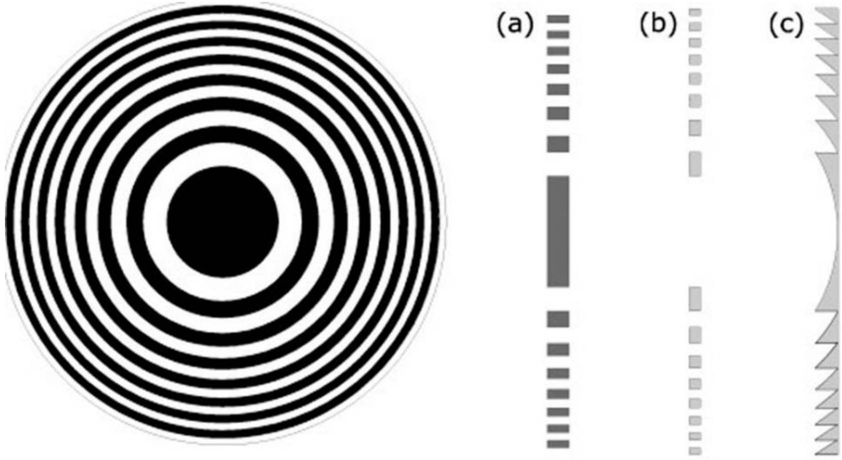


Figure 10 Fresnel Zone Plate (from http://spie.org/Images/Graphics/Newsroom/Imported/1015/1015_fig1.jpg) (b) is the equivalent phase plate and (c) the equivalent Fresnel Lens

Of course, we don't actually need to block the alternate zones, we can simply phase shift them, as shown in Figure 10 Fresnel Zone Plate (from http://spie.org/Images/Graphics/Newsroom/Imported/1015/1015_fig1.jpg) (b) is the equivalent phase plate and (c) the equivalent Fresnel Lens We can further improve the intensity by sloping the surfacing so that the entire wavefront is perfectly in phase (16c).

If we let $h^{'} \rightarrow \infty$ in equation (2) then h become the focal length. We can see that the focal length (L) is

$$L = \frac{R^2}{\lambda}$$

A more complete treatment of Fresnel diffraction is given on pages 128-134 of your textbook, which we will look at briefly. If we have a slit from x_1 to x_2 then the sum of the phases of the light will be

$$U_p = C\int_{x_1}^{x_2} e^{ikx^2/2L} = U_1\int_{x_1}^{x_2} e^{i\pi u^2/2}$$

Where C and U_1 are constants and $u = x\sqrt{\frac{2}{\lambda L}}$ (where L is the length of the slit). Since we can't solve this analytically we normalise the equation and plot it out. This gives the famous Cornu Spiral. For an infinity wide slit we have a distance of $\sqrt{2}$ so to get the resulting field strength as a fraction of the total field strength we need to divide by $\sqrt{2}$. For an edge we let one side go to -∞ and measure the length. If s=0 represents the point directly above the edge, what will the resulting intensity be (remember to square the field to get the intensity).

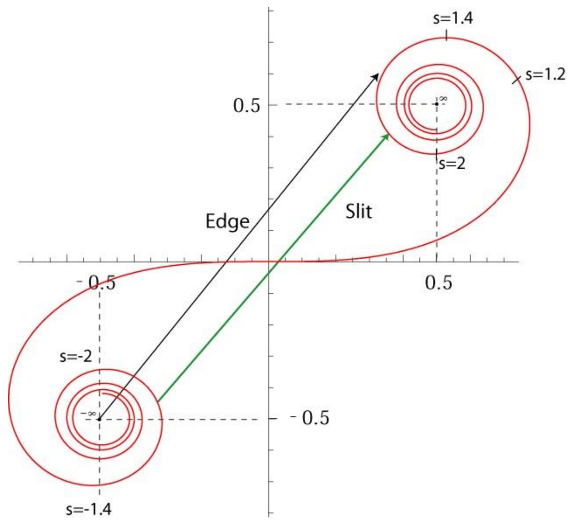
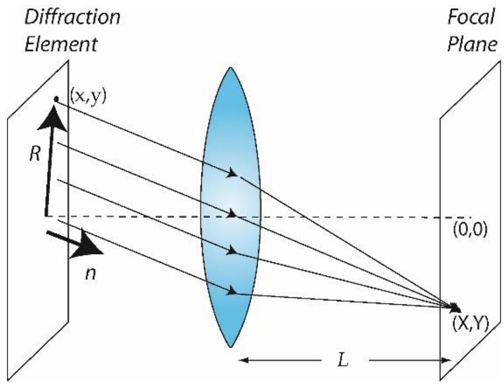


Figure 11 The Cornu Spiral can be used to calculate the resulting power from a slit of arbitrary width and offset

14.3. Diffraction and Fourier Transforms

Consider the focal plane of the focused light from a diffraction element (which could be a picture). We're back doing Fraunhofer diffraction, so assume all the beams are parallel.



We can write the direction of the beam as

$$\hat{\boldsymbol{n}} = \hat{\boldsymbol{\imath}}\alpha + \hat{\boldsymbol{\jmath}}\beta + \hat{\boldsymbol{k}}\gamma$$

where $\alpha^2 + \beta^2 + \gamma^2 = 1$ since it's a unit vector (length of 1). The focal point will be at a point *(X,Y)* where $X \approx \alpha L$ and $Y \approx \beta L$, where L is he focal length of the lens.

The path length difference between the beam coming from $\boldsymbol{R} = (x, y)$ and the origin on the diffraction element will be

$$\boldsymbol{R}.\hat{\boldsymbol{n}} = \frac{xX}{L} + \frac{yY}{L}$$

The electric field at *(X,Y)* is

$$U(X,Y) = \iint g(x,y)e^{ik\boldsymbol{R}.\hat{\boldsymbol{n}}}dA = \iint g(x,y)e^{ik\left(\frac{xX}{L}+\frac{yY}{L}\right)}\,dx\,dy = \iint g(x,y)e^{ik(\mu x+\nu y)}\,dx\,dy$$

where $g(x,y)$ is the amplitude of the wave coming from (x,y), $\mu = \frac{X}{L}$ and $\nu = \frac{Y}{L}$.
This form should look very familiar to you. It's just a Fourier transform, but in 2 dimensions. This allows all sorts of interesting filtering, either to remove or enhance periodic elements. Much image compression is also done using Fourier transforms.

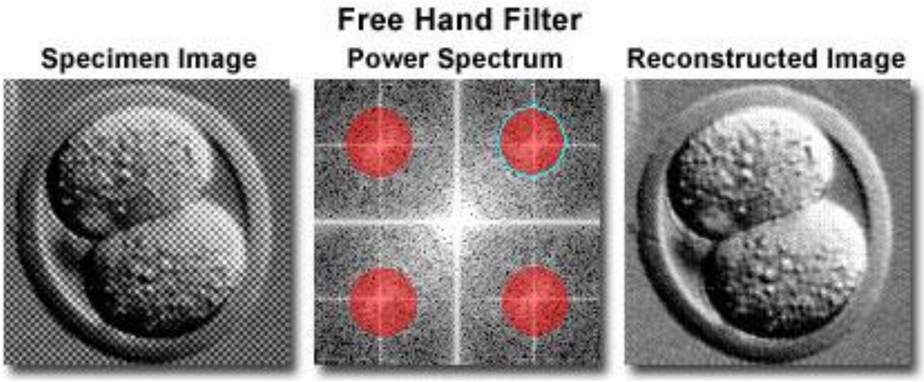


Figure 12 Removing frequency components from an image (from http://micro.magnet.fsu.edu/primer/java/digitalimaging/processing/fouriertransform/index.html)

14.4. How does a hologram work?

Consider a reflection of a plane wave from an object.

$$U(x,y) = a(x,y)e^{i\varphi(x,y)}$$

And similarly an interfering plane wave

$$U_o(x,y) = a(x,y)e^{ik(\mathrm{x}\sin\alpha + y\sin\beta)}$$

When we interfere this with the original beam then we get

$$|U(x,y) + U_o(x,y)|^2 = a^2 + a_0^2 + 2aa_o\cos(\varphi(x,y) - \mathrm{k}.(\mathrm{x}\sin\alpha + y\sin\beta))$$

The first two terms are the direct beam. The last term is the interference term that contains the phase and amplitude information of the object. If we just consider one point on the object we can see that if we convert the resulting intensity into a proportional phase delay (on say a photographic film) we create a Fresnel lens. If we then illuminate the photographic film then it will focus the light to recreate the point.

Chapter 15. The Optics of Solids Part1: Isotropic dielectrics

15.1. What slows the light?

To understand the optics of solids we need to think about atoms. An atom is a positively changed nucleus (effectively a point), with some number of changed electrons in orbit around it. The electronics sit in orbitals which can be distorted by an electronic field but will want to spring back into shape once the field is removed. If the orbital is distorted too far then the electron can move to a different orbital, usually further away from the nucleus, absorbing energy. If you've ever studied oscillators in physics (e.g. a child's swing or a weight on a spring), the maths will be very familiar. Any linear oscillator can be described by the force equation

$$m\frac{d^2\boldsymbol{r}}{dt^2} + m\gamma\frac{d\boldsymbol{r}}{dt} + K\boldsymbol{r} = -e\boldsymbol{E} \qquad (1)$$

The first term is the mass multiplied by the acceleration of the moving object, the second term is a dampening term proportional to the velocity of the moving object, the third is the force being exerted by the spring or similar object which is proportional to how much it's stretched (i.e. $\boldsymbol{r}$). On the other side of the equation is the driving force, in our case the oscillating electric field produced by the light. If we assume that the electrons orbital is distorted in a harmonic fashion (frequency ω) and is driven by an electric field (also frequency ω) then we can let

$$\boldsymbol{r} = \boldsymbol{r_o}e^{-i\omega t} \qquad (2)$$

and $\boldsymbol{E} = \boldsymbol{E_o}e^{-i\omega t}$

and the equation becomes

$$(-\mathrm{m}\omega^2 - im\gamma\omega + K)\boldsymbol{r} = -e\boldsymbol{E_o} \qquad (3)$$

So the average distance that the electrons are displaced is

$$\boldsymbol{r} = \frac{-e\boldsymbol{E_o}}{(-\mathrm{m}\omega^2 - im\gamma\omega + K)} \qquad (4)$$

The "polarisation" (nothing to do with optical polarisation) is the effective field created by this movement. I.e. the distance the changes are moved multiplied by the change density, multiplied by the charge on the electron

$$\boldsymbol{P_o} = -Ne\boldsymbol{r} = \frac{Ne^2\boldsymbol{E_o}}{(-\mathrm{m}\omega^2 - im\gamma\omega + K)} \qquad (5)$$

We can define $\omega_o = \sqrt{\frac{K}{m}}$ which is the resonant frequency. This becomes clear when we put it into equation (5)

$$\boldsymbol{P_o} = \frac{Ne^2\boldsymbol{E_o}/m}{(\omega_o{}^2 - \omega^2 - i\gamma\omega)} \qquad (6)$$

Note that, without a damping term, $\boldsymbol{P_o}$ will go to infinity at resonance.

In the absence of electric currents, the first two of Maxwell's equations become

$$\nabla \times \boldsymbol{E} = -\mu_o\frac{\partial \boldsymbol{H}}{\partial t}$$

and

$$\boldsymbol{\nabla} \times \boldsymbol{H} = \epsilon_o \frac{\partial}{\partial t}\left(\boldsymbol{E} + \frac{\boldsymbol{P}}{\epsilon_o}\right)$$

Taking the time derivative of the second equation and substituting in the first to eliminate $\boldsymbol{H}$ and using the identity $\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \boldsymbol{E}) = -\boldsymbol{\nabla}^2\boldsymbol{E}$ when $\boldsymbol{\nabla}.\boldsymbol{E} = 0$ lets us get back to the wave equation

$$\nabla^2\boldsymbol{E}(x, y, z) = \frac{1}{c^2}\frac{\partial^2(\boldsymbol{E} + \boldsymbol{P}/\epsilon_o)}{\partial t^2}$$
$$= \frac{1}{c^2}\left(1 + \frac{Ne^2/m\epsilon_o}{({\omega_o}^2 - \omega^2 + i\gamma\omega)}\right)\frac{\partial^2\boldsymbol{E}}{\partial t^2}$$

(using $c^2 = \epsilon_o\mu_o$). The expression in brackets is the refractive index squared. If the dampening term, γ, is small then

$$n^2 \approx 1 + \frac{Ne^2/m\epsilon_o}{({\omega_o}^2 - \omega^2)}$$

as we approach the resonance peak the refractive index will increase as the inverse square of the distance from resonance. At resonance it will become less than 1 and again increase back to 1.

At frequencies much less than the resonance

$$n^2 \approx 1 + \frac{Ne^2/m\epsilon_o}{{\omega_o}^2}$$

At angular frequencies much greater than ω_o, $n \rightarrow 1$. The absorption peaks on the left hand side of Figure 17 illustrate these effects reasonably well. In real materials, of course, there is not one resonance but many so the square of the complex refractive index is

$$\mathfrak{N}^2 = 1 + {}^{Ne^2}\!/_{m\epsilon_o}\sum_j \frac{f_j}{({\omega_j}^2 - \omega^2 - i\gamma\omega)}$$

where f_j is known as the oscillator strengths.

If we split the real and imaginary part of the refractive index $\mathcal{N} = n + i\kappa$, then

$$2n\kappa = {}^{Ne^2}\!/_{m\epsilon_o}\left(\frac{\gamma\omega}{(({\omega_o}^2 - \omega^2)^2 + \gamma^2\omega^2)}\right)$$

And $\mathbf{E} = \mathbf{E_o}e^{\alpha z}e^{i(kz-\omega t)}$ where $\alpha = \frac{\omega}{c}\kappa$.

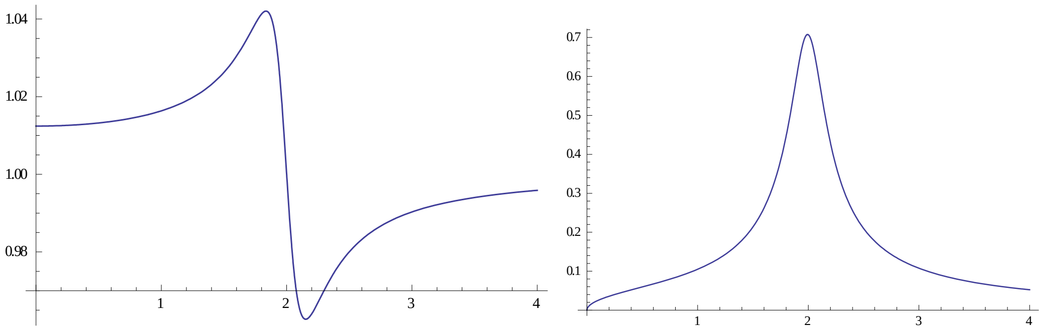


Figure 13 Shape of the real and imaginary parts of the refractive index curve at resonance

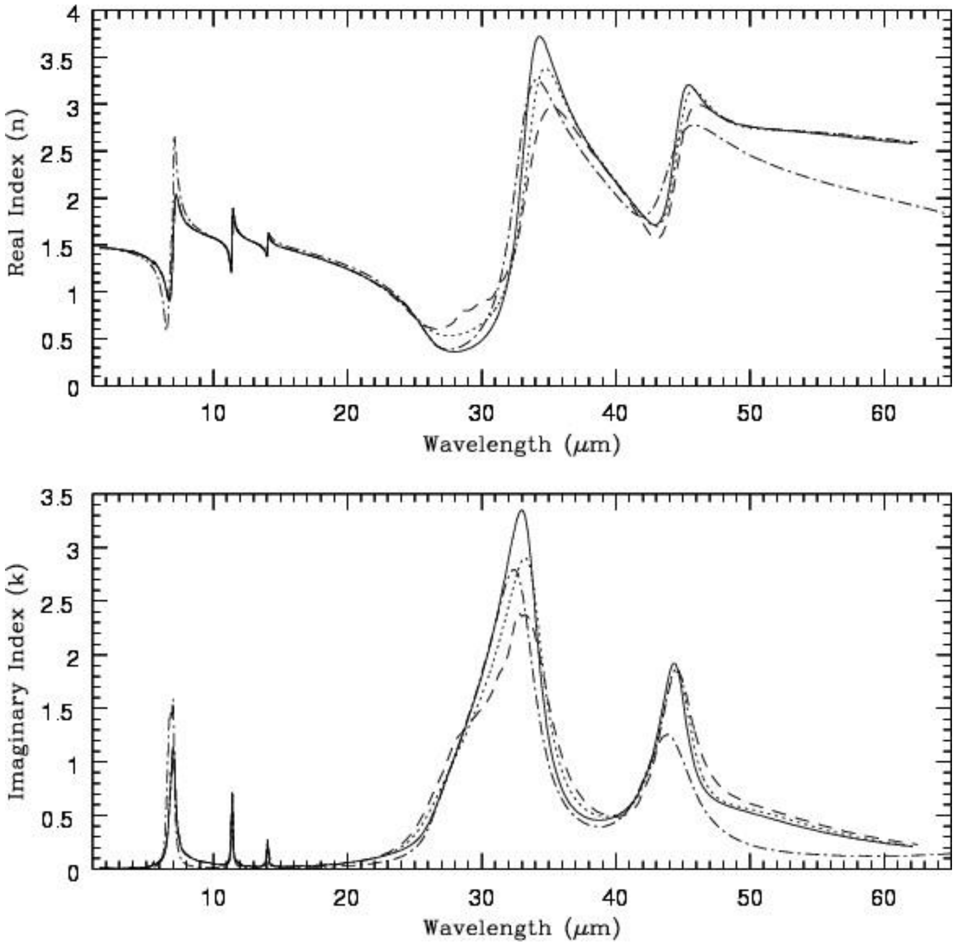


Figure 14 Refractive index (real and imaginary) from **Optical constants of powdered limestone obtained by taking into account the grain shapes: Applicability to Martian studies**
A. Jurewicz, V. Orofino, A. C. Marra and A. Blanco A&A **410 (3)** 1055-1062 (2003)

15.2. What Glasses do we use?

15.2.1.1. Abbe Number

When you pick up a glass catalogue most of the terminology will look familiar but the Abbe number is a peculiar way to talk about dispersion

$$V_D = \frac{n_D - 1}{n_F - n_C}$$

Where n_D is the refractive index at 589.3nm, n_F the refractive index at 486.1nm and n_C at 656.3nm. Large values mean low dispersion; a good crown glass has a value around 60. Low values mean high dispersion; a "flint" glass (designed for high dispersion) will have an Abbe number around 30.

