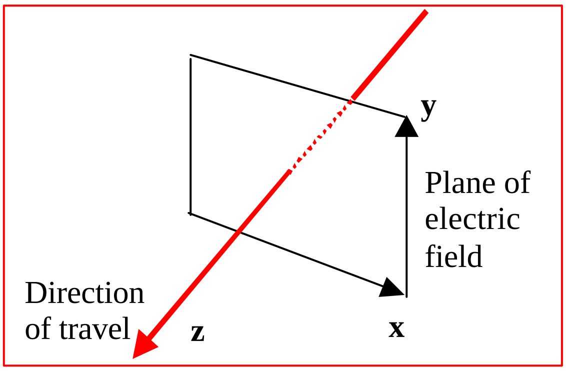
Polarisation

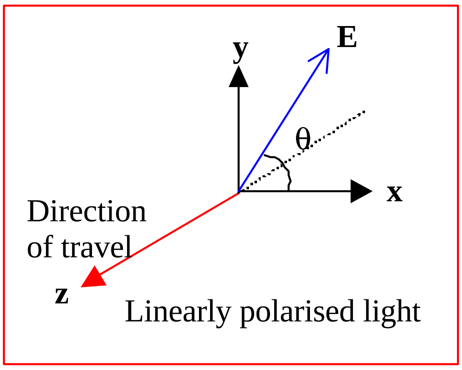
The amount of light reflected and transmitted at an interface between two given materials depends on the angle of incidence and, to a lesser extent, on the frequency of the light. However, those investigating the phenomena at the beginning of the 19th century became rather puzzled. They obtained inconsistent results that suggested there was another, hidden, property of light involved. This property seemed to produce no sensible effect on the human eye. What emerged from this work was the property called the polarisation of light [spelt 'polarization' in Hecht's textbook and other American works]. Early investigators still associated with polarisation phenomena are the Frenchman Etienne-Louis Malus and the Scottish natural philosopher David Brewster. Today, there are a myriad of widespread applications for polarised light. These include simple polarising sunglasses and glasses that allow you to watch a 3D film in the cinema, though well established techniques to make visible the otherwise invisible in the polarising microscope, to the exploitation of polarisation in LCD displays and video projectors.

When you know where to look for them, polarisation phenomena can be found in nature and in the 'daily world'. The natural polarisation of the blue sky has led to some animals acquiring polarisation sensitive vision. Measuring the polarisation of light from sunspots tells us the strength of the magnetic field associated with these spots. The natural polarisation of reflected light has led to the development of Polaroid sun-glasses and camera filters that can be used to reduce the glare from reflections or, on occasions, enhance reflections (for example to improve the contrast when viewing a rainbow). Polarisation encoding has been used to transmit 3D colour pictures. Examining rocks, minerals and other materials through the polarizing microscope can tell us a great deal more about our specimens, as we'll see later in this chapter. In the laboratory, in daily life and in nature, a great deal is made of the polarisation of light.

Four words are used to describe different kinds of polarisation – **linear**, **circular**, **elliptical**, and **unpolarised**. Circular can be considered a special case of elliptical, and so can linear polarisation. Elliptical is therefore the most general case.



Polarisation is all to do with the direction of the electric field associated with light. We shall follow *Hecht's* description in chapter 8. The existence of polarisation phenomena is a direct consequence of **light being a transverse wave, or nearly so**. In isotropic media, light is purely transverse and our sketches will assume that the direction of the electric field **E** is perpendicular to the direction of propagation, denoted **z** in this chapter, i.e. the electric field is always in the **xy** plane. Light propagating through some solids may not be purely transverse, as we shall see.



What phenomena do you associate with polarisation?

Können's book "Polarized Light in Nature", referenced in the course book list, covers many examples.

Linear polarisation

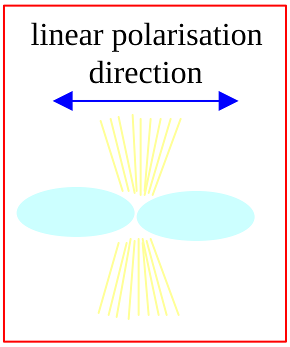
In discussing and explaining polarization phenomena, you are looking all the time at what is happening to the electric field of light. **With linear polarisation, the direction of the electric field E stays constant at a point in space.** That is, the direction doesn't vary with time, though **E** itself varies sinusoidally with time. Let **z** label the direction of travel of the light and **x** and **y** be directions in the plane of the electric field. The electric field makes a constant angle θ to the **x** direction, as in the illustration. Remember that **E** varies in time with the frequency of the light wave. If the direction of **E** is constant in space, then the variations of **E** must be in-step along both **x** and **y** directions. Thus, mathematically:

$$\begin{aligned} E_x(z,t) &= E_{ox}\cos(kz-\omega t) \\ E_y(z,t) &= E_{oy}\cos(kz-\omega t) \end{aligned} \quad .$$

For the mathematicians, and Hons physicists in the class, the complete electric field is therefore given by

$$\begin{aligned} \boldsymbol{E} &= \mathrm{E_x}(z,t)\,\mathbf{i} + \mathrm{E_y}(z,t)\,\mathbf{j}\text{ , where }\mathbf{i}\text{ and }\mathbf{j}\text{ are unit vectors along }\mathbf{x}\text{ and }\mathbf{y} \\ &= \left(\mathrm{E}_{ox}^2 + \mathrm{E}_{oy}^2\right)^{\frac{1}{2}}\cos(kz-\omega t)\text{ in the direction of the blue arrow marked }\mathbf{E}\text{ .} \end{aligned}$$

The key point is that the phases of the x and y components not only change in step but are always equal, though their amplitudes may be different. This doesn't happen with unpolarised light.



Can you tell if light is linearly polarised with your eye? If light is completely linearly polarised, then some people can tell through the phenomena of Haidinger's brush. Light transmitted through a sheet of polaroid is completely linearly polarised and, given a sheet, I can tell you in which direction it is polarised. Some of you can, too. Rotate the polaroid against a bland, uniform, bright background and look for a small, pale yellow, double brush in the centre of your field of view and, at right angles, a pale blue, filled-in, figure of 8 or just a faint, blue cloud, the whole rotating with the polaroid. If this sounds the product of a hallucination, see the adjacent figure in colour. Better, look through some polaroid. (*Demo*).

Relation between electric field and intensity (irradiance) of light

A light meter, such as is built into most cameras these days, records the intensity of incident light (more properly called irradiance, as those who have taken the general course *Tools for Science* might remember). Our eyes open and close according to decreasing or increasing irradiance of light. The irradiance is proportional to the energy contained in the light. How is this irradiance, denoted by the letter I, related to the electric field? Quite simply, the irradiance is proportional the average square of the electric field:

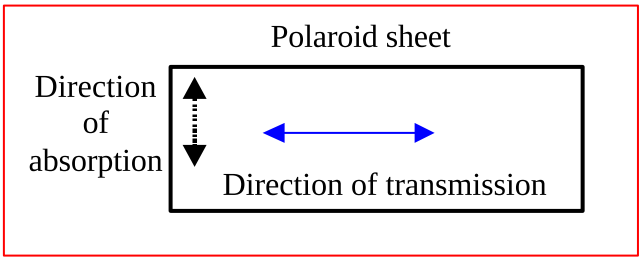
$$I \propto \langle E^2 \rangle$$

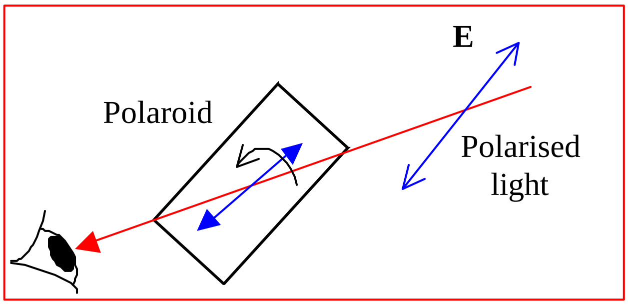
is how the mathematicians write it, where <> stands for "the average value of".

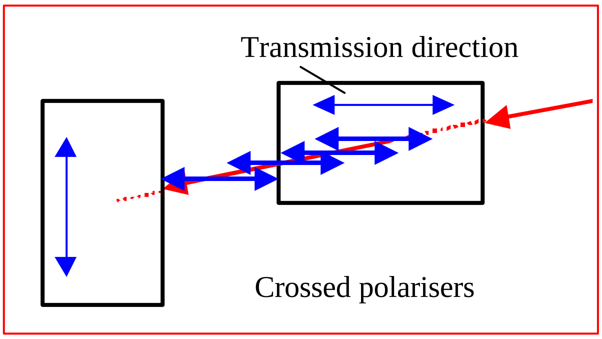
This is relevant to understanding polarisation phenomena because polarisation is about the direction and amplitude of the electric field, whereas what is seen by us or recorded on film in the camera is the irradiance of light. The expression above shows the relationship between the irradiance and the electric field.

Polaroid sheet

Polaroid sheet works by absorbing light that is polarised in one direction and transmitting perpendicularly polarised light. Old fashioned sheet exhibited an extreme example of *dichroism*, the name for a material that differently absorbs polarisations in different directions.





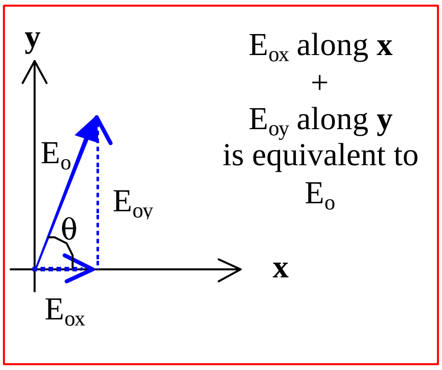


If you are looking to see if light is already linearly polarised, then a piece of polaroid rotated in front of the light will transmit all the electric vector that coincides with the direction of transmission of the polaroid. We'll say a bit more later about how the light intensity varies as the polaroid is rotated.