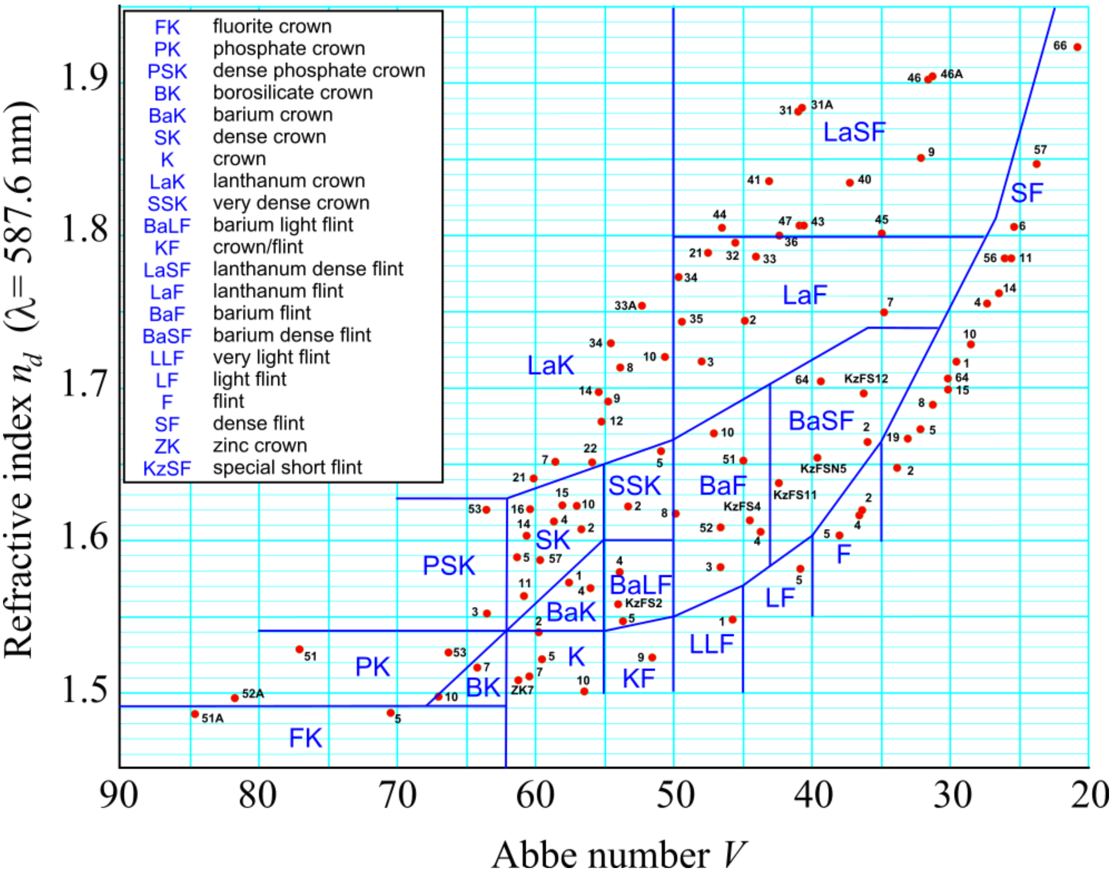
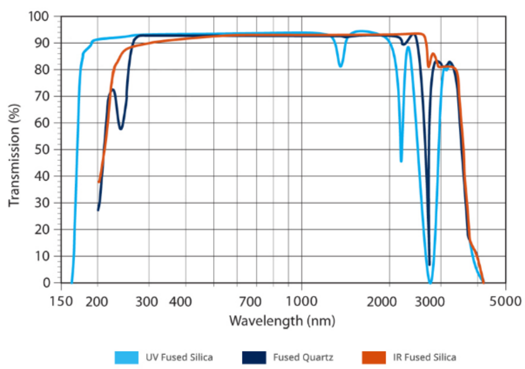


Figure 15 Abbe-diagram (by Bob Mellish Abbe-diagram.svg: Eric Bajart derivative work: Eric Bajart - Abbe-diagram.svg, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=10472043)

15.2.2. Fused Silica

Perhaps my favourite glass is fused silica. Optical fibers are typically nearly pure fused silica. It has a very low thermal expansion coefficient (0.55×10^{-6} C^{-1} , for comparison BK-7 has an expansion coefficient of 7.1×10^{-6} C^{-1}), and close to perfect transparency, chemically tough, very hard (600kg/mm^2) and an insanely high melting point (1650°C) and able to be cooled very quickly without shattering. It has an Abbe number of 67.8.



It's confusing that "pure" SiO_2 can have very different absorption curves (and names). "Fused quartz" is made by melting quartz and cooling it rapidly. It has trace amounts of metals and hydrogen creating absorption in the UV and IR respectively. By making silica by burning SiF_4 gas in oxygen, metal-free SiO_2 can be made which is usually called "UV fused silica". This reaction is used to make the silica in optical fibers.

By removing O-H bonds from the silica (less than 10 parts per million (ppm)) we can make a fiber very transparent in the infrared: "IR fused silica". For optical fibers we use both tricks since the metals also contribute to the loss at longer wavelengths.

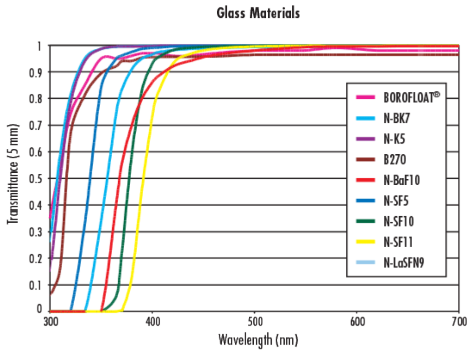


Figure 16 Absorption vs. RI (from http://www.edmundoptics.com/resources/application-notes/optics/optical-glass/)

15.2.3. N-BK7

Schott's BK7 was the standard "crown" glass (i.e. a low wavelength dependence to its refractive index) and was commonly used in lenses and as backings for filters. It's now been replaced by a lead and arsenic-free version called N-BK7. It's chemically stable, bubble free, colourless and easy to polish.

15.2.4. N-SF5

Schott's SF5 is a common flint glass (originally these glasses were manufactured by heating ground flint with lead making a silica/lead oxide mix, now titanium oxide is used to replace the lead oxide).

The reason for using a flint glass is to take advantage of its very high dispersion. In lens design we use this to counteract the effect of the dispersion of crown glasses. Both glasses have dispersion with the same sign but a weak concave flint glass lens will balance a strong convex crown glass lens. Note that this works well over visible wavelengths for these two glasses, but not over the IR, as you can see in Figure 5. The other common flint glass you'll hear about is SF6 that has a higher refractive index.

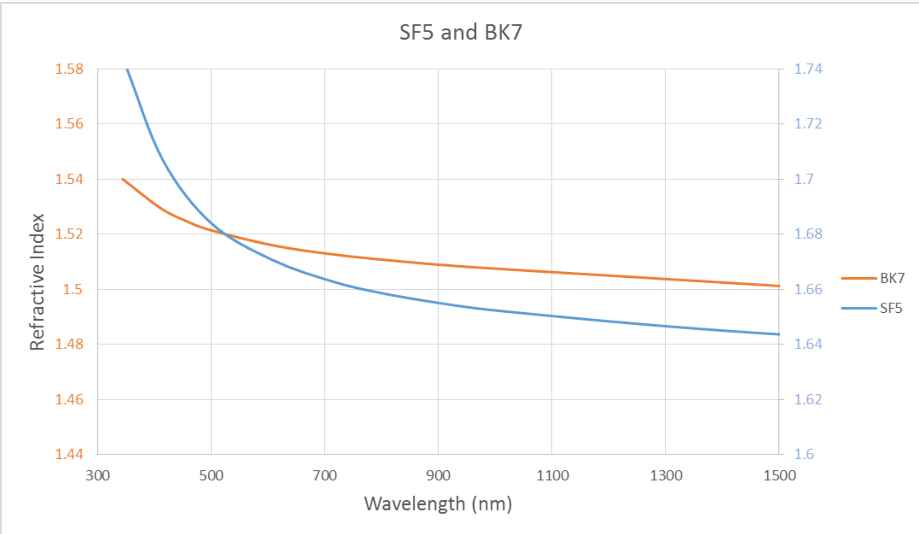


Figure 17 Comparison between a common crown glass (BK7 and a flint glass SF5)

Chapter 16. The Optics of Solids Part 2: Dielectric Crystals

Double diffraction occurs in all crystals, other than cubic crystal like common salt. Last week we saw how the refractive index could be derived by assuming that each atom was a tiny resonator. We assumed that these resonances were the same in every direction, but in a crystal this isn't true. The polarization $\boldsymbol{P}$ is related to the electric field via a tensor.

$$\begin{bmatrix} P_x \\ P_y \\ P_z \end{bmatrix} = \epsilon_o \begin{bmatrix} \chi_{11} & \chi_{12} & \chi_{13} \\ \chi_{21} & \chi_{22} & \chi_{23} \\ \chi_{31} & \chi_{32} & \chi_{33} \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix}$$

Or $\boldsymbol{P} = \epsilon_o \boldsymbol{\chi} \boldsymbol{E}$

The wave equation then becomes

$$\boldsymbol{k} \times (\boldsymbol{k} \times \boldsymbol{E}) + \frac{\omega^2}{c^2}\boldsymbol{E} = -\frac{\omega^2}{c^2}\boldsymbol{\chi}\boldsymbol{E}$$

Ignoring absorption, there are always principle axes so that $\boldsymbol{\chi}$ is diagonal.

In a uniaxial crystal such as quartz or calcite

$$\boldsymbol{\chi} = \begin{bmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & c \end{bmatrix}$$

In a biaxial crystal such as mica or gypsum

$$\boldsymbol{\chi} = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$$

The crystals we use are all uniaxial, so I'll concentrate on those. We call the polarisation in the XY plane the ordinary ray, and the corresponding refractive index, the ordinary index (n_o). We call the other refractive index the "extraordinary" ray and the corresponding refractive index, the extraordinary index (n_e). For example quartz is a uniaxial positive crystal where $n_o = 1.544 < n_e = 1.553$.

In the plane where the two refractive indices are equal, the light with behave as we expect. In the other polarisation however it's always being "pushed" sideways and this skews everything.

To convert between the matrix and the refractive index for a uniaxial crystal

$$n_o = \sqrt{1 + a}$$

And

$$n_e = \sqrt{1 + c}$$

Calcite is a uniaxial negative crystal where $n_o = 1.658 > n_e = 1.486$. Notice that for any direction a beam travels through a uniaxial crystal, one of the polarization states lie on the ordinary plane. This is referred to as the "ordinary beam"

Let's look at the maths… Going back to the wave equation

$$\boldsymbol{k} \times (\boldsymbol{k} \times \boldsymbol{E}) + \frac{\omega^2}{c^2}\boldsymbol{E} + \frac{\omega^2}{c^2}\chi\boldsymbol{E} = \boldsymbol{0}$$

So

$$\boldsymbol{k} \times (\boldsymbol{k} \times \boldsymbol{E}) + \frac{\omega^2}{c^2}(\boldsymbol{I} + \chi)\boldsymbol{E} = \boldsymbol{0}$$

Then we assume that we choose axes along the principle axes of the crystal. We can then define

$$n_1 = \sqrt{1 + \chi_{11}}$$

$$n_2 = \sqrt{1 + \chi_{22}}$$

$$n_3 = \sqrt{1 + \chi_{33}}$$

We can expand this out to give equation 6.92 in your book. Here we have three equations and three unknowns, but unless we want the answer $E_x = E_y = E_y = 0$, we can only have two independent equations, so the determinate of the corresponding matrix must be zero.

$$\begin{vmatrix} \left(\frac{n_1\omega}{c}\right)^2 - k_y{}^2 - k_z{}^2 & k_x k_y & k_x k_z \\ k_x k_y & \left(\frac{n_2\omega}{c}\right)^2 - k_x{}^2 - k_z{}^2 & k_y k_z \\ k_z k_y & k_z k_y & \left(\frac{n_3\omega}{c}\right)^2 - k_x{}^2 - k_y{}^2 \end{vmatrix} = 0$$

Assuming we are in the XY plane (i.e. $k_z = 0$), this reduces to the product of the equation of a circle and of an ellipse:

$$\left(\frac{n_3\omega}{c}\right)^2 - k_x{}^2 - k_y{}^2 = 0$$

and

$$\frac{k_x{}^2}{\left(\frac{n_2\omega}{c}\right)^2} + \frac{k_y{}^2}{\left(\frac{n_1\omega}{c}\right)^2} - 1 = 0$$

In other words, our $\boldsymbol{k}$ vector can have two different values, one in which $\boldsymbol{k}$ expands in a normal circular fashion and the other in which it's distorted into an ellipse.

16.1.1. Walk Off Crystals

The obvious result of this effect is that for a beam normal to a crystal surface, one of the polarisation states will "walk-off" the optic axis.

The walk-off angle for a beam at an angle of θ to the optical axis is

$$\varphi = -\frac{1}{n_e}\frac{\partial n_e}{\partial \theta}$$

This will generally be largest at nearly 45° to the optical axis (think about an ellipse and you'll see why).
Here are some examples

YVO_4 n_e = 2.2154, n_o = 1.9929, Δn=0.2251, φ =6.042°@ 633nm (walk-off angle)
Calcite n_e = 1.48520, n_o = 1.65578, Δn=0.17058, φ =6.205°@ 633nm

$LiNbO_3$ n_e = 2.20263, n_o = 2.28629, Δn=0.08366, φ =2.135°@ 633nm

The general rule for YVO_4 at 1550nm is 1:10, that is for every 1 mm of travel the beams will be separated by approximately 100 microns.

16.1.2. Double Refraction

Since the phase at the boundary must be constant

$$\boldsymbol{k_o}.\boldsymbol{r} = \boldsymbol{k}.\boldsymbol{r}$$

So the projection back to the boundary must be constant, i.e.

$k_o \sin\theta = k_1 \sin\theta = k_2 \sin\theta$

We can then plot out the circle and the ellipse as shown in figure 6.14 of your textbook. The ray direction as the wave leaves the crystal will be the normal to the "k surface".

16.2. What Materials do we use – and why

16.2.1. Quartz

All the waveplates we use in Finisar Australia are made from a single layer of quartz crystal sandwiched between two layers of glass. The refractive indices of quartz are $n_o = 1.54422$ and $n_e = 1.55332$ (so $\Delta n = +0.00910$). What thickness of quartz do you need to make a zero-order half-wave plate? If we made the same waveplate with calcite ($n_o = 1.65838$ and $n_e = 1.48643$ so $\Delta n = -0.1719$), what thickness would we need? Waveplates are made by sending one polarization along an ordinary axis and one on the extraordinary axis however if we send both polarisations along ordinary axes then the crystal exhibits optical activity (one circular polarisation travels faster than the other). With green light it's about 30°/mm, but this becomes very small in the infrared. Quartz comes in both left-handed and right handed crystals.

The advantages of quartz as a birefringent material is that it's very strong, easy to polish to very thin sheets, has a refractive index close to BK7 and has a very low expansion coefficient.

16.2.2. YVO_4

YVO_4 has a 24 times larger Δn that quartz are ($n_o = 1.9929$ and $n_e = 2.2154$ (so $\Delta n = +0.2225$). It's no use for waveplates but perfect for walk-off crystals as we mentioned earlier. Note though that the thermal expansion coefficents are very different for the different directions (4.46×10^{-6}/K in the ordinary axes, but 11.37×10^{-6}/K in the extraordinary axis). When gluing two rotated pieces together, this can cause lots of problems. The big difference in refractive index also limit how good an anti-reflection coating you can put on the crystal.

16.2.3. MgF_2

Magnesium fluoride has a very different wavelength dependence to quartz so a combined waveplate can be made that has very little wavelength dependence.

16.3. More on Waveplates

The cheapest waveplates are simply a relatively thick piece of quartz (or, even cheaper, polymer). The complete phase delay may be 5.25 say for a quarter waveplate. These are very wavelength dependent, and angularly dependent but if you only have one wavelength and can make sure you hit the waveplate straight on then they are just fine.

In telecommunications applications we generally use single-order waveplates which are very thin and so need to be mounted between glass sheets. These give excellent angular dependence and pretty good wavelength dependence.

Shivaramakrishnan Pancharatnam (1934–1969), the Indian Physicist came up with a way of combining 3 or 6 waveplates to create quarter-wave plates that stayed nearly constant over a huge wavelength range (typically 600 – 2700nm to within less than 1%). However you need to be careful with these that you go into them at exactly 90 degrees; at angles they can be very unpredictable.

Chapter 17. The Optics of Solids Part 3: Building Useful things…

17.1. Polarisers

One of the most common uses of a birefringent crystal is to build a Glan-Taylor polariser. These are typically made from calcite. The cut angle is near the Bewster angle ensuring nearly perfect transmission of the horizontal polarisation and total internal reflection of the vertical polarisation.

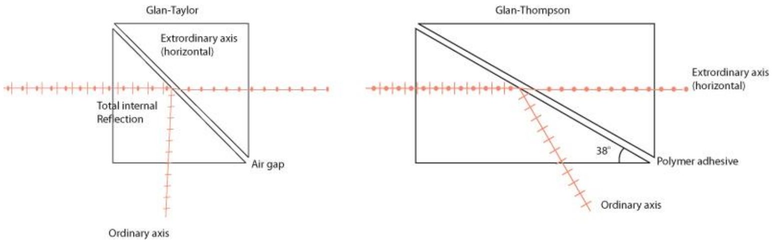


Figure 18 Glan-Taylor and Glan-Thompson polariser

The transmitted polarisation typically has an extinction ratio of 100,000:1, however the reflected beam has a small but significant horizontal component. The Glan-Thompson has a better acceptance angle (before it stops being a polariser) but the Glan-Taylor design can handle much high powers. A similar design, called the Glan-laser polariser has better transmission but a lower acceptance angle.

17.2. Isolators and Optically active Materials

Consider an atom with four covalent bonds. If there is a different structure at the end of each of its four bonds then it has a "handedness". Quartz is an example of a crystal with handedness (as is sugar, but it has disadvantages as an optical material). If $\boldsymbol{\chi}$ is of the form

$$\boldsymbol{\chi} = \begin{bmatrix} a & ib & 0 \\ -ib & a & 0 \\ 0 & 0 & c \end{bmatrix}$$

then it's eigenvectors (for a wave in the z-direction) have the form

$$E_x = \pm i E_y$$

(there's a long proof in your textbook but you can see it works by multiplying it out). Quartz has a large optical activity in the visible (49°/mm at 400nm) but is small in the infrared. Interestingly sugar is also optically active.

However there is another sort of optical activity which we use regularly called the Faraday effect. Remember when we derived the refractive index from the equations of motion of a damped oscillator we used the force equation

$$m\frac{d^2\boldsymbol{r}}{dt^2} + m\gamma\frac{dr}{dt} + K\boldsymbol{r} = -e\boldsymbol{E} \qquad (1)$$

Now let's consider what happens when we add a magnetic field to the equation

$$m\frac{d^2\boldsymbol{r}}{dt^2} + m\gamma\frac{dr}{dt} + K\boldsymbol{r} = -e\boldsymbol{E} - e\left(\frac{dr}{dt}\right) \times \boldsymbol{B} \qquad (2)$$

Now $\frac{dr}{dt} = i\omega\boldsymbol{r}$ and (assuming the effect is small) $\boldsymbol{r}$ is in the direction of $\boldsymbol{E}$ so $\boldsymbol{r} \times \boldsymbol{B}$ is perpendicular to the electric field direction. Assuming the field is in the z-direction $\boldsymbol{\chi}$ will have the form

$$\chi = \begin{bmatrix} a & ib & 0 \\ -ib & a & 0 \\ 0 & 0 & a \end{bmatrix}$$

The magnitude of the effect is defined by the Verdet constant. An angular rotation of

$$\theta = VBl$$

where B is the magnetic field strength in the direction of $\boldsymbol{k}$. “l” is the length being considered and V is the Verdet constant. In silica V is about 3.84 radians/(Tesla.meter) for red light (632.8 nm). However for Yttrium iron garnet (YIG) 1600 rad/T.m @ 632nm.

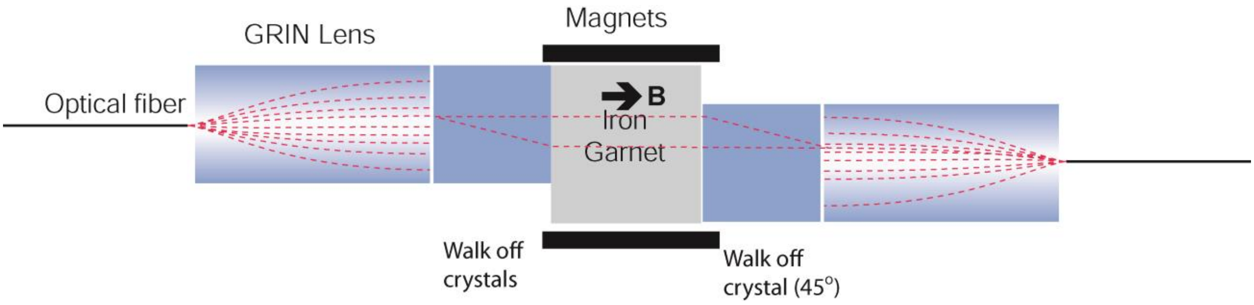


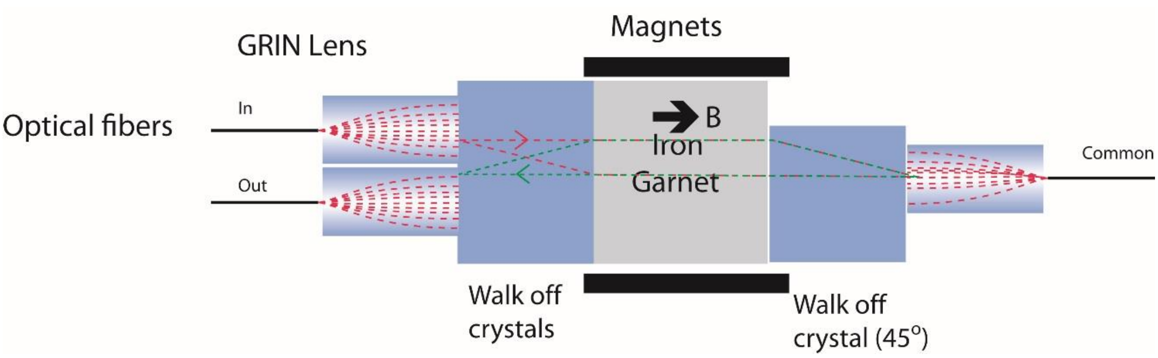
Figure 19 Design of an optical isolator. The garnet rotates the light by 45 degrees

Because the Faraday effect is directional (the rotation direction is determined by the direction of the field), it is the perfect tool to make an optical isolator (the optical equivalent of a diode). As shown in figure 20, light can travel one way through the isolator, but in the other direction, it will “walk-off” and not focus back into the fiber.

A second pair of walk-off crystals and a second garnet can be used to make a two stage isolator. These have a number of advantages. The extinction window can be much broader, and PDL and PMD can be largely cancelled out by rotating the second isolator by 90 degrees.

Isolators are used wherever you don’t want light going the wrong direction. Around WSS they are often used to prevent unintended connections between data lines. Around amplifiers they are used to prevent MPI (multi-path interference).

17.3. Circulators



A circulator works on the same principle but it doesn’t throw away the returning light. This can be very useful if your using a device that reflects light, like a sensor or a Bragg grating.

17.4. What Goes Wrong with Isolators and Circulators

17.4.1. Performance

Isolators split the polarisations so PDL (polarisation dependent loss) and PMD (polarisation mode dispersion are always potential issues. A good isolator (dual stage or with a compensating element) should have a PDL of around 0.05dB, a PMD of around 2 fs and a loss of about 0.6dB for single stage and 0.9dB for double stage. Typical cost in volume is about $20.

17.4.2. Failures

Isolators are glued together optics, made at low cost. Classic failure modes are

1. Glue breaking/delaminating from temperature cycling (test your isolators if you care!)
2. Glues "cooking" from green or blue light in the system (optical amplifiers produce green light as we'll discuss next week) – so "epoxy-free" isolators should be used with optical amplifiers
3. People packing isolators close together causing demagnetization
4. Isolators spliced facing the wrong way (it happens all the time)

17.5. Question

1. Usually we use Terbium Gallium Garnet (TGG) crystal for 1064nm high power isolator due to the high transparency. The verdet constant is 40rad/T.m. How long the crystal is if we want a 45deg rotation angle (assuming B= 1T)

Chapter 18. Photons and Optical Spectra

At the beginning of the Twentieth Century it became clear that the energy on a light wave was quantised, or lumpy. This was mathematically necessary in order to explain the optical spectra given off from hot objects, but the most powerful proof was the photoelectric effect. Light will dislodge electrons from a charged surface, but only if the light is of a sufficiently high frequency. Increasing the intensity will increase the number of electrons dislodged proportionally but increasing the intensity of light that is too low a frequency has no effect. Each photon is dislodging one electron.

The energy of a photon can be measured:

$$E = h\upsilon$$

"h" is Plank's constant 6.62606896×10^{-34} J.s

The momentum of a photon can be calculated using the equation $E = mc^2$. If

$h\upsilon = mc^2$ then the momentum

$p = mc = \frac{E}{c} = \frac{h\upsilon}{c} = \frac{h}{\lambda}$ (Equation 19.1)

So how much energy does a photon have? Consider a photon at 1550nm, $\upsilon = 193$THz, so the energy of the photon is 1.28×10^{-19}J. This number is close to the charge on the electron (1.6× 10^{-19} C), so often we write these values in electron volts (the energy needed to move an electron through 1V). One 1550nm photon has enough energy to move an electron through 0.8V ($\frac{1.28\times10^{-19}}{1.6\times10^{-19}} = 0.8\text{eV}$).

Conversely, the maximum voltage that can be produced by a (unbiased) PIN photdiode corresponds to the lowest energy photons that can be detected. If the longest wavelength a detector can measure is 1550nm (0.8eV), then the maximum voltage the PIN photodiode can produce will be 0.8V.

18.1. Angular Momentum

A rotating electric field will twist a surface it hits. This means that right and left circularly polarised light has angular momentum. The angular momentum of a photon is $\frac{h}{2\pi}$ (often written $\hbar$). Notice that this is independent of frequency.

Louis de Broglie proposed that electrons also were "wavelike" with a wavelength of

$$p = mu = \frac{h}{\lambda}$$

so $\lambda = \frac{h}{p} = \frac{h}{mu}$.

18.2. Spectra

To understand lasers and optical amplifiers we need to first have some intuition regarding electrons in orbitals around an atom. If we ignore quantum mechanics, we could simply say that the electrostatic attraction between the electron and the nucleus must equal the centripetal force needed to keep the electron moving in a circle, i.e.

$$\frac{e^2}{4\pi\epsilon_0 r^2} = \frac{mu^2}{r} \qquad \text{(Equation 19.2)}$$

This would allow the electron to be any distance from the nucleus. However we also know that we can only have a whole number of waves in an orbit, otherwise we would get destructive interference. I.e.

$$2\pi r = n\lambda = \frac{nh}{mu}$$

or rearranging this, the angular momentum

$$mur = \frac{nh}{2\pi} \qquad \text{(Equation 19.3)}$$

so only particular angular momentums are allowed. We say that it is "quantised".

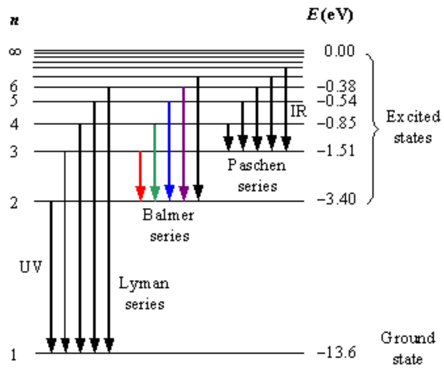
We can take equations 19.2 and 19.3 and eliminate the velocity *u* to get the allowed radii of the orbitals

$$u^2 = \left(\frac{nh}{m2\pi r}\right)^2 = \frac{re^2}{m4\pi\epsilon_o r^2} \quad \text{so} \quad r = \frac{\epsilon_o h^2}{\pi m e^2} n^2 = a_H n^2$$
(Equation 19.4)

These are all values we know so we can calculate the first Bohr radius $a_H = 0.0529\text{nm}$.

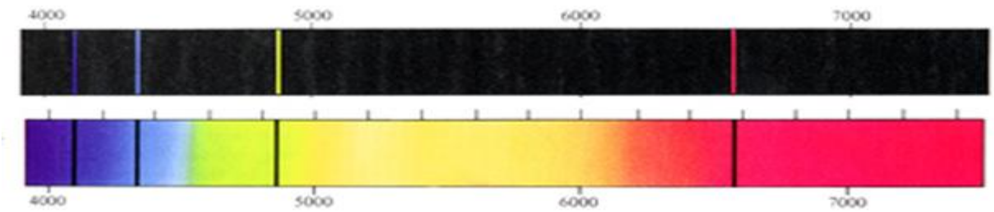
The total energy of the orbital is simply the potential energy plus the kinetic energy

$$E = \frac{1}{2}mu^2 - \frac{e^2}{4\pi\epsilon_o r} = -\frac{e^2}{8\pi\epsilon_o r} = -\frac{e^2}{8\pi\epsilon_o a_H n^2} = -\frac{R}{n^2}$$
(Equation 19.5)



Energy levels of the hydrogen atom with some of the transitions between them that give rise to the spectral lines indicated

where *R* is called the Rydberg constant. So now we can calculate the energies needed to move electrons between these orbitals. This corresponds to the absorption and emission lines of hydrogen atoms. We've assumed that the neculus is perfectly stationary and the electron moves, but in reality they both have mass. So the nuculus wobbles as the electron orbits it (in this model) so we need to adjust the effective mass of the electron to



The Hydrogen Spectrum
Above: Emission
Below: Absorption

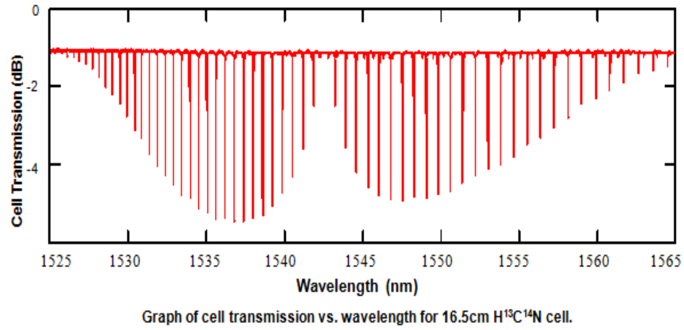
Figure 20 Hydrogen absorption and emission spectrum (from http://honolulu.hawaii.edu/distance/sci122/Programs/p27/p27.html)

$$\mu = m\frac{1}{1+\frac{m}{M}} \quad \text{(Equation 19.6)}$$

For normal hydrogen $\frac{m}{M} = \frac{1}{1836}$. From equations 19.4 and 19.5 you can see that the Rydberg constant (and the energy of the electron in the orbital and the frequency of the emitted light) is proportional to the electron mass.

18.3. HCN Absorption lines

In telecom the most useful spectral absorption lines at from HCN (hydrogen cyanide gas). As you can see in figure 2, the lines cover the whole C-band and are very sharp and can fairly easily let you calibrate a system down to 0.2 GHz.