You will get no transmitted light when the polaroid is turned so that its transmission axis is 90° to the electric vector of the light. The condition in which two polarisers one behind the other are at right angles is called **crossed polarisers**. Crossed polarisers play an important rôle in many applications of polarised light. More later, also.

*A note on components of **E***

E, the electric field, has a direction and a size. It is a **vector**, like a displacement. Students of first-term, first year physics will be very familiar with displacement vectors. Imagine coming into a room through the door and walking to a spot near the centre of the room. The displacement from the door is exactly equivalent to a

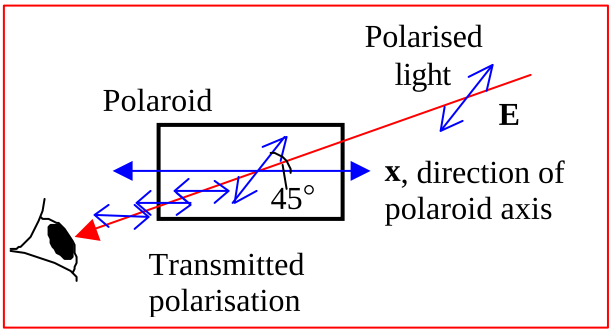


certain distance along the side wall plus another distance along an end wall. The distances along the walls are the **components of the displacement**. Likewise on an Ordnance Survey map, the location of a point is given by the *easting* and the *northing*. The easting is the easterly component whose coordinate is given along the bottom of the map and the northing is the northerly component, whose coordinate is given up the map.

The size of the components is just determined by the angle of the vector (displacement, electric field or whatever) to the **x** axis. This is so from the very definitions of the functions sine and cosine as the ratio of sides of a right angled triangle. See the figure. The result for the electric vector of magnitude E_0 is:

$$E_{ox} = E_o \cos(\theta)$$
$$E_{oy} = E_o \sin(\theta) .$$

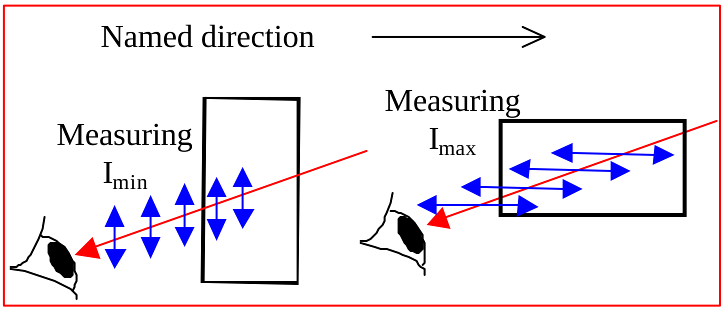
The subtlety of polaroid is that it lets through all the components in the direction of its transmission axis. For example, if E_o is inclined at 45° to the transmission axis, then application of the expressions above shows that E_o has equal components $E_o/\sqrt{2}$ along the **x** and **y** directions, since cos45° = sin45° = $1/\sqrt{2}$. A following sheet of polaroid with its transmission axis along the **x** direction will let through an electric vector of size $E_o/\sqrt{2}$.



% polarisation

For light that is partially polarised, it is useful to define the % of polarisation. Rotating a polaroid in partially polarised light produces a maximum intensity (call it I_{max}) and a minimum intensity with the polaroid at right angles (call it I_{min}). Then

$$\%\ polarisation = \frac{(I_{max} - I_{min})}{(I_{max} + I_{min})} \times 100$$



For example:

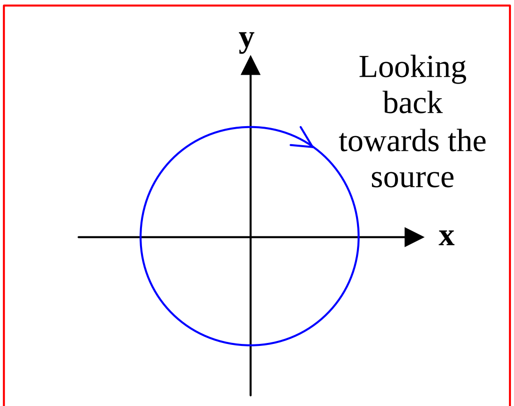
a) No change on rotating the polaroid gives $I_{max} = I_{min}$ and the % polarisation is zero.
b) If $I_{max} = 2I_{min}$, then % polarisation = 100/3 = 33%.

Circular polarisation

With circular polarisation, the x and y amplitudes of E are both equal but there is a phase difference of p /2 between them.

E.g.

$$E_x = E_o \cos(kz - \omega t)$$
$$E_y = E_o \sin(kz - \omega t)$$.



Right circularly polarised light

This pair of equations defines *right circularly polarised light*, whose rotation is *clockwise when looking back towards the source*. If you point the thumb of your right hand towards the source, then your fingers curl in the same direction as the **E** vector rotates. Look back at the previous pair of equations. Choose a time such that (kz-ωt) is zero. In this case **E** lies along **x.** At slightly later times, the **y** component is negative and the **x** component decreases and hence you can see that **E** rotates as shown. This is a description in terms of the time variation of **E.**

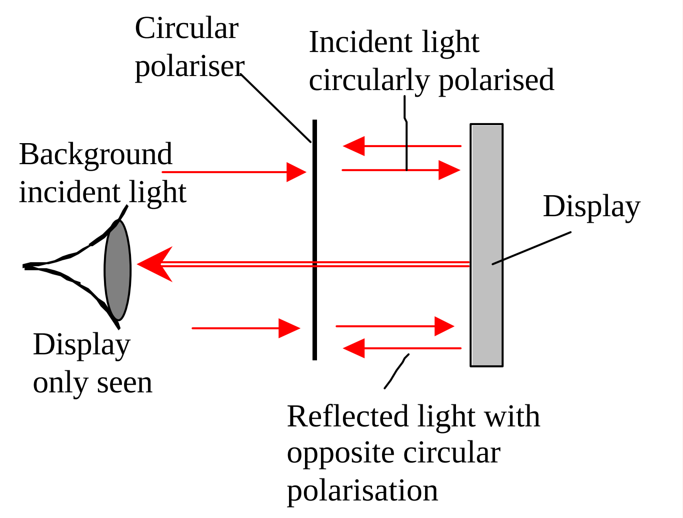
An alternative picture in terms of the space variation of **E** is that if you were to look at the electric field at a fixed time along the direction of propagation (**z**), **E** would curl around like a corkscrew. In right-circularly polarised light, it curls around like an ordinary right-handed corkscrew. I find I need a cork-screw beside me to imagine this properly.

Analogously, *left circularly polarised light* is defined by

$$E_x = E_o \cos(kz - \omega t)$$
$$E_y = -E_o \sin(kz - \omega t)$$.

In one sense, circular polarisation is the most basic polarisation because a single photon can be considered circularly polarised.

A sheet of polaroid on its own won't tell you if light is circularly polarised. Why is this? The average value of the electric vector is the same in every direction in the plane of the sketch above and hence there is no difference in the amount of light getting through a sheet of polaroid as it is rotated.



We'll see near the end of this chapter how to make a circular polariser. Circular polarisers are used to enhance the visibility of digital displays. They can cut out the extraneous light reflected from the face of the display, improving the contrast of the display. Here's how they work. They use the fact that mirror reflection changes the handedness of the polarisation of the light. Right circular polarisation becomes left circular polarisation upon reflection. Place a sheet of circular polariser in front of a digital display. External light that falls on the polariser becomes circularly polarised before it reaches the front of the display. The reflected light has the handedness of its polarisation reversed, and its return path is blocked by the polariser. Light generated by the display passes through the polariser and hence is seen without the background reflecting light. See the accompanying figure. (*Demo*).

A final point about circular polarisation, which we'll see is relevant later on, is this. If you combine together right and left-handed circular polarisation, you get linear polarisation. Take the equations above, for example. Adding together RH and LH circular polarisation (of the same amplitude) gives:

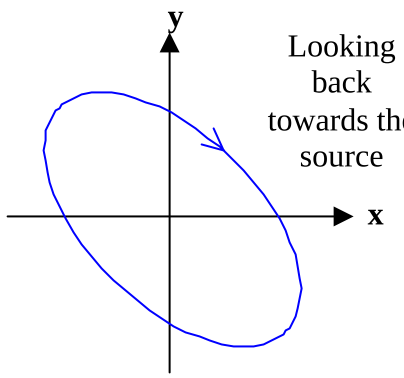
$$\begin{aligned} E_x &= E_o\cos(kz-\omega t)+E_o\cos(kz-\omega t)=2E_o\cos(kz-\omega t) \\ E_y &= E_o\sin(kz-\omega t)-E_o\sin(kz-\omega t)=0 \quad . \end{aligned}$$

i.e. the result is a vibration only in the **x** direction and hence is linearly polarised in this direction. If the phase of the two circular polarisations were different from the example above, the result would be inclined to the x axis.

Elliptical polarisation

With elliptical polarisation, the amplitudes of x and y components are generally not equal and neither are phases between the components anything special.