R Branch	Wavelength (nm)	P Branch	Wavelength (nm)
26	1527.63342(12)	1	1543.11423(5)
25	1528.05474(15)	2	1543.80967(18)
24	1528.48574(9)	3	1544.51503(8)
23	1528.92643(6)	4	1545.23033(7)
22	1529.37681(7)	5	1545.95549(7)
21	1529.83688(6)	6	1546.69055(8)
20	1530.30666(8)	7	1547.43558(24)
19	1530.78615(8)	8	1548.19057(7)
18	1531.27537(7)	9	1548.95555(4)
17	1531.77430(8)	10	1549.73051(4)
16	1532.28298(8)	11	1550.51546(5)
15	1532.80139(7)	12	1551.31045(9)
14	1533.32954(8)	13	1552.11546(10)
13	1533.86745(7)	14	1552.93051(9)
12	1534.41514(6)	15	1553.75562(12)
11	1534.97258(6)	16	1554.59079(10)
10	1535.53981(5)	17	1555.43605(11)
9	1536.11683(4)	18	1556.29141(15)
8	1536.70364(5)	19	1557.15686(15)
7	1537.30029(6)	20	1558.03240(15)
6	1537.90675(13)	21	1558.91808(14)
5	1538.52305(7)	22	1559.81389(14)
4	1539.14921(12)	23	1560.71983(10)
3	1539.78523(9)	24	1561.63593(9)
2	1540.43120(10)	25	1562.56218(13)
1	1541.08703(10)	26	1563.49859(16)
0	1541.75280(6)	27	1564.44519(21)

Figure 21 HCN absorption lines from http://www.wavelengthreferences.com (note that wavelength references sell these cells for about $800 with fiber optic pigtails).

For the L-band use a CO (carbon monoxide) cell, and for the S-band, an acetylene gas cell. This is a NIST standard method for measuring frequencies, so any measurement using this technique is "NIST traceable" (NIST SRM 2519).

To accurately find the bottom of an absorption peak it's important to curve fit. Simply looking for the lowest value in a band isn't very accurate because the bottom of the curve is flat.

18.4. What goes wrong…

The numbers given in figure 2 are for a pressure of 25 Torr. There's a shift of +1.6 to +2.3 MHz/Torr (12 to17 MHz/kPa) for the P20, P23, P24 and P27 lines according to Sarah Gilbert's NIST paper, other lines move in the opposite direction. Temperature changes also have a slight effect, so measuring as close to the standard 22±5°C is a good idea.

If you need more details:

https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication260-137.pdf

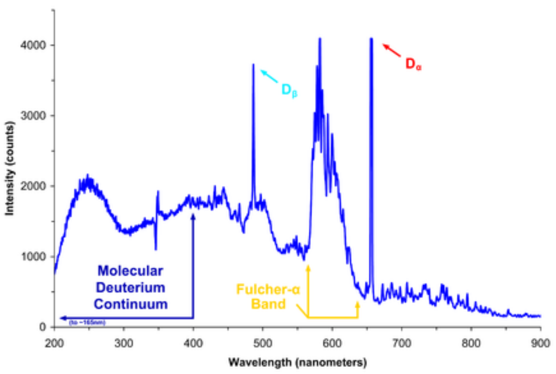
18.5. Questions

1. What minimum voltage would you expect to need to produce red light at 775nm from an LED?

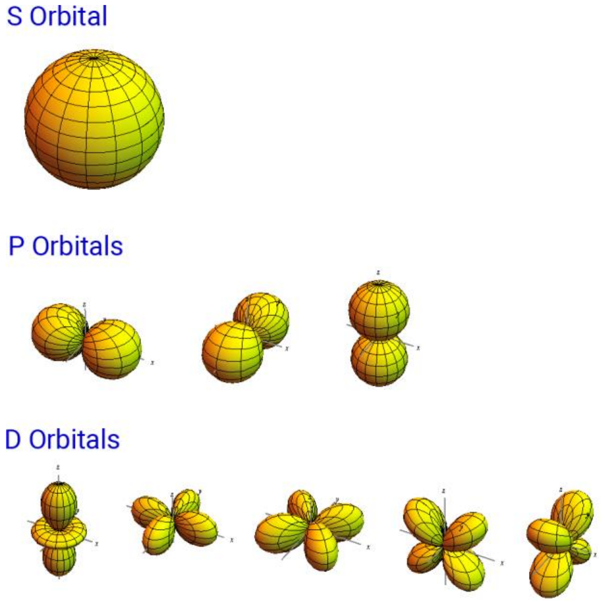
2. A perfect photodiode converts one photon into one electron. What is the maximum electric current 10mW of 1500nm light can produce? (use h= 6.6×10^{-34} J.s, and the charge on an electron of 1.6×10^{-19} C)

3. Using equation 19.1 ($p = \frac{E}{c}$) figure out if you could use the momentum of photons to sail using light in space? Consider a reflecting aluminium sail that's 100x100m square (10^4 m^2). In space sunlight has an intensity of about 1.4kW/m^2. Calculate the total force from the sunlight (remember that by reflecting the light you get to transfer twice the momentum). How long would it take to get a 1000kg space craft to 100m/s?

4. Deuterium has twice the mass of hydrogen. The first Balmer line for hydrogen is 656.3nm (456.8THz). Using equation 19.6 work out the approximate frequency for the first Balmer emission line for deuterium.



Chapter 19. Meta-stable states, and optical amplifiers



19.1. Angular momentum

Last week we considered the s-orbitals of the hydrogen atoms.

While Niel Bohrs model of the hydrogen atom gives a non-zero angular momentum to the "s" (circular) orbitals, a proper Quantum mechanical treatment gives a zero momentum. The "p" orbitals have an angular momentum "number" of $l = 1$ and the "d" orbitals an angular momentum number of $l = 2$, and the "f" orbitals have an angular momentum number of $l = 3$ where the angular momentum

$\boldsymbol{l} = \sqrt{l(l+1)}.\hbar$

and $\hbar = \frac{h}{2\pi}$.

The z-component of the angular momentum of atoms turns out to be

$$l_z = m.\hbar$$

where $-l \leq m \leq l$ and is integer. The total angular momentum $\boldsymbol{j}$ is the orbital angular momentum $\boldsymbol{l}$ plus the "spin" $\boldsymbol{s}$, where $s = \pm\frac{1}{2}$.

19.2. Fine Structure

In orbitals other than the s-orbital, the magnetic fields from the spin interact with the magnetic field from the orbit to give the electrons different energy levels depending on the spin. In the resulting spectrum, this is called the "fine structure".

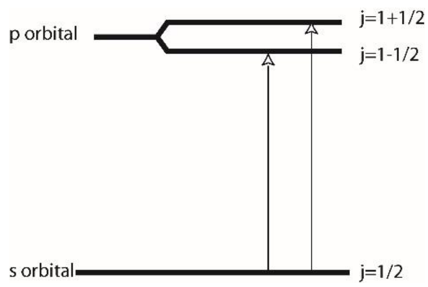


Figure 22 Fine structure splitting

19.3. Zeeman Effect

If there is an external magnetic field, this field will interact with the orbitals as well, not only will the resulting emmisions be split, but they will have different polarizations.

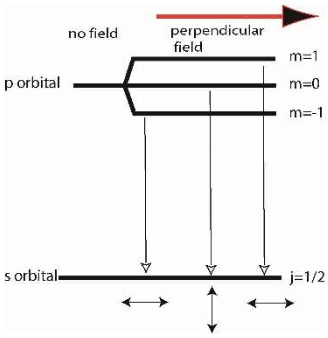


Figure 23 Zeeman splitting (l only)

19.4. Metastability

The total angular momentum for an atom is the sum of the angular momentums from each of the electrons (i.e. $\boldsymbol{L} = \sum \boldsymbol{l_i}$, $\boldsymbol{S} = \sum \boldsymbol{s_i}$ and $\boldsymbol{J} = \sum \boldsymbol{J_i}$). States are written

${}^{2S+1}L_J$ Where the "*L*" is replaced by *S,P,D,F,G,H,*I etc for *L*=0,1,2,3,4,5,6,7… respectively. When a photon is emitted by an atom, angular momentum is transferred to the photon. As we discussed in the last talk, the angular momentum of a photon is $\hbar$. So this angular momentum must be lost from the atom. The "selection rules" (shown on page 252 of your textbook) are

$$\Delta L = \pm 1,0$$

$$\Delta S = 0$$

$\Delta J = \pm 1,0$ but the transition $J = 0$ to $J = 0$ is not allowed.

Furthermore, the total angular momentum of the atom ($\boldsymbol{L} = \sum \boldsymbol{l_i}$) must go from an odd number to an even As a result, there are excited state orbitals that can't be escaped by simply emitting a photon. These excited states are called "metastable" states.

19.5. The effect of temperature

Heat can be thought of as how much energy each atom has for each "degree of freedom" (or way it can move). Each degree of freedom has an energy of ½kT, where "k" is Boltzmann's constant 1.38×10^{-23} J K^{-1} . An individual atom can move in 3 dimensions so has an average energy of $^3/_2$kT. A diatomic molecule can spin in two axes so has average energy $^5/_2$kT. However because electron energy can only be an certain levels (it's quantised) we have to ask what the probability electrons will make it to higher energy levels. The ratio of electrons in one level N_1 compared to another N_2 is

$$\frac{N_2}{N_1} = \frac{e^{-E_2/kT}}{e^{-E_1/kT}}$$

What this means in practice is that when we heat a material more of the electrons will move into higher energy levels but there will always be more electrons in the lower levels.

19.6. Optical Amplifiers

19.6.1. A side track: Optical Couplers

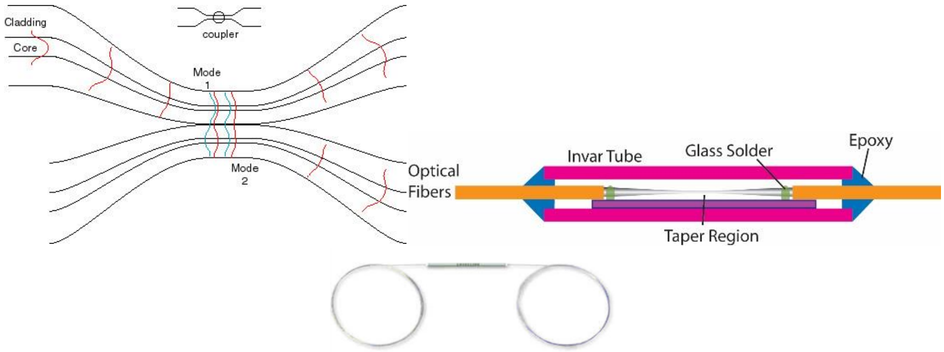


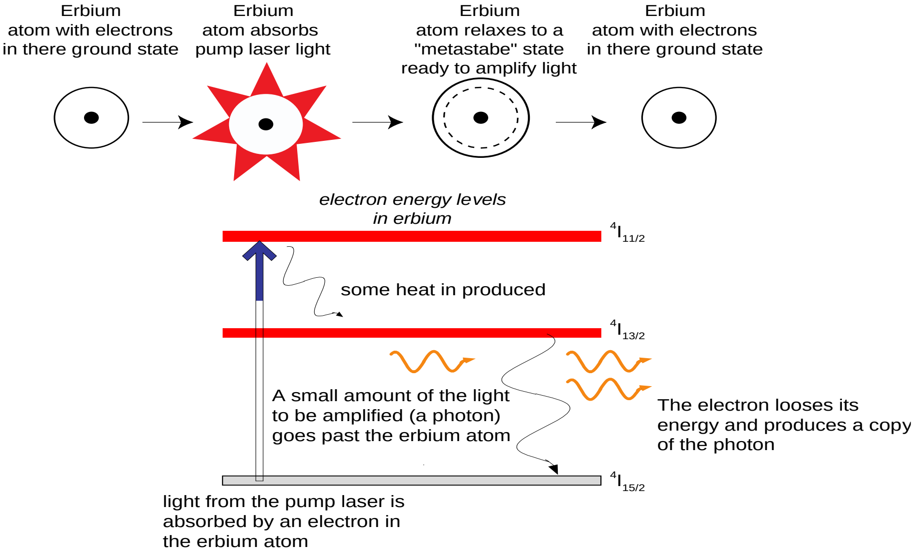
Figure 24 Three ways of looking at an optical fiber coupler

Before explaining optical amplifiers, you'll need to understand how optical couplers work. Couplers are made by heating two fibers and stretching them until the mode fields overlap and power can couple from one to the other. This can be used to tap off a small amount of light, or to combine two different wavelengths.

19.6.2. Optically Stimulated Emission

Optically stimulated emission was proposed by Einstein in 1917 and first observed in 1960. A photon can distort an excited metastable orbital, allowing the electron to move to a lower energy state. The resulting photon that is emitted will have the same polarisation, frequency and direction as the photon that created the distortion- the light is amplified. We'll consider this in much more detail next week, but this week I wanted to introduce the optical amplifier.

19.6.3. Gain in Erbium



By adding a small amount of erbium to the core of an optical fiber, we can create an optical amplifier.

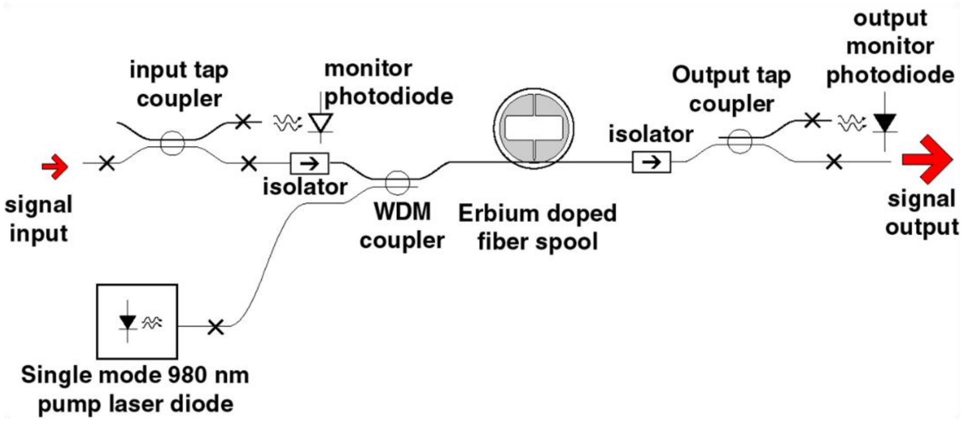


Figure 25 Design of a single stage optical amplifier

Note that this is a three level system. A four level system is nicer because then you can drain the lower state, making it easier to get an "inversion".

In practice, most amplifiers use two or more stages with a gain flattening filter and often a variable attenuator in the middle to ensure that the gain is approximately equal over the spectrum. In the past a "mid-stage access point" was often placed in the middle of a two-stage amplifier to allow various lossy elements to be added such as spools of negative dispersion fibre, but this is less common now. The pump light can go in either direction, but "forward pumping", as shown here, is more common.

The critical parameters for an amplifier are:

- Gain (typically about 17-25 dB with 20 dB and 23 dB fixed gain amplifiers being very common)
- Maximum output power (this depends on the pumps. Calculate the output power in mW for 100 channels each at -3dBm)
- Gain tilt – how much the gain varies across the spectrum.

19.6.4. Noise figure

This is a measure of how much additional noise is added to a perfect (shot-noise limited) signal. This is always at least 3dB, with 4.5 to 6dB being typical. Any losses before the start of the erbium fibre add directly to the noise figure. Because some electrons emit light spontaneously and this light is also amplified, there is always noise.

Figure 1 sketches the input (red solid line), output (GREEN), and the component of the output that represents the amplified gain (red dotted line). The gain is:

$$G = \frac{P_{out} - P_{ASE}}{P_{in}}$$

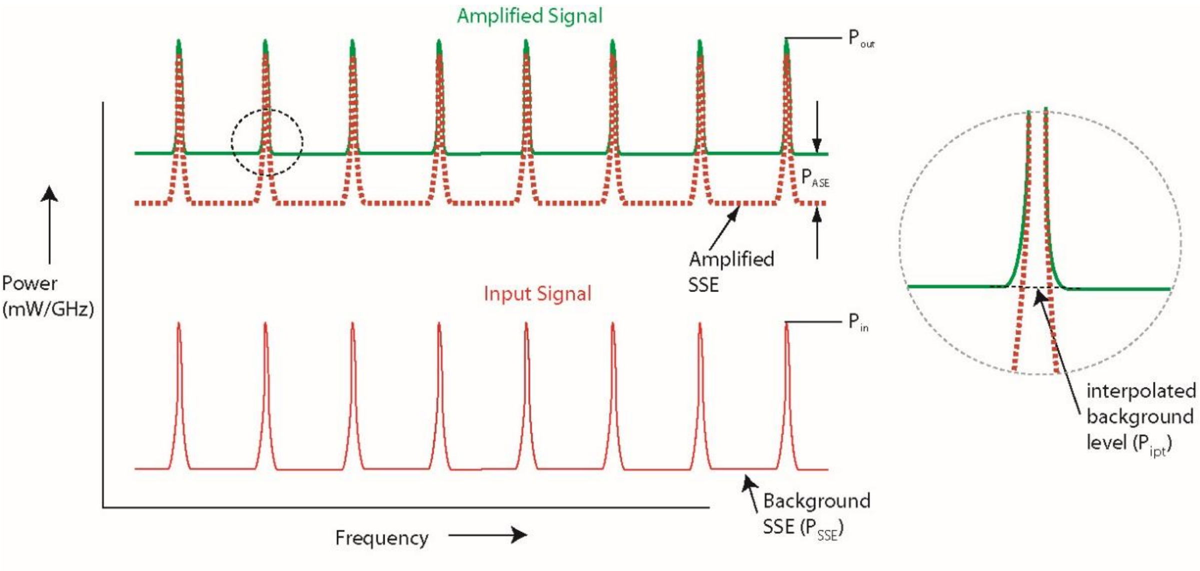
Note that the signal peaks and the laser SSE (spontaneous source emission) are all amplified. To calculate the amplified spontaneous emission (P_{ASE}) we need to interpolate under the amplified signal to find the power level without the peak, and then subtract the amplified SSE.

$$P_{ASE} = P_{ipt} - G.P_{SSE}$$

This calculation can be extremely inaccurate if the gain is small and P_{SSE} is large. The noise figure is the ratio of the noise from the amplifier ($\frac{P_{ASE}}{GB_o}$) divided by the shot noise $h.v$

$$F = \frac{P_{ASE(\text{interpolated})}}{G.h.v.B_o}$$

Where G is the amplifier gain, h, Plank's constant, v the frequency and B_o the bandwidth of the OSA.



Note that losses before an amplifier add directly to the noise figure (in dB). Losses after an amplifier have no significant effect.

19.7. Other important things to know about optical amplifiers

The gain of an erbium fiber is temperature dependent so it is almost always necessary to include a heater under the fiber. As the temperature increases the gain goes down, but more at the center of the C-band than at the edges making gain equalization difficult.

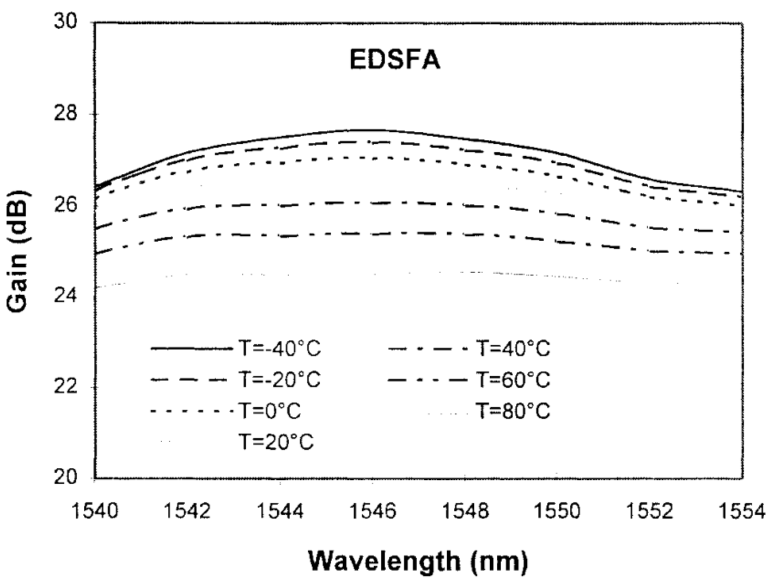


Figure 26 Temperature dependence of the gain on an EDFA (from : "Gain temperature dependence of erbium-doped silica and fluoride fiber amplifiers in multichannel wavelength-multiplexed transmission systems" J. Kemtchou; M. Duhamel; P. Lecoy , Journal of Lightwave Technology, Year: 1997, Volume: 15, Issue: 11)

Erbium also upconverts into the green so any components in the optical path before the isolators must be tolerant to green light (The erbium atom absorbs a photon pumps in the ground state to get to 4 I 11/2. It can then absorb another photon, or take energy from another excited ion, to populate 4 F 7/2 that non-radiatively decays to 4 S 3/2. Transitions from here to the ground state produce photons at 546 nm).

19.8. Transients

Perhaps the most difficult part of EDFA design is minimising gain transients. A sudden drop in the number of channels being carried can result in an increased inversion in the erbium atoms if the automatic gain control circuit is too slow.

19.9. Modelling Programs

If you are using OFS fibers then they have free modelling programs available (OASiX). Finisar has (or had) its own internal modelling program, and there are various commercial programs such as RPFiber Power.

19.10. Raman Amplifiers

Distributed Raman amplifiers – which we'll talk about later, can be used to reduce the noise figure by creating gain in the transmission fiber before the light reaches the amplifier.

19.11. Questions

1. (From Q1 in your textbook) Sodium lamps are bright yellow (590nm) due to a transition from 3p to 3s (actually two lines very close together). What is the energy in electron volts?

   What proportion of the electrons are in an excited state at 250°C?

2. Calculate the noise figure of an optical amplifier with gain of 20dB (100), ASE power of 50nW/GHz (note this includes the bandwidth so you can ignore that), at 1550nm ($hv = 6.6 \times 10^{-34}$Js). Express your answer in dB.

Chapter 20. Lasers

A laser is a light amplifier with feedback, so a laser consists of a gain medium and a cavity.

Boltzmann proposed that at thermal equilibrium, the ratio of electrons in energy state E_1 to the electrons in energy state E_2 will be

$$\frac{N_1}{N_2} = \frac{e^{-E_1/kT}}{e^{-E_2/kT}}$$

So in equilibrium, if $E_1 > E_2$, then N_1 will never be equal or greater than N_2. This is a problem, because the probability of an incoming photon pushing an electron from E_1 to E_2 (B_{12}) is the same as the probability of an electron in the upper (metastable) state being stimulated to emit a photon. Since, at equilibrium, there are more electrons in the lower state, on average more photons will be absorbed than emitted.

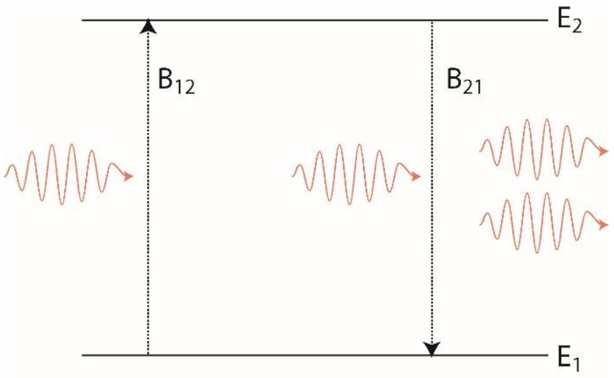


Figure 27 The probability of an electron in the lower energy state absorbing a photon and moving up to the upper state is exactly the same as the probability of an electron in the upper state being stimulated to emit a photon and dropping down to the lower state.

20.1. Creating an Inversion

To overcome this, we need to create a population inversion. There are four common ways of producing an inversion:

1. Optical pumping: The material is illuminated with a wavelength that excites electrons into an orbital higher than the meta-stable state. The material emits a phonon (vibration energy) to transfer the electron into the meta-stable state. Examples include optical fiber based lasers and Nd-YAG slab lasers
2. Electrical pumping: Semiconductor lasers often allow electrons to be injected directly into the excited (conduction) band.
3. Collision pumping: An electron can be excited in an atom by colliding with another excited atom.
4. Chemical reaction: When H_2 reacts with F_2 excited HF is formed. This forms the basis if the HF laser.

20.2. Three and Four Level Systems

In order to get lasing we need to be able to get electrons out of the ground state without de-exciting the electrons in the excited state. The simplest way to do this is a four-level system, such as a Helium -Neon laser, where electrons get excited into a high energy orbital, decay into a metastable state and transition (via lasing) into a state above the ground state. This requires much less than 50% inversion since the lower state is continually draining into the ground state.

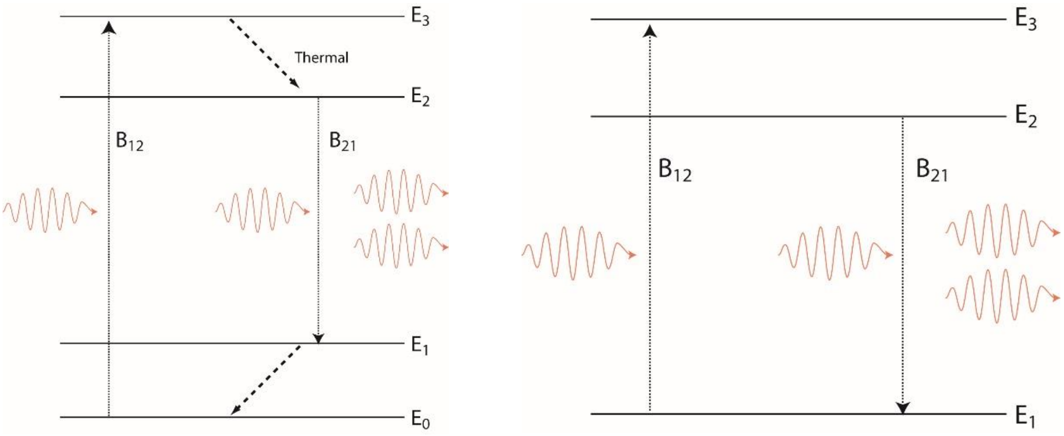


Figure 28 Three and four level laser systems

Three level systems, such as Er^{3+} in glass also work, but require more than 50% inversion between the ground state and the metastable state.

20.3. Conduction and Valance Bands

In a semi-conducting material, the orbitals of the atoms in the material merge into bands. The valance band is usually full and the conduction band empty. If there are a few electrons in the conduction band, the material is refer to as an N-type semiconductor (generally caused by adding a tiny amount of dopant material with an extra electron). If there are a few gaps in the valance band, then it's referred to as a P-type (made by adding a dopant with fewer electrons in its outer shell). If there are no dopants present then the semiconductor is "I" type. Unfilled orbitals in in the valance band are called "holes" (the equivalent of bubbles).

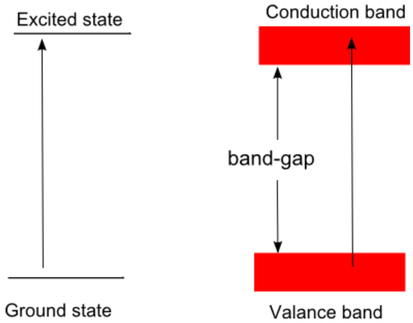


Figure 29 In a semiconductor the orbitals of the whole crystal are coupled forming bands

However, there are still energies that an electron can't take. These energies are called the band-gap. InGaAs (Indium Galium Arsenide) has a direct band gap. The lowest energy states in the conduction band have the same momentum as the highest energy states in the valance band. This allows a photon with energy equal to the band gap to stimulate an electron to move from the conduction band to the valance band producing a new photon.

Electrons in the conduction band tend to move to the lowest energy level within the band (by emitting phonons - heat energy). Similarly, holes in the valance move to the top of the valance band. If these two regions have essentially the same electron momentum then a photon with energy equal to the band gap can stimulate a downward transition. We can make efficient lasers from InGaAs which is a direct band-gap material, but not from silicon, which is not.

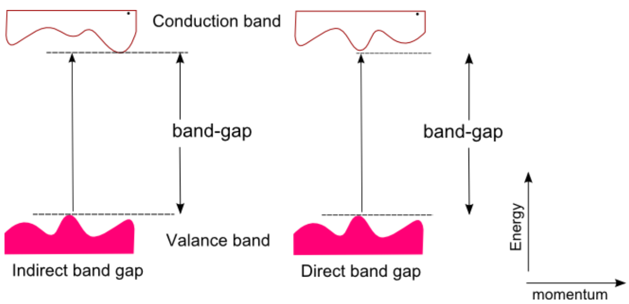


Figure 30 Direct and indirect semiconductors

20.4. The Gain Curve

The rate α, at which a beam, frequency range $\Delta\upsilon$, increases in power in a gain medium is proportional to the transition constant B_{12}, the energy gain per transition $\frac{h\upsilon}{c}$, and the difference between the electron density in the excited state to the ground state $(N_2 - N_1)$. I.e.

$$\alpha_\upsilon.\Delta\nu = \frac{h\upsilon}{c}(N_2 - N_1)B_{12}$$

And the resulting intensity will increase as

$$I_\upsilon = I_{o\upsilon}e^{\alpha_\upsilon x}$$

To achieve gain $N_2 - N_1 > 0$ and the total gain must be enough to cancel out the losses in the laser cavity. Of course there is a natural feed-back situation here, since the gain is depleting N_2 and increasing N_1.

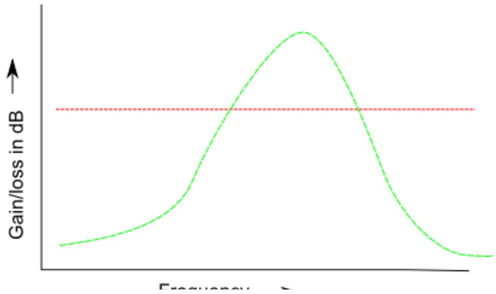
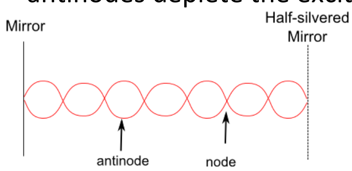


Figure 31 Laser gain depends on frequency

However not all frequencies are allowed in a laser cavity. For a standing wave in the laser cavity, the round trip distance must be a whole number of wavelengths. The laser modes differ by the free spectral range of the cavity.

$$\upsilon_{n+1} - \upsilon_n = \frac{c}{2d}$$

This will cause electron orbitals with the matching energies to be depleted (spectral hole burning). The standing waves associated with the modes will also cause “spatial hole burning” as the antinodes deplete the excited electrons and the nodes accumulate them. This will give an advantage to other modes which can exploit these regions.



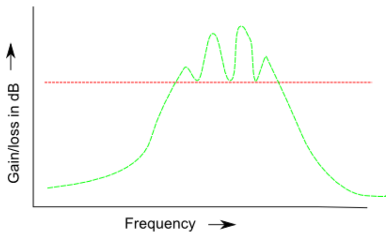


Figure 32 Gain and loss is affected by spectral hole burning

20.5. Mode Locking

Modes will naturally order themselves to maximise the rate of depletion. By adding a material that can absorb light, but only to a certain level and then saturate, to the cavity we can encourage the modes to self-organise into pulses. There was a great example of one of these at a CLEO laser conference. A Japanese group presented a tiny optically-pumped composite Nd:YAG laser that fits inside a spark plug fitting. The laser was pumped by laser diode light delivered by fibers and consisted of a crystal with Nd:YAG at the back acting as the gain medium and Cr4+ YAG at the front acting as a saturable absorber. The laser was pumped with a pulse from the laser diodes. Being a single crystal, N. Pavel’s laser could withstand the extreme engine temperatures and allowed a three-point ignition. It looked very impressive. It certainly looks like an interesting mass market for lasers. The flashes in the photo below are real air breakdown events as the beams focus in the two-point ignition version. It’s so simple it might just change the motor engine. They achieved pulse of a few tens of fs.

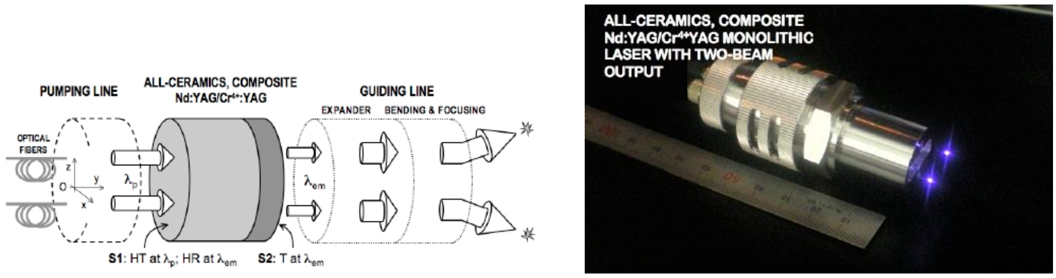


Figure 33 Spark plug laser presented at CLEO 2011

20.6. Q-switching

Q-switching produces less frequent, higher power pulses than mode locking. In this case the inversion slowly builds up until suddenly the Q of the cavity changes. This can be done passively through a saturable absorber such as a Cr:YAG crystal or carbon nanotubes. It can also be done actively using a fast optical switch (Pockels cell or Kerr cell).

20.7. Cavity Dumping

In a mode locked laser we can make both mirrors 100% reflecting, in which case the pulse power accumulates through multiple passes. Then by activating a fast switch the entire pulse can be dumped out of the cavity in one go.

20.8. Relaxation oscillations

In a fiber laser, the inversion builds up (and down) slowly, over a few milliseconds. During this time the poorly inverted erbium at the end of the fiber furthest from the pump can act like a saturable absorber causing the laser to pulse. These “relaxation oscillations” can produce power peaks a hundred or more times greater than the normal power.