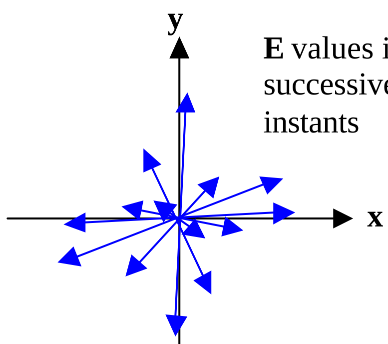
$$\begin{aligned} E_x &= E_{ox}\cos(kz-\omega t) \\ E_y &= E_{oy}\cos(kz-\omega t+\varepsilon) \end{aligned} \quad .$$



Elliptically polarised light

You can see why special cases give circular polarisation ($\varepsilon = \pm\pi/2$ and equal amplitudes) and linear polarisation ($\varepsilon = 0, \pm\pi, 2\pi$, etc). *Hecht* shows what happens when the amplitudes and phase shift change. The orientation of the ellipse depends on the ratio E_{0x}/E_{0y} and the eccentricity of the ellipse on ε. Elliptical polarisation comes into our story when we look at the phenomenon of birefringence.

Unpolarised light



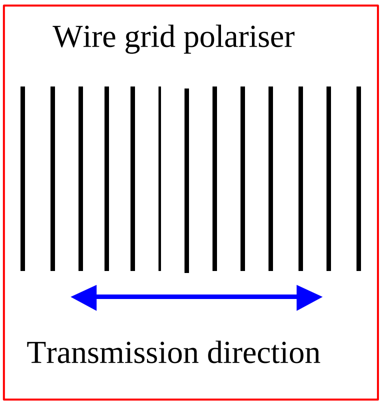
Unpolarised light

Unpolarised light consists of light where the direction of **E** varies at random between successive measurements at one point. Any direction is equally likely. Unpolarised light can be considered as a combination of equal amounts of linear polarisation in two directions at right angles, where the **two components are incoherent**, i.e. have a randomly changing phase difference.

If you place a piece of polaroid in front of unpolarised light, the result is linearly polarised light of **half the intensity** of the unpolarised light. You will soon see why.

Polarisers

The next sections are concerned about linearly polarised light, which is relevant in at least five different sets of circumstances. As a result, linear polarisation can be produced by at least five means:

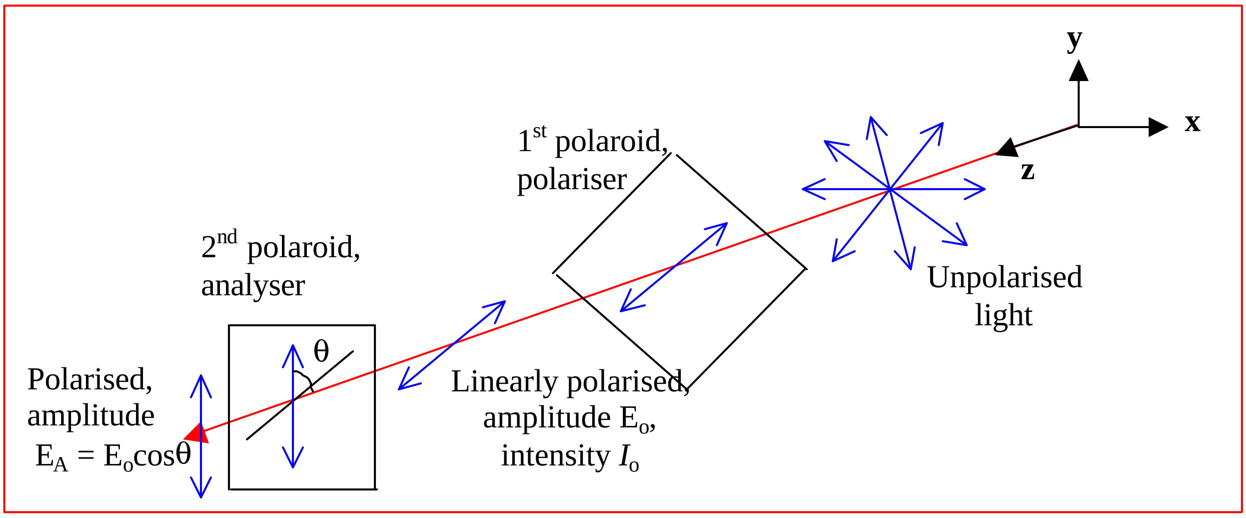


- Polaroid sheet
- Wire grid (*Hecht 8.3*)
- Scattering (*Hecht 8.5*)
- Reflection and multiple reflection (*Hecht 8.6*)
- Double refraction (*Hecht 8.4*)

Each of these mechanisms has its interest and applications. We'll pass quickly over wire grid polarisers, except to say that the spacing between the wires should be $<\lambda/4$, which means that they are most usually used for electromagnetic radiation of longer wavelength than light. They work on the principle that the incoming radiation forces the electrons in the wires to move lengthwise. An electron moving lengthwise re-radiates an electric field parallel to the wires in the forward direction that is exactly out-of-phase with the incident wave. The result is that beyond the wire there is no net field parallel to the wire. The radiation in the backward direction is not cancelled by the incident radiation and hence for polarisation parallel to the wire the grid acts as a reflector. The electrons hardly move perpendicular to the wires and hence the component of **E** in this direction is transmitted almost unaffected. Common polaroid sheet (called H sheet) essentially works like a wire grid polariser. In this material the sheet is stretched out so that its constituent PVA (polyvinyl alcohol) molecules align. These molecules are doped with iodine to make the polymer chains conducting. By this means, a 'wire' grid is created on a molecular scale. Cunning, really.

Malus' law

The sketch below shows a polariser producing linearly polarised light of amplitude E_o. If this light meets a second polariser (called an analyser) at angle θ, then *Malus' law* tells us that the amplitude of the polarised light transmitted through the analyser is $E_A = E_o\cos\theta$. Now



remember that the intensity of the light is proportional to the square of its amplitude. Hence Malus' law tells us that the intensity of the emerging polarised light ∝ $\cos^2\theta$.

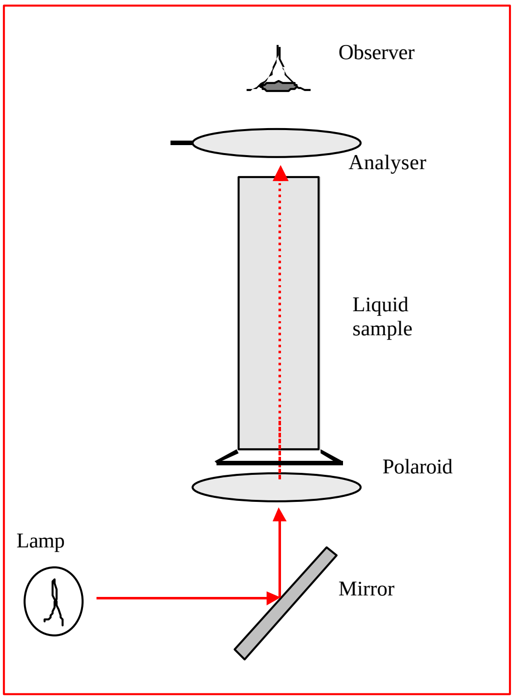
$$I_A = I_0 \cos^2 \theta \qquad .$$

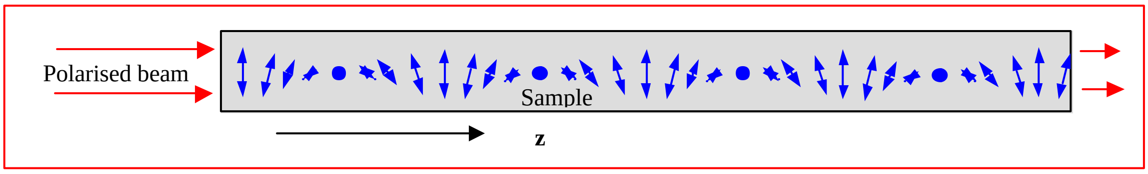
a) If the unpolarised light incident on the first polaroid has intensity I, then 50% of the light will be transmitted through an ideal polaroid. There is no transmission of the perpendicularly polarised component.
b) If the analyser is at 90° to the polariser, then, to recap, the two polaroids are said to be **crossed** and the intensity getting through is zero.
c) You should be able to see (and show) that if a polaroid is inserted at 45° between crossed polaroids, then light will appear from the final polaroid, equal to 0.25 of the intensity of the light emerging from the first polaroid.
d) If you insert two polaroids between crossed polaroids, each inserted polaroid inclined at 30° to the previous one, then the light appearing from the final (4^{th}) polaroid will be 0.42 of the intensity emerging from the first polaroid. The more (ideal) polaroids you put in at equal angles to each other, the more the direction of polarisation will be slowly rotated round, with less loss.

Some materials have the property of being able to rotate the direction of linear polarisation, on account of the handedness of their molecules. We shall look into this in more detail in the next section.

Optical activity

Optical activity is the ability of materials to rotate the direction of linear polarisation as the light travels through. It is of particular interest to structural chemists and biologists. The archetypal example is shown by natural dextrose ($C_6H_{12}O_6$) extracted from cane or beet sugar. The illustration shows how it can be observed. Putting water in the cylinder and placing it between crossed polaroids has no effect on the dark field seen by the observer. Putting in dextrose (or other sugars) causes the direction of polarisation of the light to be rotated and hence some light to emerge through the top analysing polaroid. Dextrose always rotates the direction of polarisation clockwise, when looking towards the source. Such materials are called **dextrorotatory** or d-rotatory. This is how dextrose has got its name. Materials that rotate the direction of polarisation counterclockwise are called **levorotatory** (sometimes laevorotatory) or l-rotatory.





The previous sketch tries to show what is happening for a long sample. Natural turps, quartz and sodium chlorate are other optically active materials. A measure of the optically activity of a sample is the rotation produced for a 1 mm slab for a solid, or a 100 mm path length for a liquid. This measure is called the **specific rotation**. For quartz it is 21.7° (mm^{-1}). For turps it is -37° $(100\ mm)^{-1}$. Liquids usually rotate the light much less than solids. Solutions of solids will obviously show an effect that depends on the concentration of active material and, to a small extent, both on temperature and the solvent.

What is happening here? The explanation in general terms is both subtle and elegant. What is happening is that left and right circularly polarised light travel through the material at different speeds. The refractive index is therefore different for right and left-handed circularly polarised light. Call these refractive indices n_R and n_L. When the linearly polarised light propagates into the material it excites both right and left circularly polarised light. Because one travels more slowly than the other, a phase shift builds up between them. When you combine right and left-handed light together at any point, the result is always linear polarisation but the angle of the polarisation depends on the phase shift *between* the two circular polarisations. The mathematics is not too difficult for Hons physics students to work out. The result is that the angle of polarisation, θ, is given by:

$$\theta = (n_R - n_L)k_{vac} z / 2$$

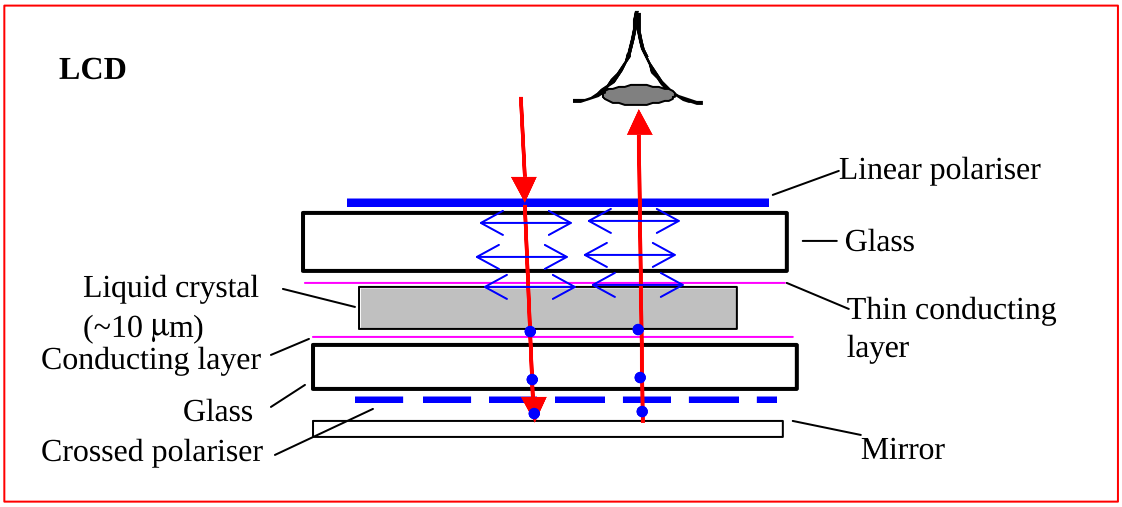
where z is the coordinate in the direction of travel of the light through the material. As a result, the direction of polarisation spirals around **z** in space.

The cause of this special behaviour for circularly polarised light is, basically, that the underlying molecules are **chiral**, meaning that they have a helical twist in them. Any arrangement of atoms with a helical structure can form left-handed helices or right-handed helices. One is a mirror image of the other. Now there is a curious property of helices not possessed by rotating circles. If you look at a circle rotating clockwise (say a bicycle wheel as the bike passes you from left to right) and then go around to the other side, the same wheel appears to go around anticlockwise. The 'wise' description just depends on your viewpoint. Helices are different. If you turn round a right-handed helix (in which the spiral appears to go around clockwise as you look into the helix), then it is still a right-handed helix. So, helical molecules may interact with circularly polarised light differently, depending on whether their chirality is the same handedness as the circular polarisation or opposite handedness. This is the fundamental reason for the two refractive indices n_R and n_L. It's interesting that the cause of the phenomenon lies at a molecular level and does *not* have anything to do with spatial ordering of molecules. This is why the phenomenon occurs with liquids as well as solids.

Why does the effect interest biologists? If you make sugar by chemical reaction, then you find it is not optically active. In fact it contains equal amounts of dextro- and levo-rotatory molecules. Biologically produced sugar is always dextro-rotatory. There seems to be a handedness in all biologically chiral molecules. All but one of the basic amino acids of life are levo-rotatory (the one is not optically active). If you are synthesizing a drug that works by locking into a biologically active site, if that site has a handedness, if it is chiral, then you must produce the drug with the right optical activity. Is all life based on levo-rotatory amino acids? Almost. Why? [This question is left as an exercise for someone's career]. Structural chemists need to be aware that the left-handed and right-handed forms of molecules will have exactly the same chemical energy and may be present together when they are trying to work out the arrangement of atoms in the molecules. What is a single chemical species of sample, may be a mixture of physical structures. Another simple application of optical activity by chemists is to measure its effect in a sugar solution and take the activity as a measure of the sugar concentration (e.g. in differently concentrated fruit juices). This is much quicker than detailed chemical analysis of each solution. Hecht gives the example of using the effect to measure sugar concentrations in urine, to test for diabetes.

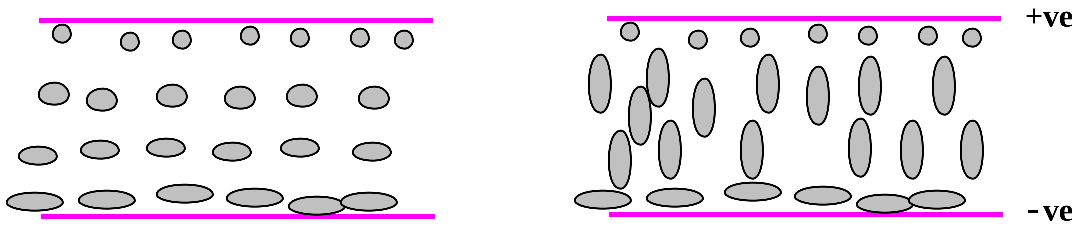
Liquid crystal displays

Liquid crystal displays (LCD) first became common in calculators, digital clocks and digital watches but are now used on much bigger flat screen imaging displays as well. Vast sums of money have been spent developing them and it's perhaps not surprising that a whole range of different types are made, some of them pretty complicated. They all consist of picture elements (pixels), not necessarily small dots but sometimes elements quite large and specially shaped in the form of symbols. The pixels are either bright (often silvery, sometimes coloured) or dark. Virtually all achieve the transition from dark to light by using crossed polarisers to produce the dark and then, to produce the light, they change the state of polarisation of the light in the pixel so that it is no longer absorbed by the second polariser. Nice idea, how does it work?



The story is a bit complicated. The most popular type is the **twisted nematic** LCD. At the heart of the LCD 'cell' is a liquid whose long molecules naturally tend to order. This order makes a powerful optically active material that can rotate the direction of linear polarisation by 90° in a distance of typically 10 μm. When an electric field is applied (e.g. 5V over 10 μm produces a field of 5×10^5 Vm^{-1}) the optically active ordering is destroyed very quickly and the rotation disappears. That's the optics. What's needed to put it into practice is roughly the following.

Look at the diagram. There are eight layers in the device. External light shines onto a linear polariser. The polarised light shines through a thin glass sheet and a transparent conducting layer into the liquid crystal. Its polarisation is rotated 90° by the liquid crystal material and it passes out the back through a second transparent conducting film, a second polariser crossed with respect to the first and finally onto a mirror. Reflected from the mirror, the light tracks back its path in reverse. The observer sees a bright region of light. Now apply an electric field across the conducting layers. The rotation of the direction of polarisation in the cell is destroyed. The incident light reaches the back polariser and is absorbed. No light is reflected back so the cell now appears dark.



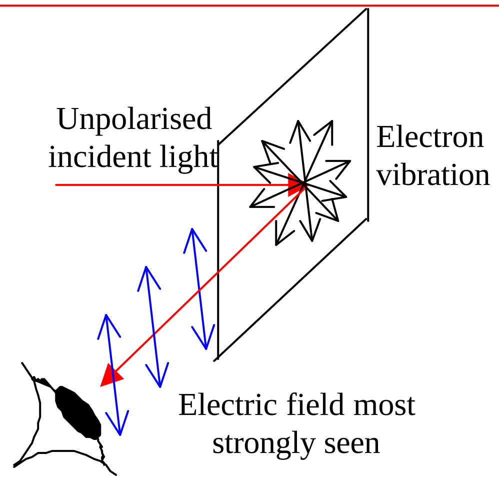
No voltage
Natural state with built-in 90° twist.

Applied voltage
Molecules away from surface re-align

The wizardry lies in the liquid crystal material. The lozenge-shaped molecules themselves naturally orient in line with a chemical 'surfactant' spread onto the conducting coatings. The surfactant is spread on the bottom coating at 90° to the direction on the top coating, so a twist of 90° throughout the liquid crystal is 'built in'. It is this twist, along with a careful choice of liquid, that gives the cell is huge optical activity. This is the natural state when no electric field is applied. Now apply an electric field and the molecules respond (apart from a very thin layer near each surfactant) by aligning themselves with the field that directly runs between the conducting layers. The optical activity is destroyed and the cell goes dark, for the reasons given above. Now you see how it's done.