20.9. Fiber Fuse

When we send 2 Watts or more light energy through an optical fiber it is easy to create a fiber fuse. This starts with a strong absorption of light at the light surface. This creates a plasma (the “blackest” material in the universe) that absorbs the light energy and travels back along the fiber towards the source and destroys the fiber core leaving a string of bubbles. A fuse can be generally be stopped with a fiber coupler which expands the beam out.

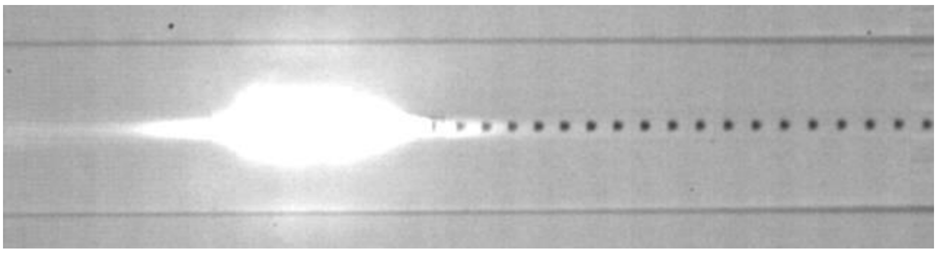
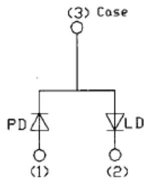


Figure 34 A fiber fuse propagating to the left. Note the characteristic string of bubbles where the core once was (from www.fiberfuse.info)

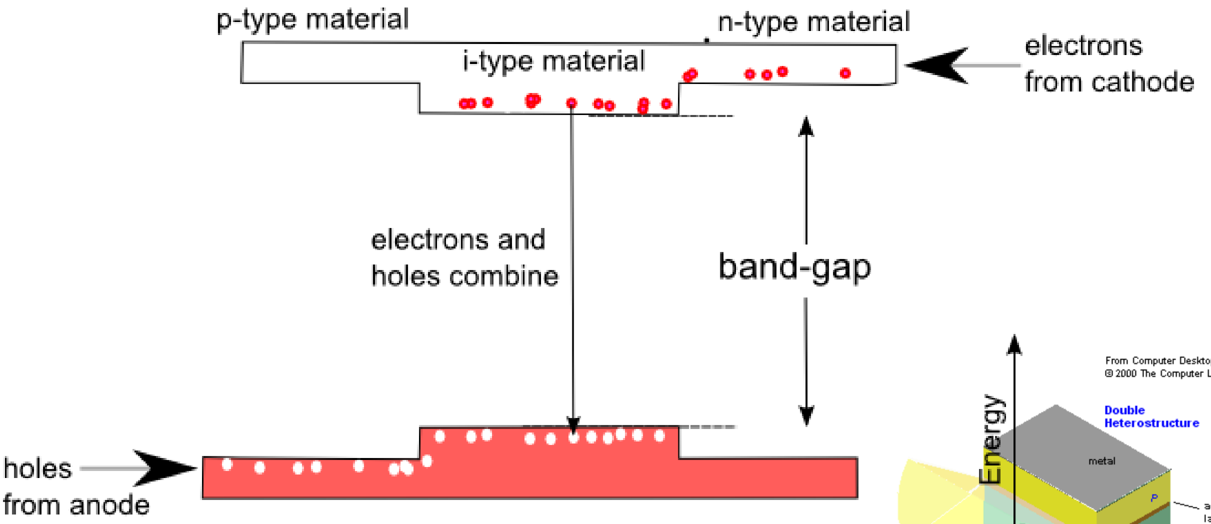
Chapter 21. Laser Diodes and Photodiodes: How they work and how to wire them up so they still work...

21.1. Laser diodes

The most common lasers are laser diodes. They are easy to use and easy to destroy. As we discussed last week, a laser diode consists of a PIN junction with two mirrors, often produced by simply cleaving the facets. Laser diodes usually have three pins. The laser we'll be using has its anode connected to the case (pin 3) and its cathode on pin 2. The third pin is reserved for the photodiode anode. The photodiode can be used to monitor the laser output power.

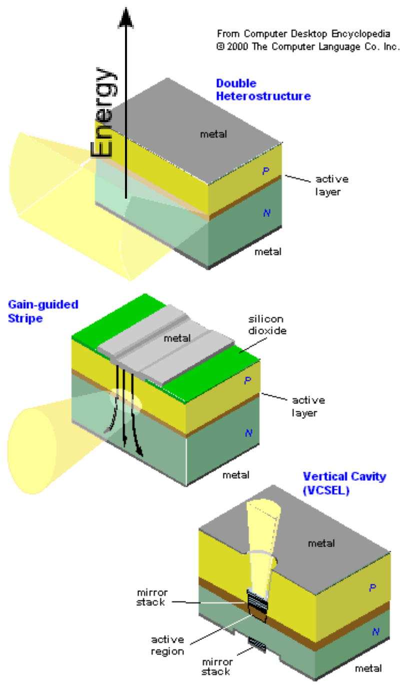


The inversion in a laser diode is created by injecting holes into the valance band from the anode, and electrons from the cathode. These get trapped in the undoped "i" region, where they can recombine through stimulated emission.



21.2. Types of laser diodes

The simplest laser diodes consist of a cleaved block of InGaAs with a PIN configuration. In order to get efficient lasing, it's usual to add a "gain-guiding" stripe to localise the current to a small area.



Often many of these lasers are made beside each other. Then a stack can be sandwiched to produce large (kW) output powers (with very poor beam quality).


Figure 35 Laser diode array (from LLNL document "Laser Programs, the first 25 years".)

All such lasers lase over several modes. To create a single wavelength laser we anti-reflection coat one facet and build in a grating. This type of laser is called a Distribute Feedback Laser (or a DFB laser diode). These are widely used in telecommunications.

Because all these lasers require cleaving the crystal, it’s impossible to make them as part of a larger optical device. This is one advantage of the VCSEL (vertical cavity surface emitting laser). The other advantage is that it produces a round beam, which is easy to focus into an optical fibre core.

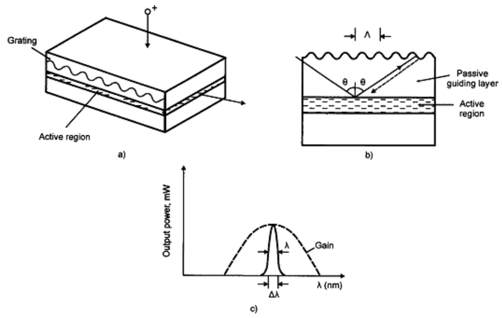


Figure 36 DFB laser

21.3. Driving a Laser Diode

Laser diodes are extremely vunerable to current spikes. A simple constant current driver is shown in figure 35. A constant current source is needed because otherwise the laser will take all the current it can (once you are over the band-gap voltage, the resistance is extremely low). The capasitor is present to smooth any current spikes. The two 10 Ohm resistors are there to stop the regulator allowing an uncontrolled current going to the laser diode. Note that the anode of the laser diode is connected to the positive and the cathode to the negative. You can get much more detail of the regulator here: http://lednique.com/power-supplies/lm317-constant-current-power-supply/ .

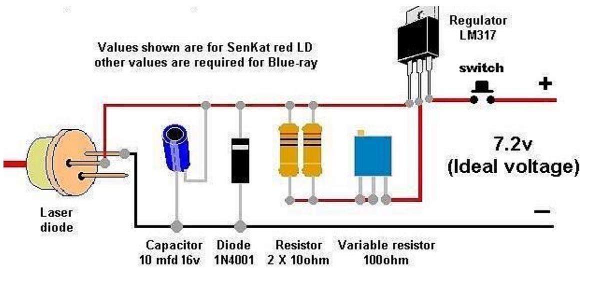


Figure 37 A simple constant current driver for a laser diode from http://www.rog8811.com/laserdriver.htm

21.4. Photodiodes

In contrast, photodiodes are reverse or unbias (or they die…).

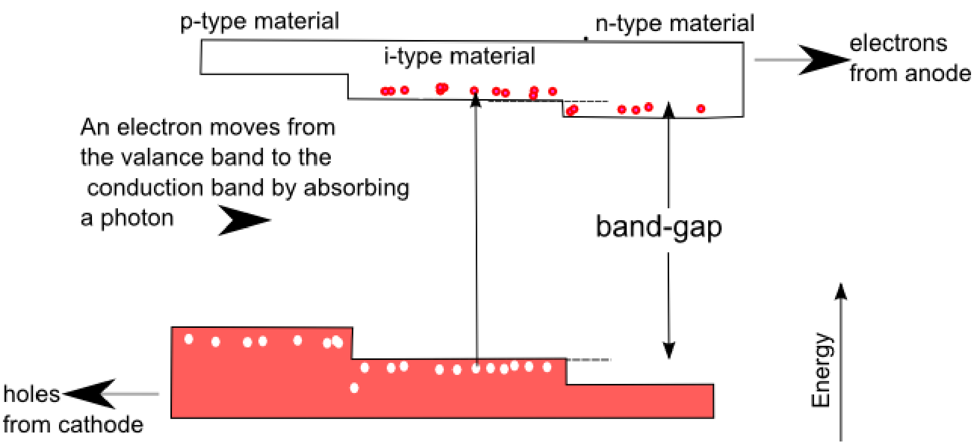
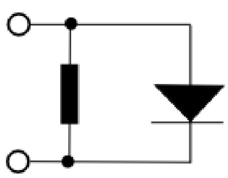
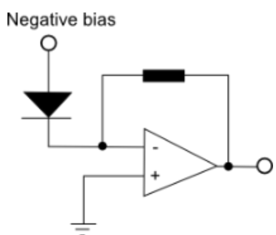


Figure 38 Photodiodes produce current through the absorption of a photon driving an electron from the valance band to the conduction band.

The circuit for an unbias photodiode is very simple. This is usually the lowest noise configuration for a photodiode. To increase the response speed we need to bias the photdiode, pulling the holes and conduction band electrons out of the “i” layer.





An example datasheet is shown in Figure 43. Note that the “dark current” is only specified when a -5V bias is applied.

GENERAL SPECIFICATION

Criteria	Specification			Remarks
Tap ratio (%)	2	5	10	
Fiber		8/12 ribbon fiber		Bend insensitive available
Number of Channels		~ 12 channels		Standard – 8, 12
Material (Lid/Substrate)		Si/Si		4 points contact
Insertion Loss (dB)	< 0.6	< 0.8	< 1.0	
Responsivity Range (mA/W)	> 15	> 30	> 50	
Input Power (dBm)	< 19	< 17	< 15	
Return Loss (dB)		> 50		@ angle polished
PDL (dB)		< 0.15		
TDL (dB)		< 0.15		
WDL (dB)		< 0.15		
Bandwidth (MHz)		> 500		
Dark Current (nA)		< 1.0		Typ. 0.3 nA @ –5 V, R.T.
Crosstalk (dB)		> 35		
Capacitance (pF)		< 3.0		Typ. 2.3 pF @ –5 V.
Core Pitch (μm)		250.0 ± 0.5		
Fiber Alignment Error (μm)		0.0 ± 0.5		
Polishing Angle (°)		0, 8, 12 (± 0.3)		Customisation available
Polishing Flatness		< 4 fringes @ 633nm		
Operating Temperature (°C)		0 ~ 70		
Storage Temperature (°C)		-40 ~ 85		
Dimensions (mm^3)				

Figure 40 A typical pigtailed photodiode specification (from Hantech)

21.5. Tunable Laser Diodes

When Finisar acquired Syntune, they became a major manufacturer of tunable diode lasers. There are various ways to build a tunable diode laser. The simplest is to anti-reflection coat one of the laser diode surfaces and install an external grating and back-reflect in Littrow (can you remember what this means?). Instrument tunable lasers from KeySight, Viavi, etc. use this technology.

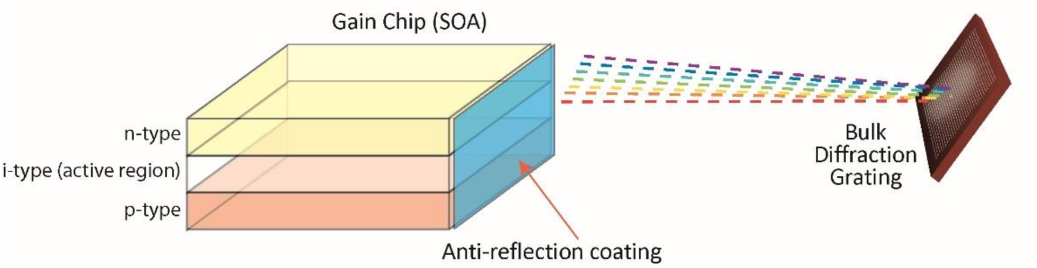


Figure 41 Laser diode with external tuning mechanism (used in Agilent tuneable lasers)

Another method, more commonly used in telecommunications, is the Y-junction design shown in Figure 40. This is often referred to as a Modulated Grating Y-branch laser (MGY). The gratings in the two arms are current tuned and match at the choosen wavelength giving constructive interference.

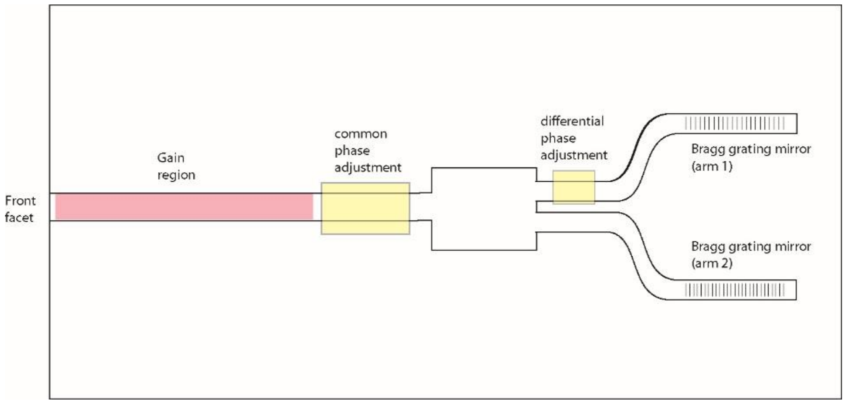
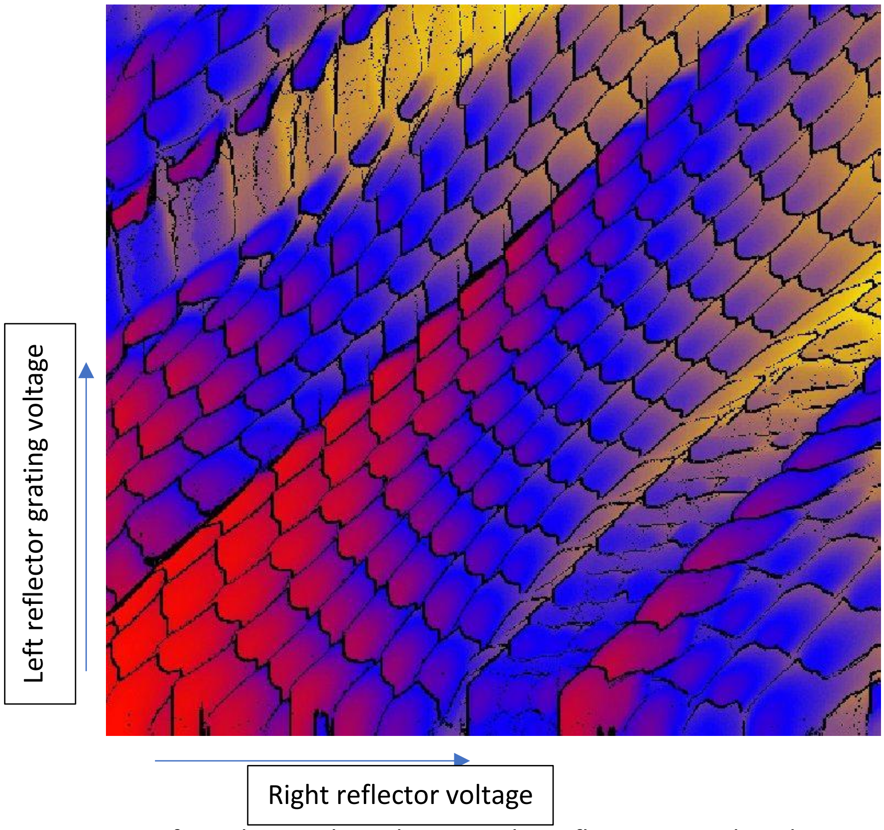
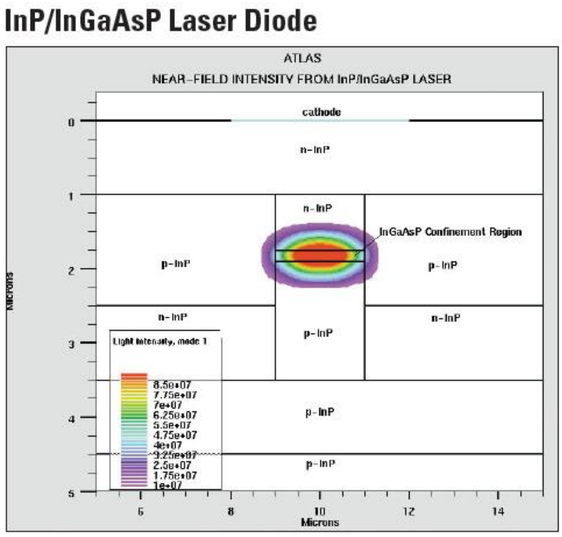


Figure 42 Y-junction laser diode (see Jan-Olof Wesström's paper at OFC 2004)

With careful design the differential phase adjustment can be dropped and an SOA (semi-conductor optical amplifier added after the front facet to boost the power.