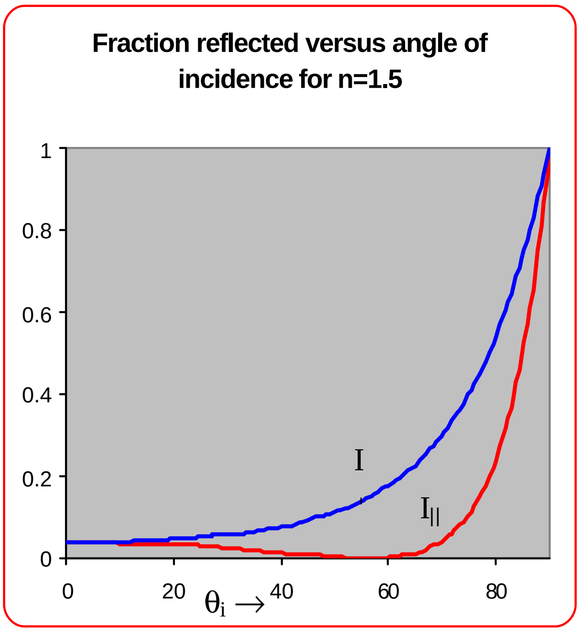
Polarisation by scattering

The phenomenon comes about because incident light causes electrons in the scattering medium to vibrate. A vibrating electron emits most light in a direction perpendicular to its vibration and none along the direction of its vibration. The electric field of the emitted radiation is parallel to the direction of electron vibration. Hence light scattered through about 90° is strongly polarised. In the accompanying diagram, an observer sees no polarisation perpendicular to that shown, because the electrons do not vibrate in this perpendicular direction, which is parallel to the direction of travel of the incident light.



The light from a blue sky is quite strongly polarised, particularly at 90° from the sun. It is not completely polarised because a significant amount of sunlight has undergone multiple-scattering, i.e. has been scattered more than once. Light scattered twice through a total angle of 90° will be less polarised than light scattered once.

Bees and other animals make use of the polarisation of blue sky for navigation. See Können's book, referenced in a class list issued near the beginning of the course, for more on the natural occurrence of polarised light.



Polarisation by reflection

The curves that give the percentage of reflection and transmission of light incident at angles from 0° to 90° are known as the Fresnel reflection coefficients. They were derived by Fresnel from an elastic model of the response of a material to light. Nowadays, they are considered as consequences of Maxwell's equations with accompanying boundary conditions applied to the reflecting interface. This generality allows reflection to be calculated for any wavelength of electromagnetic radiation and for complicated interfaces such as multi-layer devices.

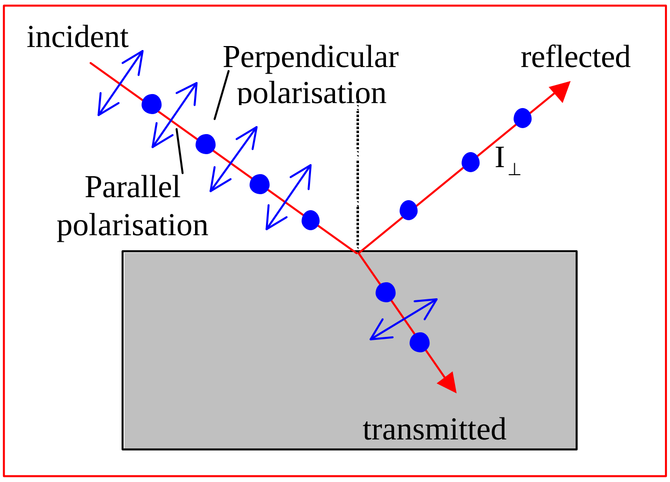
The fraction of light reflected depends on whether the light has electric vector parallel ($I_{||}$) or perpendicular ($I_{\perp}$) to the plane of reflection. $I_{\perp}$ increases steadily from an angle of incidence of 0° to 90°; $I_{||}$ goes to zero at an angle θ_B called the **Brewster angle,** which is simply related to the refractive index n.

Brewster angle θ_B:

$$\tan\theta_B = n \quad .$$

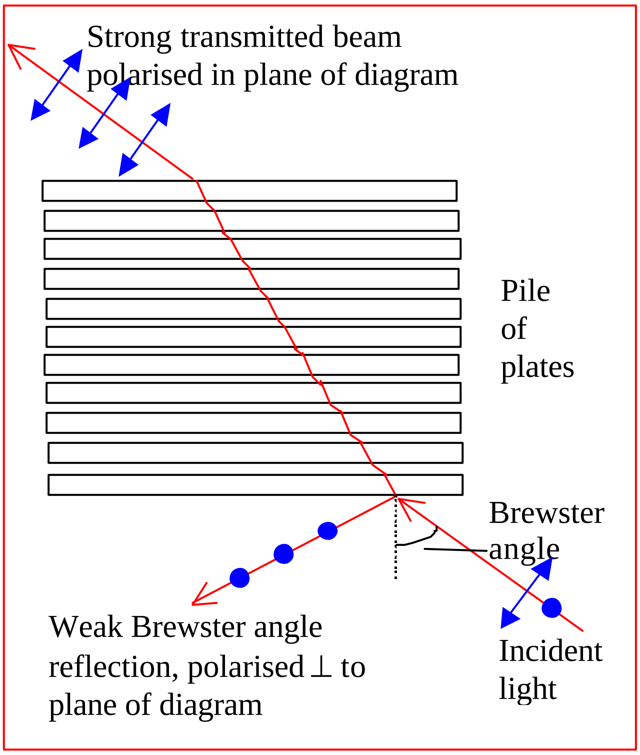
e.g. for n = 1.5, θ_B = 56.3°, as in the diagram above.

You can deduce the formula for the Brewster angle if you know that the reflected and transmitted rays are at 90° to each other when the angle of incidence is the Brewster angle.



Polarisation by reflection can be achieved in principle by a single reflection of a ray of light exactly at the Brewster angle. In real life, this method doesn't produce 100% polarised beams because in a beam of light, all the rays within the beam are never at the correct angle. Only some of them are. Another disadvantage of the method is that the fraction of incident light that is polarised perpendicular to the plane of incidence that is actually reflected is about 20% for ordinary glass. The efficiency of selecting the one polarisation is therefore not very high.

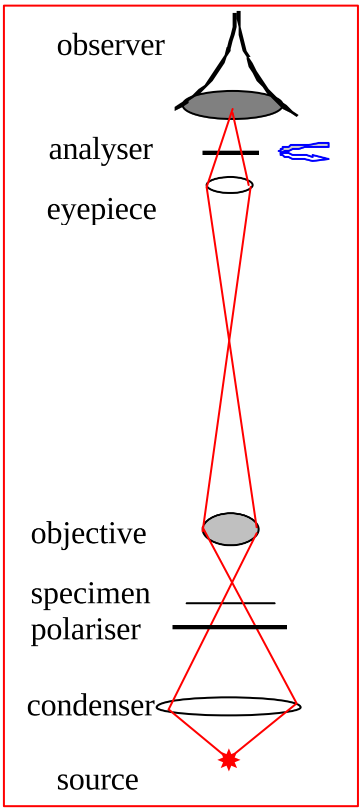
A better arrangement is a 'pile of plates' polariser. This selects the transmitted ray, which is more polarised than the incident ray and contains virtually **all** the light polarised in the plane of incidence when the angle of incidence is near the Brewster angle. At each successive plate, the light becomes more polarised without any loss of the parallel polarised component. After transmission through say 10 plates, the light is highly polarised in the plane of incidence, with little loss of intensity of the polarisation in the direction required.



The polarising microscope (petrological microscopic)

The pioneering Scotsman William Nicol invented the first really efficient man-made polariser of light, called the Nicol prism after him. The Nicol prism is made from two suitably cut pieces of the mineral calcite. It lets through almost all of the light polarised along its direction of polarisation. In fact it's better in this respect than modern polaroid sheet, transmitting a greater amount of light over a much wider wavelength range, which extends well into the UV. However, a Nicol prism is bulky, expensive and restricted in size by the availability of calcite.

William Nicol is also credited with co-developing the technique of examining thin mineral samples in polarised light and bringing to the notice of geologists a technique that has served them well for over a century and half. This technique is the basis of the petrological microscope. [William Nicol was an itinerant lecturer who made his living giving public lectures on science to audiences of several hundred at a time, a couple of centuries ago. He once refused an invitation to take up a University professorship, because he found his own way of living a more satisfying job].



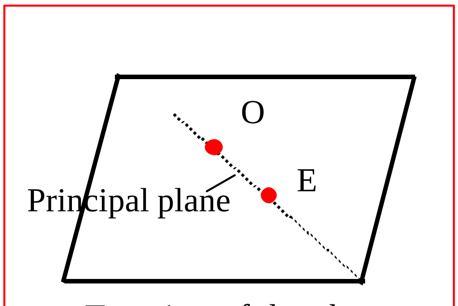
Hecht doesn't introduce the polarising microscope into his book. A pity. The essential difference between a petrological microscope and an ordinary microscope is that the petrological microscope linearly polarises the light incident on the sample. Just before the light emerges from the microscope, there is an analysing polariser, which can be rotated by any angle with respect to the initial polariser. If the analyser is set at 90° to the polariser, then no light gets through the microscope unless the object alters the direction of polarisation of the light. Remember the discussion of Malus' law. One of Nicol's discoveries was that most minerals do do something to the polarisation of incident light. This is because they are **birefringent**. How do you get light through rock samples like granite, sandstone, etc.? As Nicol showed, you cut very thin sections of them and place these thin and fragile samples on a glass microscope slide. The empty slide does nothing to the polarisation. The mineral may well do. To get a better idea of what might happen, we need to look at the phenomenon of birefringence.