If we change the voltages in the reflectors together then we simply add a wavelength to the cavity, shifting about 70GHz (moving up on the diagonals in the above diagram). If we change them by different amounts then we match up a different pair of modes on our "Vernier" scale and jump to a new "super mode"

Typical laser diode types include

InGaAsP – ~1060nm

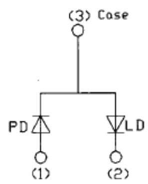
InGaAs – 980nm

InP 1400-2000nm

21.6. Questions

1. If a InP laser is 3mm long (n=3.4), what's the frequency separation of the modes of the laser?

2. Here's a laser diode. Sketch out a circuit that would let you run the laser and measure the light output:



3. How can a laser diode get damaged in a circuit?

4. How can a photodiode get damaged?

5. How can you increase the response speed of a photodiode?

Chapter 22. Week 22: Designing with lenses and mirrors

Many optical systems include several lenses, mirrors and other components. While software like ZeMax can be used to model such systems, it very much helps to understand how to do such modelling by hand before using complex software. ZeMax is great for optimising a good design, but won't do your design for you.

22.1. The Lens Makers formula and other equations

Consider a spherical mirror, as shown in Figure 43.

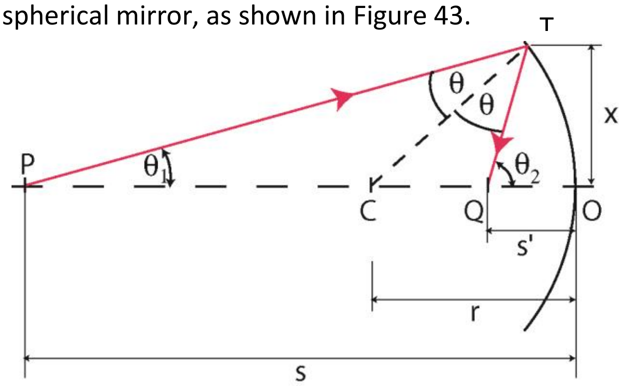


Figure 43 Optics of a spherical mirror

The "sine rule" tells us that, in a triangle, the ratio of the sine of an angle to the opposite side is constant. So using triangle PTC

$$\frac{\sin\theta_1}{r} = \frac{\sin\theta}{s-r}$$

Similarly using triangle CTQ (and the fact that $\sin(180-\theta_2) = \sin\theta_2$)

$$\frac{\sin\theta_2}{r} = \frac{\sin\theta}{r-s'}$$

So therefore

$$\sin\theta = \frac{(\mathrm{s}-\mathrm{r})\sin\theta_1}{r} = \frac{(r-s')\sin\theta_2}{r}$$

and

$$(\mathrm{s}-\mathrm{r})\sin\theta_1 = (r-s')\sin\theta_2$$

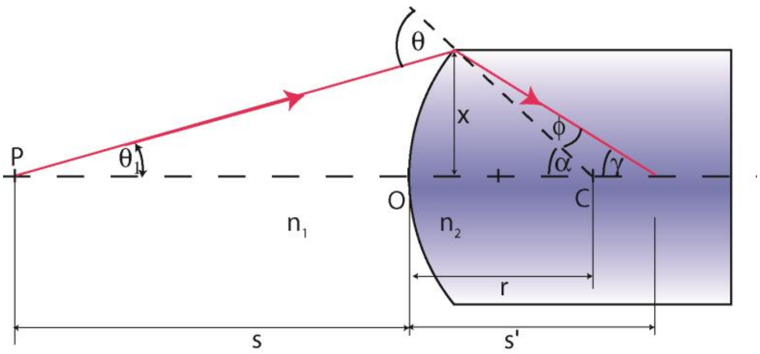
If the angles are small (the paraxial approximation), then $\sin\theta_1 \approx \theta_1 = \frac{x}{s}$ and $\sin\theta_2 \approx \theta_2 = \frac{x}{s'}$ and so, in this approximation, with a simple rearrangement

$$\frac{2}{r} = \frac{1}{s} + \frac{1}{s'}$$

Notice that a ray coming from infinity focuses at $\frac{r}{2}$. We call this the focal point and rewrite the equation

$$\frac{1}{f} = \frac{1}{s} + \frac{1}{s'}$$

We can do virtually the same derivative with a spherical surface



And derive the equivalent equation

$$\frac{n_2 - n_1}{r} = \frac{n_1}{s} + \frac{n_2}{s'}$$

Notice that we are always assuming that the distances are large ($\sin\theta \approx \theta$). In general rays don't all focus at the same point. This is called spherical aberration. When would you expect this to be worst?

For a spherical lens the rays hitting the edge of the lens focus first (the peripheral focus) and the rays from the center focus further away (the paraxial focus). By keeping the light rays close to orthogonal to the glass surfaces, spherical aberration can be minimised. An asherica lens can also be used to remove spherical aberration, but this generally only makes economic sense when the lens is molded or etched.

22.2. Single thin lens

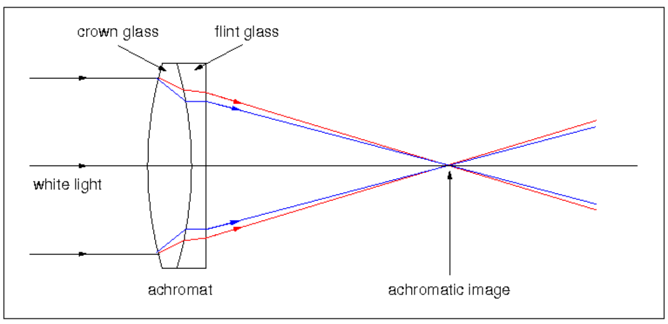
To model a single thin lens we need to do a near identical derivation twice to get

$$\frac{1}{s} + \frac{1}{s'} = (n-1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right) \qquad (22.1)$$

If we let $s \to \infty$ we can see that the inverse of the focal length is

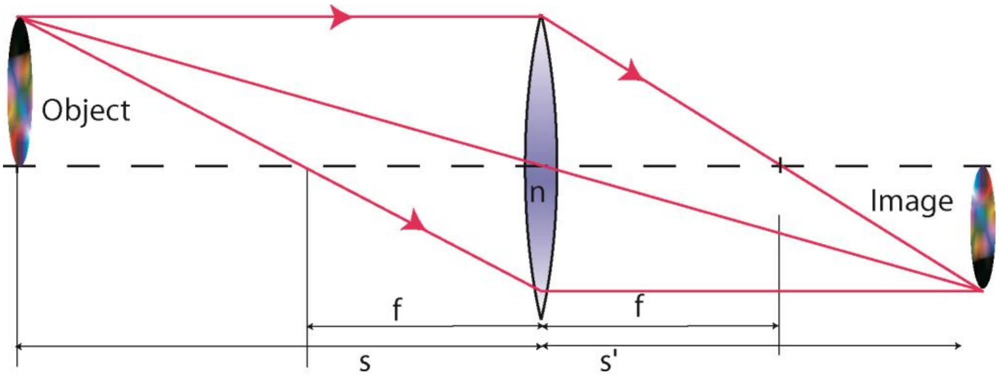
$$\frac{1}{f} = (n-1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right)$$

This is called "the lens makers formula". In real life, different wavelengths focus at different distances- this is called chromatic aberration. This can be partly overcome by using an "achromatic doublet"; two lens combined with opposite wavelength dependences. Usually the main lens is a "crown glass" lens. A weak negative "flint glass" lens is used to reverse the spherical aberration.

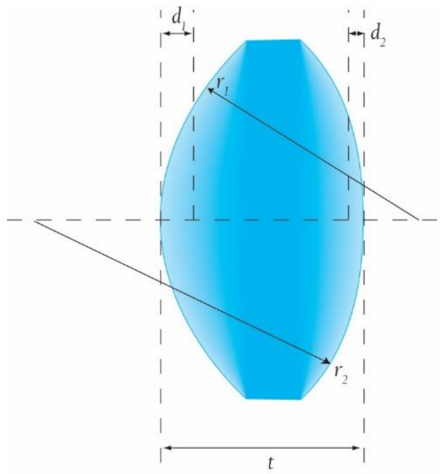


We can also work out graphically where an image will form using three simple rules

1. Rays through the centre of a lens are unaffected by the lens.
2. Rays parallel with the axis will go through the focus.
3. Rays through the focus will become parallel with the axis.



Of course, real lens aren't always thin.



The focal length of a thick lens is given by

$$\frac{1}{f} = (n-1)\left[\frac{1}{r_1} + \frac{1}{r_2} - \frac{(n-1)^2 t}{n r_1 r_2}\right] \qquad (22.2)$$

Where r_1 is the radius of the first face, r_2 is the radius of the second face, *t* is the thickness of the lens at the thickest point and *n* the refractive index of the glass. *"f"* is the distance from the focus to center of the lens.

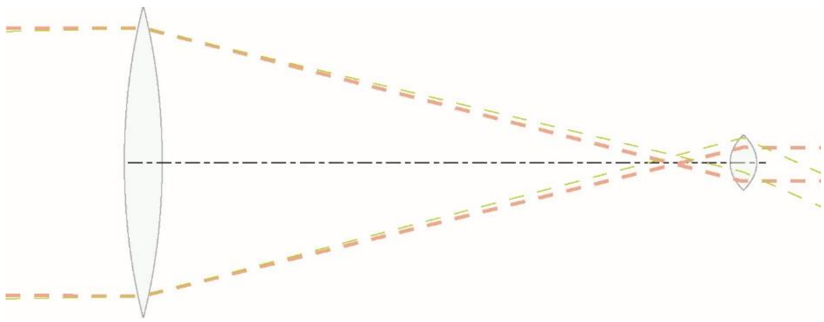
However to use equation 22.1 we need to measure s_1 and s_2 from the principal plans. The distance from the surface to the "principal plane" (inside the lens) is

$$d_1 = ft\left[\frac{1-n}{r_2}\right] \quad (22.3)$$

$$d_2 = ft\left[\frac{1-n}{r_1}\right] \quad (22.4)$$

22.3. Telescope

There are hundreds of useful optical designs, but one that keeps being useful is the telescope, or beam expander. The idea here is simple. We take two lens with different focal lengths and arrange them so their foci are at the same point. The combination acts as a beam expander. In a telescope, the lens with the short focal length is referred to as the eye piece. The lens with the long focal length as called the objective.



The magnification is simply the ratio of the focal length of the objective lens (the one nearest the object) to the eye piece. If you're designing a telescope then you will want the size of the exit beam at the eye piece to be between 2 and 6mm (that's about as big as your iris can get). A typical focal length for a telescope eyepiece is 12mm. At the extemes a 30mm eyepiece will give you a huge field of view but limited magnification, and a 3mm eyepiece will give huge magnification but you'll need a very steady telescope or the vibrations will be similar in size to the field of view. As a rule-of-thumb it's hard to get better than 60x magnification per inch of objective size and almost impossible to do better than 300x in air because of fluctuations in air density.

We often use a telescope as a beam expander, the beam expansion ratio is also the ratio of the focal lengths.

22.4. Matrix form

We can trace the beam propagation through a system by tracking the beams angle and its position relative to the axis (for a spherically symmetric system).

We can represent this by the vector $\begin{bmatrix}\rho\\ \theta\end{bmatrix}$ where ρ is the distance from the optical axis and θ is the angle the beam is travelling.

Each optical element is then represented by a matrix.

A beam travelling distance d is $\begin{bmatrix}1 & d\\ 0 & 1\end{bmatrix}$. A plane dielectric interface is $\begin{bmatrix}1 & 0\\ 0 & {n_1}/{n_2}\end{bmatrix}$. A thin lens is is $\begin{bmatrix}1 & 0\\ -{1}/{f} & 1\end{bmatrix}$. To model an optical system we can simply multiply these together.

$$M = M_n.M_{n-1}\dots M_o$$

Ray tracing programs generalise this method to include both axes, and trace many rays, but the principle remains the same.

We're going to try a simple software version of this, which you can try at http://demonstrations.wolfram.com/ConstructingASimpleOpticalSystem/ .

This has been a very superficial introduction to lens design. If you need to know more then I recommend "Introduction to Lens Design" by Joseph M Geary. If you need to borrow my copy, let me know!

22.5. Key Terms

Spherical aberration – For a convex spherical lens the light further from the axis of the lens will have a shorter focal length

Coma- Distortion caused by light coming in at an angle to the lense or parabolic mirror

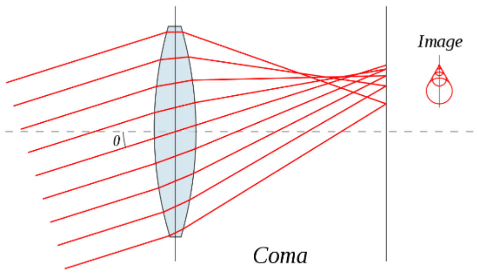
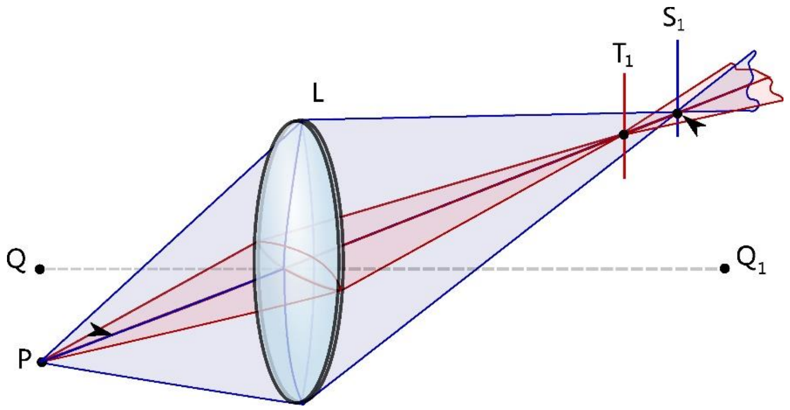


Figure 44 Coma - from By http://upload.wikimedia.org/wikipedia/en/3/31/Lens-coma.svg, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=996267

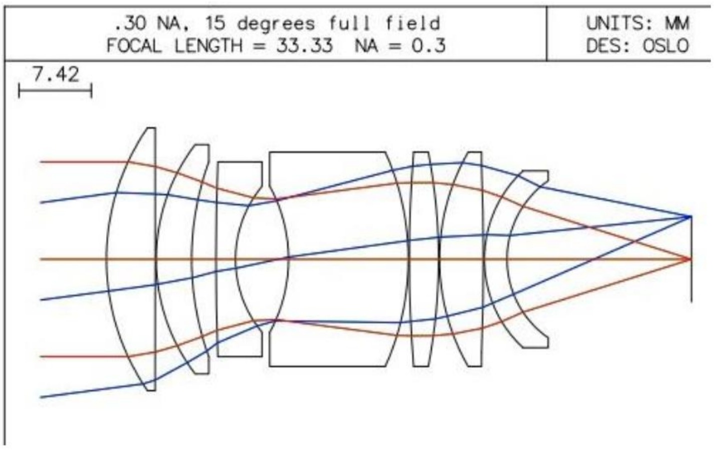
Astigmatism – light in different planes has different focal lengths



22.6. Key Principles of Optical Design

1. Keep the angle between the surface and the light as close to 90° as possible on all lens surfaces
   (the exception being where you are using an opposite spherical aberration to compensate)
2. Minimise spherical aberration by using multiple lenses
3. Use stops to minimize aberations
4. If two sides of a lens are nearly the same – make them the same (because they will be put in backwards)
5. Big mirrors are cheaper than big lenses – and parabolic mirros have no spherical or chromatic abberation (but do have coma and astigmatism)
6. Use standard glasses and standard lenses/mirrors wherever you can (Edmund, Thorlabs are good starting points). Custom lens are expensive!
7. Microlens arrays can be etched or molded and are almost always aspherical

Let's look at an optimised lens design for an eyepeice from Dave Shafer's excellent notes at https://wp.optics.arizona.edu/jsasian/wp-content/uploads/sites/33/2017/07/Design-methods-David-Shafer.pdf .



What principles can you see that he's employed?