Birefringence

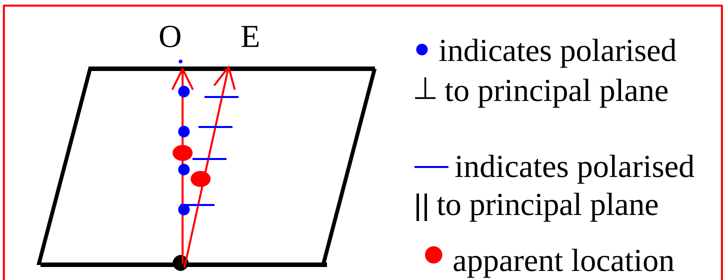
As light became more and more investigated in times past, people became very familiar with how it behaved when it travelled through transparent materials such as water, glass, alcohol, diamond and so on. Then came the discovery of 'Iceland spar'. This mineral, known more commonly now as calcite, is transparent and could be found in large chunks. When you look through a big piece, you see clearly a double image, even when the surfaces are perfectly flat. Something is happening to the light as it travels through the crystal, something that is quite different from the Snell's law refraction we've met so far. Iceland spar, which is just crystalline $CaCO_3$, is an example of an **anisotropic solid**.

Anisotropic materials are materials that behave differently in different directions. It turns out that most solids are anisotropic. Solids with an underlying cubic symmetry arrangement of their atoms are isotropic. So are solids with random atomic positions, on the scale of light waves, such as glasses and liquids. The rest are anisotropic.

Now, optically anisotropic solids don't just behave slightly differently from usual. They exhibit a whole range of new phenomena. Since most minerals are anisotropic, the geologist and gemnologist needs to have a working knowledge of anisotropic optical effects. So do many materials scientists, since these effects are used for identification, quality control, as a measure of strain, and so on. $CaCO_3$ is a very good example to study, since the effects are very strong in this material. It's used in all textbooks as the archetype of a **birefringent** material. Birefringent because light is generally transmitted as two rays, not simply one.



Top view of rhomb

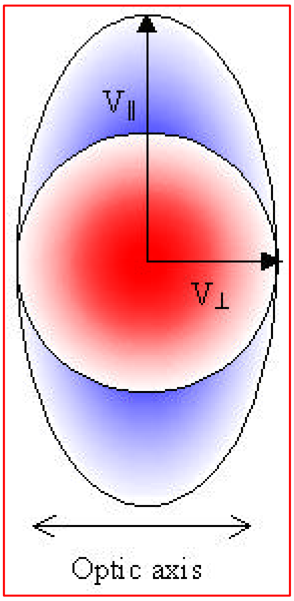


Side view of rhomb in principal plane

When you look more closely, you'll find that one of the rays coming through Iceland spar obeys Snell's law; the other does not. For example, one ray incident perpendicularly on a face goes straight through, the other goes off at an angle. *Demo*. Why does it go at an angle in the particular direction it does? Why do the images appear at different depths within the crystal? You'll also find that the rays are both polarised. For crystals cut in certain special ways, only one image is seen. When you take the emerging rays and pass them into another crystal, sometimes four emerge and sometimes, depending on the orientation of the second crystal with respect to the first, just two emerge. The phenomena to be explained are complicated, unusual and apparently baffling.

In his astonishingly perceptive *Traité de la lumière*, my friend Christiaan Huygens gave the essentially modern explanation for the behaviour of Iceland spar, and hence the behaviour of any birefringent material. His reasoning wasn't entirely modern, but the results are. He began to make sense of what was happening by deducing that the properties seem to be centred on a direction inclined equally to three intersecting natural crystal faces. This direction is called the **optic axis**. Huygens thought that this direction was special because of the way the elementary units of the material packed together. In modern language, $CaCO_3$ is a trigonal crystal and the optic axis is the special axis in this crystal class, namely the "c axis" in the international notation. Huygens wrote about 150 years before molecules were seriously considered as the basis of material structure and 200 years before crystal symmetry was first completely categorized.

When you turn a rhomb of calcite around over a printed dot, the ordinary image stays stationary and the extraordinary one rotates around it. The extraordinary image is always displaced in a plane containing the optic axis. Such a plane is called a **principal plane**. The principal plane contains the optic axis and is perpendicular to the surface. The real dot is directly below the O (ordinary) image. If you could see the rays in that plane, you would see two rays diverging from the dot, as in the second picture already given.

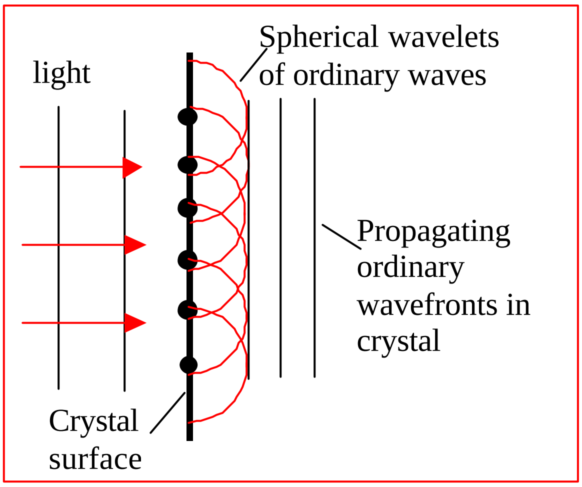


Calcite is an example of a simpler kind of birefringent material called **uniaxial**, because it has one optic axis. Guided by the fact that two rays are present within the crystal, Huygens postulated that the incident light excited two separate wavelets within the crystal, one spherical wavelet associated with the ordinary waves and one ellipsoidal wavelet. The spherical wavelet propagates equally quickly in all directions. We now know what Huygens didn't, namely that the ordinary ray is always polarised with its electric vector perpendicular to the optic axis. Polarisation was not discovered until over a century after Huygens' death. The polarisations of the waves are shown in the earlier diagram. The ellipsoidal wavelets are polarised with electric vector parallel to the principal plane. They propagate at different speeds in different directions, the slowest speed (in calcite) being in the direction of the optic axis. There are thus two refractive indices $n_o = c/v_{\perp}$ and $n_e = c/v_{\|}$. n_o stands for the refractive index associated with the polarisation perpendicular ($\perp$) to the optic axis, namely the **ordinary ray**. n_e is the smallest refractive index associated with the **extraordinary ray**, representing the speed when it is travelling perpendicular to the optic axis with electric vector parallel ($\|$) to the optic axis. Notice that the optic axis is the direction in which the two wavelets always coincide and travel with the same speed.

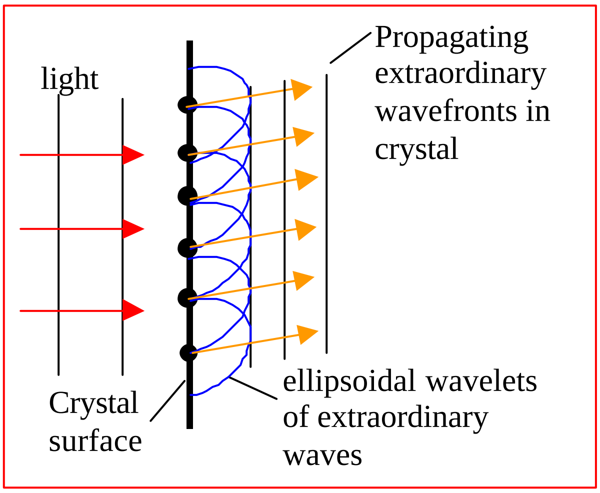
For calcite, $n_o = 1.658$ and $n_e = 1.486$ (at the d wavelength). This is a much bigger difference in refractive indices than you'll find in most materials. Because n_e is less than n_o, the material is called a **negative uniaxial** crystal.

How does Huygens' principle explain the propagation of two waves?

The ordinary waves behave much as we've met in Huygens' principle before. The plane wavefront arriving normally at the crystal surface generates spherical wavelets, travelling equally fast in all directions with speed $v_{\perp}$. The tangent to these waves lies straight ahead and, by successive application of the principle, the plane wave propagates straight ahead with speed $v_{\perp}$. The only difference between this and normal refraction is that the travelling wave is completely polarised, with its electric field perpendicular to the optic axis.



The extraordinary wavelets are more subtle. The wavelets spread out in ellipsoidal shape. The common tangent to these ellipsoids after a little time is the new wavefront. The line from the point of generation of each ellipsoid to the tangent point on that ellipsoid is off at an angle, represented by the orange arrows in the coloured diagram. These arrows define the direction of travel of the extraordinary wavefronts. What is odd is that the wavefronts are NOT perpendicular to their direction of travel. Inside the birefringent crystal, the extraordinary wavefronts are not purely transverse.



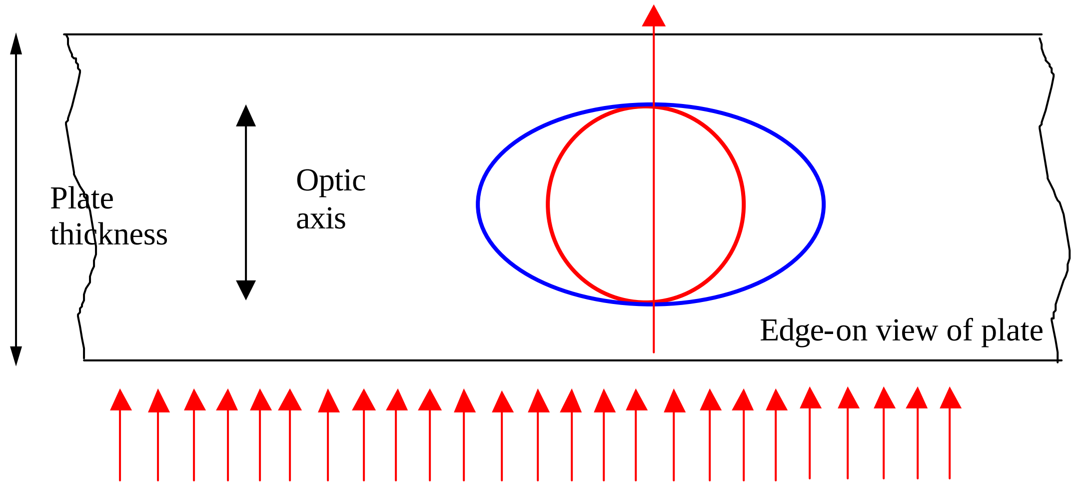
Batholinus, a contemporary of Huygens, discovered the double refraction of Iceland spar, but not the explanation. He said: "diamond may be prized by men but those who are fascinated by knowledge of unusual phenomena will find great joy in a new crystal, recently bought from Iceland, which is perhaps one of the greatest wonders that nature has produced".

You must admit that Huygens was brilliant to deduce how calcite works. I'm sure you'll have to read these notes and a textbook more than once to be confident of what's going on. Huygens didn't know that a good many minerals are a bit more complicated even than calcite. These minerals have two optic axes and are called **biaxial crystals**. In these crystals no new phenomena appear. It is just slightly harder to describe what is going on. Biaxial crystals have three refractive indices. It was the equally brilliant Augustin Fresnel who realized that biaxial crystals could be described by an ellipsoid of index of refraction, having three different axes all at right angles to each other. I think it's fair to say that you have to be a mineralogist with a need to know before you spend the effort needed to get a good idea of how to follow the propagation of light through biaxial crystals.

The final point of this section is the simple observation that the birefringence of a crystal is related to the symmetry class of its crystal structure. Any crystallography course you take will tell you that crystals fall into seven crystal classes, with the following names:

- *Cubic* crystals are isotropic. (e.g. rock salt, NaCl)
- *Tetragonal*, *hexagonal* and *trigonal* crystals are uniaxial. (E.g. zircon, apatite, calcite respectively).
- *Orthorhombic*, *monoclinic* and *triclinic* crystals are biaxial. (e.g. olivine, biotite, kaolinite respectively).

Making birefringence work for us



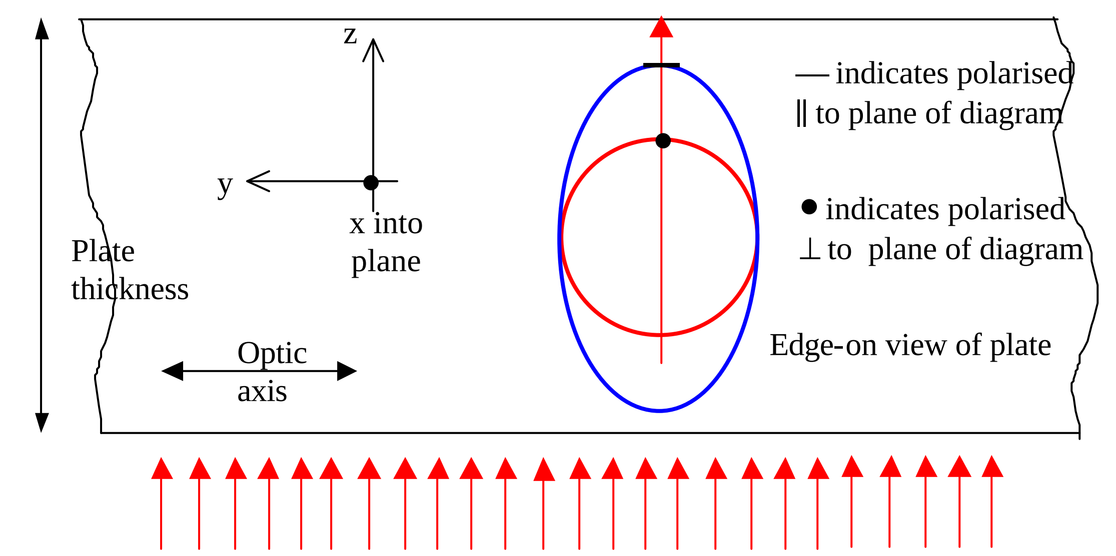
Light incident when the optic axis is ⊥ to the plate – no interesting effects

Birefringence is exploited by looking at the light travelling through thin plates of birefringent material. Light is usually shone onto the plate perpendicularly. If the optic axis is also perpendicular to the plane of the plate, as in the diagram above, then the birefringence has NO interesting effect. Both wave surfaces coincide in this direction. Both polarisations of light travel at the same speed and the propagation is more or less as happens in 'ordinary' materials.

Interesting effects happen when THE OPTIC AXIS LIES IN THE PLANE OF THE PLATE (see the diagram on the next page). Now the two polarisations travel perpendicular to the plate with different speed. As a result, there is a changing phase between the two components. They start out in phase but this soon changes. Two plane polarised waves become one elliptically polarised wave. Why?

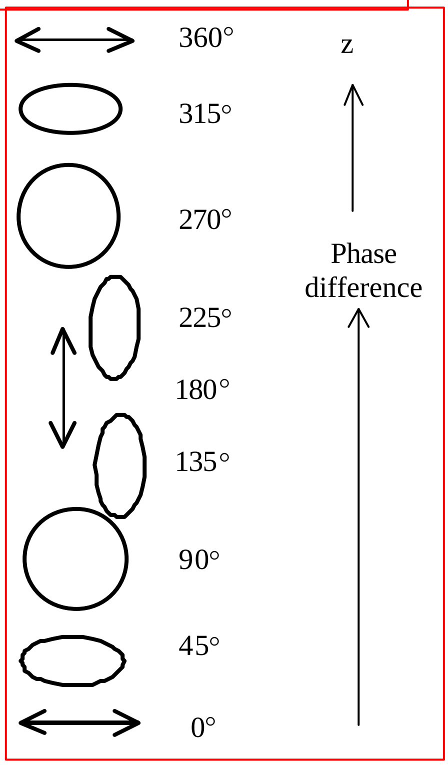
The x component of the electric field (see next diagram) is provided by the expanding ordinary wave, which encounters a medium of refractive index n_o. The optical path length for a given distance z traveled by the wave is $z \times n_o$. The y component is provided by the extraordinary wave, which encounters a refractive index n_e. The optical path length for this wave is $z \times n_e$. The difference in path length introduces a *phase shift* between the components

$$
\begin{aligned}
E_x &= E_{0x}\cos(kz-\omega t)\\
E_y &= E_{0y}\cos(kz-\omega t+\varepsilon)\\
&= E_{0y}\cos(kz-\omega t+(n_o-n_e)k_{vac}z)
\end{aligned}
$$



Light incident when the optic axis is || to the plate

of $z(n_o-n_e)2\pi/\lambda_{vac}$ in radians. The mathematical will write the two components as:

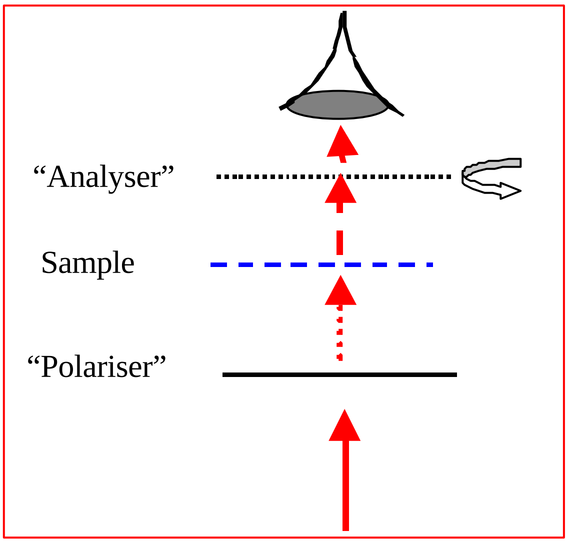


Let's take an example where both components of the light have the same magnitude, i.e. $E_{0x} = E_{0y}$. At the start, both components are in-phase and hence the light is linearly polarised out of the picture at 45° to the plane of the diagram. When the phase shift is 180°, the light is again linearly polarised; when the phase shift is 90° or 270°, the light will be circularly polarised (see the earlier description of circularly polarised light); in between it is elliptically polarised. The light therefore goes through a succession of polarisation changes as it propagates through the plate.

In calcite, the difference in the refractive indices is large at 0.172 and hence the phase shifts between the two components change by 2π (≡ 360°) every $\lambda_{vac}/0.172 \approx 3$ μm. For a more typical mineral like muscovite (a variety of mica) where its two refractive indices are 1.599 and 1.594, the change takes place across a thin plate of about 120 μm in width. This is much more manageable. A standard thickness for geological specimens is 30 μm. (*Demo on overhead*).

Minerals and the microscope

You can now understand some of what geologists see through a petrological microscope. The gist of the arrangement is to look at a thin plate of a rock sample placed between polarisers, usually under the microscope because the individual grains within the rock are very small. The first polariser, usually called "**the polariser**", linearly polarises the light incident onto the sample. An **analyser**, in the form of a polariser that can be rotated, is placed between the sample and the observer looking at the specimen. If the analyser is placed at an orientation of 90° from the polariser, then if the sample is just 'ordinary' material the observer will see black, i.e. no light. If the sample is a thin birefringent plate with its axis in the plane of the specimen, then the transmitted light will emerge as two polarised components whose phase difference on emerging will be some angle that isn't special. Hence, some of this light will be transmitted by the analyser and the birefringent crystal will be made visible. The device will therefore distinguish birefringent minerals with their optic axis approximately in the plane of the sample.