For next week, if you have a laptop, please download I:\Transfer\IanC\WinLens3DBasic_Installer.exe so we can do some lens design. You can see a manual for this software here: http://atom.lylver.org/AstroSurf/PDF/Manual_WL3D_Intro.pdf

22.7. Problems

1. Using equation 22.2 work out the focal length of a glass sphere refractive index n.

2.
   a. An object is 10cm from one side of a 10cm diameter glass sphere with refractive index 1.5. Using equation 22.3 and the focal length, figure out how far from the surface of the opposite side the image will focus?

   b. What would happen if we coated the back face with a mirror surface?

3. Play with the computer model and see how many, and how closely you can get the beams to focus in the corner in under 2 minutes.

4.
   a. If I have a 1 meter focal length objective on a telescope, what focal length do I need on the objective to get 50x magnification?

   b. If the pupil on my eye is 3mm across, how big should I make my objective so it fills my field of view?

   c. How much more light am I gathering from the objects I'm looking at?

Chapter 23. Week 23 Liquid Crystals and beam steering

Liquid crystals, first observed in 1888, are liquids in which the molecules retain some order. In the smectic phase they sit in regular layers. In the nematic phase, the regular laying is lost, but the molecules still have a preferred direction (the "director"). Heating a liquid crystal in the nematic phase beyond the "clearing temperature" will result in a further transition to a totally random (isotropic) liquid. In the remainder of this week's talk, we will only consider the nematic phase. Commercial liquid crystal materials, designed to operate in the nematic phase usually consist of a mixture of several different molecules.

Typical liquid crystals are organic molecules with a rod-like structure. An example of such a liquid crystal molecule is drawn in Figure 47. While such molecules have little or no permanent electric dipole, an electric field will produce a strong induced electric dipole, aligning the molecule with the direction of the electric field.

Figure 45 MBBA molecule (4-methoxylbenzylidene-4-butylaniline), perhaps the most studied liquid crystal molecule, first discovered by Hans Kelker1

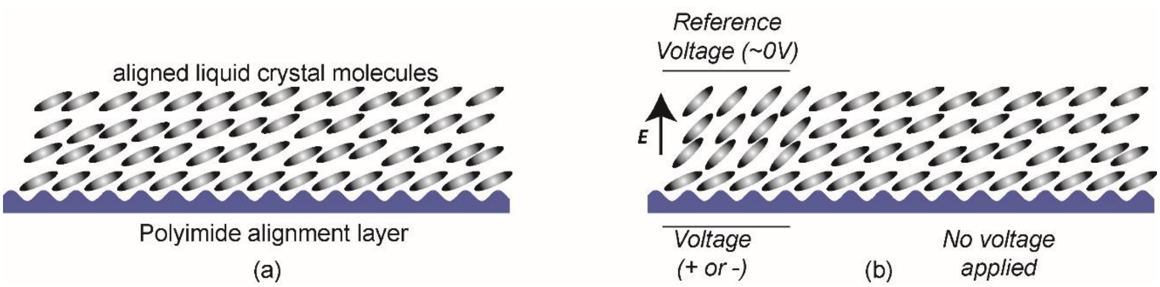


Figure 46 (a) Nematic liquid crystal with a polyimide alignment layer. (b) The addition of a voltage across the liquid crystal causes the molecules to rotate, increasingly aligning themselves with the direction of the electric field as the field increases.

23.1. Untwisted Nematic Liquid Crystal Cells (Fréedericksz cells)

Molecules at the edge of a structure can be anchored using a thin layer of shaped polyimide. The polyimide is spin coated onto the glass, baked and then buffed in the desired direction, shaping the surface (Chatelain's method[2]). This alignment layer holds the molecules in a particular (non-vertical) direction, as sketched in Figure 2a. The first layer of liquid crystal molecules typically sit at an angle of about 3° to the horizontal when no field is present and the subsequent layers conform to this alignment unless an electric field is applied. When an electric field is applied across the cell, the molecules rotate to be partially aligned with the applied field; the higher the field, the closer the alignment. A typical cell used for visible light will operate over approximately a 2V range, while a cell designed for 1550 nm light will require approximately twice the voltage range and twice the cell thickness. Once the applied electric field exceeds 10^6V/m, the molecules are nearly fully aligned with

[1] H. Kelker, B. Scheurle and B. Scheurle H. Kelker, "A Liquid-crystalline (Nematic) Phase with a Particularly Low Solidification Point," *Angew. Chem. Int. Ed.* 8, no. 11 (1969): 884.
[2] Par Pieére Chatelain, "Sur L'Orientation Des Cristaux Liquides Par Les Surfaces Frottees," *Bulletin De La Société* 66 (1944): 105–130.

applied field and there is little further change in the birefringence. In a Fréedericksz[3] cell the director is essentially uniform for the molecules, independent of their distance from the base of the cell. This is in contrast to the twisted nematic cells used in most liquid crystal displays.

23.2. Ion Migration

In practice, an applied voltage will gradually be cancelled by the migration of free ions in the solution. To counteract this, the amplitude of the applied voltage is held constant but the sign of the field is switched at a rate much faster than the time taken for the molecules to realign (typically at 1 kHz or more). The orientation of the liquid crystal molecule stays essentially constant with only the induce electrical dipole swapping direction. However, if the electric field amplitude in the two directions is not equal, either because the applied voltages are different or because of a non-symmetric ion buildup in the alignment layer, the LCoS image will flicker.

23.3. Cell construction

A sketch of a Liquid Crystal on Silicon (LCoS) Cell is shown in Figure 47. The CMoS chip controls the voltage on plates buried just below the chip surface, each controlling one pixel. An XVGA chip will have 1024x768 plates, each with an independently addressable voltage.

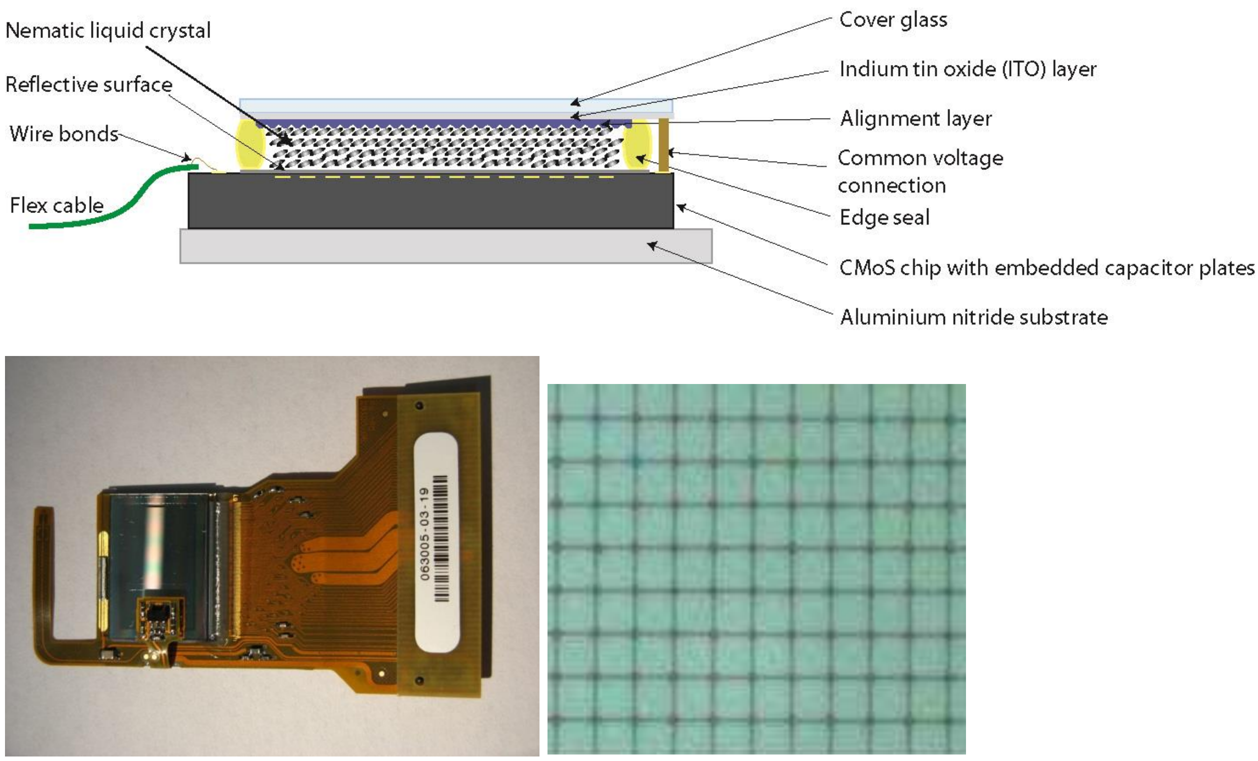


Figure 47 A conceptual drawing of a Liquid Crystal on Silicon cell with the vertical dimension exaggerated to reveal the different layers. Typical cells are about 3 centimeters square and about 2 mm thick, as shown in the photo. The reflective surface itself is pixelated as shown in the microscope image.

A common voltage for all the pixels is supplied by an indium tin oxide layer on the cover glass. This layer should be optimized for the wavelength range being controlled. Layers optimized for visible light are often absorbing in the near infra-red. Not shown in figure 49 are the spacer balls or glass rod used to ensure that a uniform thickness of liquid crystal is maintained across the cell.

[3] V Fréedericksz and V Zolina, "Forces Causing the Orientation of an Anisotropic Liquid," *Trans. Faraday Soc.* 29 (1933): 919–930.

23.4. Driving a cell

The variable voltage needed to control each pixel may be applied via an analog ramp voltage which is disconnected from a pixel at a time corresponding to the desired voltage level. This voltage source must be extremely stiff because the load on the voltage source will change as pixels reach their desired charge level and are disconnected from the supply. A lack of stiffness in the source will result in the voltage jumping as the load is removed, increasing the voltage applied to cells still connected to the source.

Alternatively, the variable voltage may be simulated using pulse-width modulation. This has the advantage of being purely digital, but the major disadvantage of introducing an extra source of flicker in the image. This causes problems in telecommunications applications but can be overcome by using liquid crystals with very slow response times.

23.5. Phase and Spatial Light Modulators

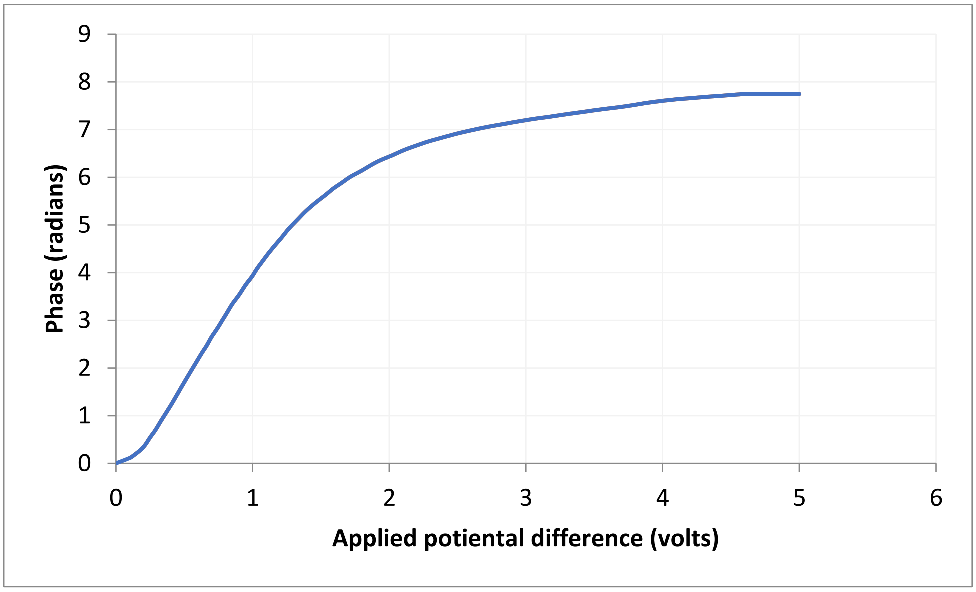


Figure 48 A typical applied voltage vs. phase shift in a Fréedericksz liquid crystal cell

In a Fréedericksz cell, the polarization orthogonal to the director is minimally affected by any voltage applied to the cell (the "ordinary" axis of the liquid crystal). However, the alignment of the liquid crystal strongly affects the phase velocity of the extraordinary polarization of any light wave passing through. This can be understood intuitively, since the molecule is polarisable along its length and so interacts strongly with any electric field with a component in that direction. If the molecule is parallel to the electric field associated with a light wave (hence perpendicular to both the light-wave direction of travel and the associated magnetic field) it interacts strongly, producing a high refractive index and corresponding phase shift, as shown in Figure 50. We can consider the LCOS as a phase (or amplitude if used with cross polarisers) matrix where pixel- by-pixel addressing of the local phase is possible. In general it is desirable to have at least 360 degrees of phase to allow phase resetting (360 degree = 0 degrees) in creating a desired phase map, but making the liquid crystal cell thicker to create more phase can have detrimental performance in terms of speed of operation and also the sharpness of the individual pixels which can affect efficiency and light scattering at the reset points.

This phase control, as sketched in figure 51, allows us to set up a phase ramp. Through this ramp we can steer an optical beam; equivalent to an optical wedge. More complex beam steering allows full head-up displays, such as those being commericalised by Holoeye (see figure 52). What would a lens look like?

.

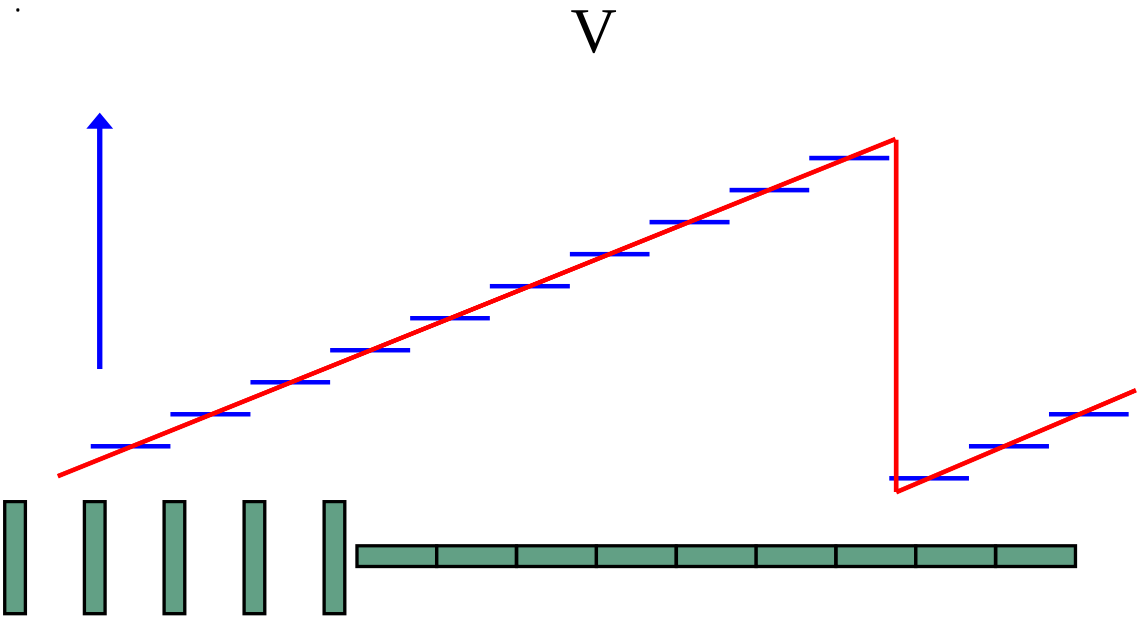


Figure 49 A phase ramp cam be built by applying gradually increasing voltages to the pixels. Resets normally occur at 2π. The spatial period determines the direction of the beam (the orders of the virtual blazed grating or Fresnel mirror.

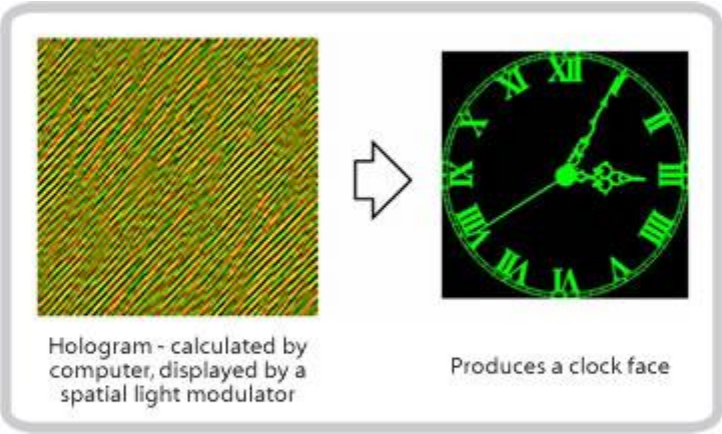


Figure 50 head-up display using LCoS (Joel Carpenter http://www-g.eng.cam.ac.uk/CMMPE/displays3d.ht