In most circumstances the incident light is white. Whatever the thickness of the specimen, for some wavelength the phase difference between the polarisations will be exactly 360°. For this wavelength, the light will emerge with the same polarisation as the incident light and hence will not be transmitted by an analyser set at right angles. The device will therefore take out some of the colour of the light and leave the transmitted light approximately the complementary colour. Hence, the birefringence not only shows up, but it produces very pretty colours in each mineral grain. The most intense **polarisation colours** are seen when the optic axis is at 45° to the orientations of the crossed polariser and analyser. The skilled mineralogist can estimate the amount of birefringence from the maximum colouring observed in grains of a mineral, and hence perhaps identify the mineral itself.

On top of the colouring produced by subtraction of some of the spectral colours from the incident light, some minerals show colouring by absorption. For example, biotite has a characteristic yellow colour. Moreover, the colour may depend on the polarisation of the transmitted light. This phenomena is called **pleochroism**. Biotite is pleochroic and if you rotate a sample in the petrological microscope then its colour will change from pale straw yellow to dark brown, depending on the orientation of the optic axis relative to the polarisation direction of the incident light.

It is astonishing what mineralogists can deduce from the behaviour of polarised light traversing their specimens. Rotate the specimen on the turntable built into the microscope. This varies the ratio of E_{0x} to E_{0y}. When the optic axis of the specimen lies ‖ or ⊥ to the direction of polarisation emerging from the polariser, only one wave is excited within the crystal and you will see black. This is called **extinction** by the mineralogists. Hence, you can tell within 90° what is the direction of the optic axis of the mineral grain that shows extinction and look and see how this direction is related to the shape of the grain. You can see that there is the making of a powerful analytical tool here and mineralogist have indeed made it so since the time of William Nicol. I'm not going to say more about what happens when the sample isn't nicely oriented.

Birefringence for materials scientists

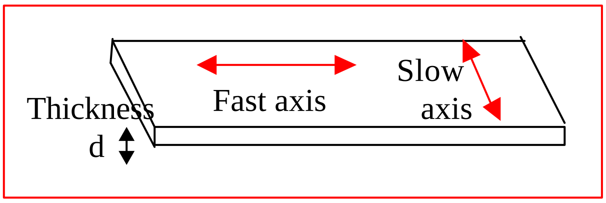
The orientation of mineral grains may not be easily predictable but induced birefringence in transparent specimens is, particularly when the inducing mechanism is applied strain. Strain can induce birefringence in glass and plastic that is normally isotropic. Sometimes the strain can be 'built-in' when the material was being formed in a mould or die. (*Demo of strain patterns in plastic protractors, rulers, etc.*)

The method is very sensitive to strain. A 1 mm thick specimen will induce a phase change of 360° between the two polarisations when $(n_o - n_e) \approx 5\times10^{-4}$. When the phase change is constant, i.e. the strain is constant, you will see lines of constant colour. When the stress increases, so the strain increases and lines of constant colour will pack closely together. Residual internal strain and strain concentrations are easily seen in plastic mouldings. Engineers used to make models of bridge cross-sections and many other objects to look at the strain concentrations. Nowadays, the preferred method is to use finite element analysis computer packages.

Once you know what to look for, you can see such patterns in objects out of the laboratory. For example, the sky can produce significantly polarised light. When this is transmitted through a car windscreen made of toughened glass that has a substantial built in strain, then if you look through the windscreen with polarising sunglasses, you may see tell-tale coloured evidence of this strain. My take-home message is that birefringence is both a fascinating phenomenon and a practical one too.

Retarders

Common optical components called retarders are made from birefringent materials. They are called quarter-wave plates, half-wave plates and full-wave plates. Their uses are to produce circularly polarised light and to analyse polarised light into its full elliptical components. I mentioned earlier another useful application of circularly polarised light for improving the contrast of digital displays. In this section, we look at retarders.



A **retarder** is just a uniform plate of birefringent material whose optic axis lies in the plane of the plate. For calcite, the **fast axis** (the refractive index *n* is least) is the axis with electric vector parallel to the optic axis. Perpendicular is the **slow axis**, so called because *n* is largest. The previous discussion of birefringent plates was essentially an explanation of how retarders work. They divide the incident wave into two polarisations that travel perpendicular to the plate at different speeds. A phase retardation of one wave relative to the other is therefore given as the waves cross the thickness *d* of the plate. This phase retardation $\Delta\varphi$ in radians becomes:

$$\boxed{\Delta\varphi = k_{vac} d \left(n_o - n_e\right)} \quad ,$$

almost in *Hecht's* notation.

In terms of the more general ideas of fast and slow axis:

$$\Delta\varphi = k_{vac} d\left(n_{slow} - n_{fast}\right) \quad .$$

The point about retarders is that they change the state of polarisation of an incident wave. Or at least they do if the incident wave is not itself polarised || or ⊥ to the optic axis, in which case a wave of only one polarisation is generated within the plate and relative retardation isn't an issue. If linearly polarised light is incident on a retarder, under most circumstances the result is a change to elliptical polarisation after travelling a distance *d*. Special cases are:

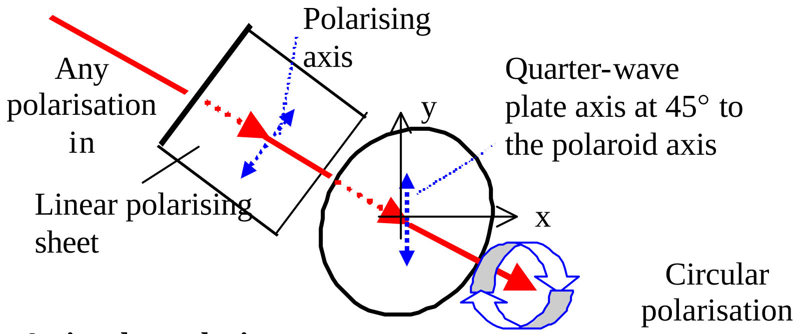
- A **full-wave plate** retards the slow wave relative to the fast wave by an integer multiple of 2π in phase, equivalent to adding a multiple of a wavelength to the path of one wavelet relative to the other
- A **half-wave plate** retards in phase by an odd number of π radians, usually just π radians or 180°
- A **quarter-wave plate** retards by an odd number of π/2 radians, usually just π/2 radians or 90°

The retarding effect of a plate is called its **retardance** and may be measured in wavelength, in nm, rather than in phase. Notice that all retarders are a set number of wavelengths (e.g. a quarter-wave) only at a particular optical wavelength. They are therefore **chromatic**, meaning that their performance is different for different colours.

You may well find a full-wave retarder in the petrological microscope, to help identify how the polarisation colours are produced in the microscope. One problem in identifying the birefringence of an unknown sample is to know how many wavelengths of retardation are producing the colours seen. A clue is obtained from the brightness of the colours. If there is only about one wavelength of path difference, then the colours are more intense. The basic idea of having a retarder you can slide in between the sample and the analyzer is that if you add a whole wavelength of retardance, then if it makes only a little difference you have several wavelengths retardation already present; if it makes a big difference, you probably have only one wavelength present. A skilled microscopist will tell you precisely how many wavelengths, upon seeing the change in colours upon introducing a full-wave plate when the sample is a standard thickness, such as 30 μm.

Making circularly polarised light

If you have a quarter-wave plate **and** you shine on incident light that is polarised at 45° to the optic axis, then you create a **circular polariser**. You can buy a circular polariser looking like a single sheet but it is in fact two sheets bonded together for this purpose. The first sheet is polaroid, the second a quarter-wave sheet.



A circular polariser

The x and y components of the emerging electric field are 90° out of phase. They can be written as:

$$E_x = E_o \cos(kz - \omega t)$$
$$E_y = \pm E_o \sin(kz - \omega t)$$

If you understand how sines and cosines are related, you'll see that the + sign happens when y component lags by 90° and hence the light is right circularly polarised, as shown in the diagram. The – sign corresponds to the y component leading the x component and the light being left circularly polarised. Which circumstance actually happens depends on whether the fast or slow axis of the quarter-wave sheet is placed along the y direction. If the slow axis is placed along y, then the y component will lag and you'll get right circularly polarised output. Therefore, rotating the quarter-wave sheet by 90° will change the output from one circular polarization to the other.

Notice that circular polarisers have an input side and an output side. Another of their properties, which you might be able to deduce for yourself, is that if you send a circularly polarised beam of the same handedness back into the output end, then it will emerge at the input end as linearly polarised light. If, instead, you shine in a beam of the opposite handedness, it will be stopped. This is how the contrast enhancer for digital displays, mentioned earlier, works.

That is certainly not the end of the story as far as polarisation is concerned. The **Faraday effect** is a rotation of the direction of polarisation of light by a magnetic field in the same direction as the light is travelling (a longitudinal field). John Kerr (a Scottish scientist, working in the 1890s) discovered that a transverse electric field applied to materials like CS_2 and nitrobenzene caused them to become birefringent, in proportion to the square of the electric field strength. A **Kerr cell** is used to switch light beams off and on, which it can do very quickly, or modulate the intensity of the beam according to an applied voltage. Likewise, Friedrich Pockels discovered at about the same time another electro-optical effect in which a longitudinal electric field makes non-centrosymmetric crystals birefringent, in proportion to the strength of the field. The effect is strong in materials like barium titanate, KDP (potassium dihydrogen phosphate), lithium niobate, etc., all well known materials that have been subject to a lot of practical and theoretical investigation in recent decades. A **Pockels cell** is also used for switching light beams off and on and is more compact than the Kerr cell.

The polarisation of light is a fascinating phenomenon. Polarised light occurs quite widely in nature; it has many technological applications; it has complexities and deep subtleties, of which there are certainly more than I've mentioned. I hope these lectures have aroused your interest.

End of polarisation

JSR