University of Victoria

Physical Optics

Jeremy Tatum

The LibreTexts libraries are Powered by NICE CXOne and are supported by the Department of Education Open Textbook Pilot Project, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

This text was compiled on 04/15/2025

TABLE OF CONTENTS

Licensing

*A detailed breakdown of this resource's licensing can be found in **Back Matter/Detailed Licensing**.*

CHAPTER OVERVIEW

1: Reflection and Refraction via Fermat's Principle and Huygens' Construction

Thumbnail: Diffraction of a plane wave when the slit width equals the wavelength. (CC BY-SA 3.0; Lookangmany).

1.1: Reflection and Refraction

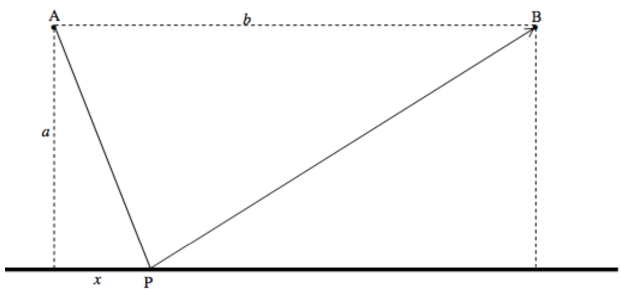
Reflection of light from a smooth, shiny surface is called *specular reflection*. (Latin speculum a mirror.) At the other extreme we have the sort of diffuse scattering that occurs when you shine light on blotting paper. And there are lots of situations in between these extremes. In this chapter I am going to deal solely with specular reflection, the law of specular reflection being that the angle of reflection is equal to the angle of incidence.

When light passes from one medium to another, the angles of incidence and refraction, and the two refractive indices are related by the familiar *Snell's Law,*

$$n_1 \sin\theta_1 = n_2 \sin\theta_2. \tag{1.1.1}$$

In this chapter we are going to look at the laws of reflection and refraction from the point of view of Fermat's Principle of Least Action, and Snell's law of refraction from the point of view of Huygens' construction. We'll start with the law of reflection (angle of reflection equals angle of incidence).

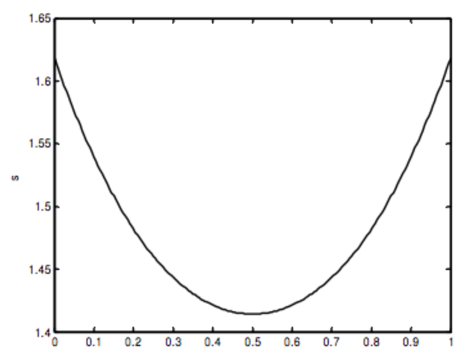
Light goes from A to B via reflection from a point P on a mirror.



The distance s travelled is given by

$$s = \sqrt{a^2 + x^2} + \sqrt{a^2 + (b-x)^2}. \tag{1.1.2}$$

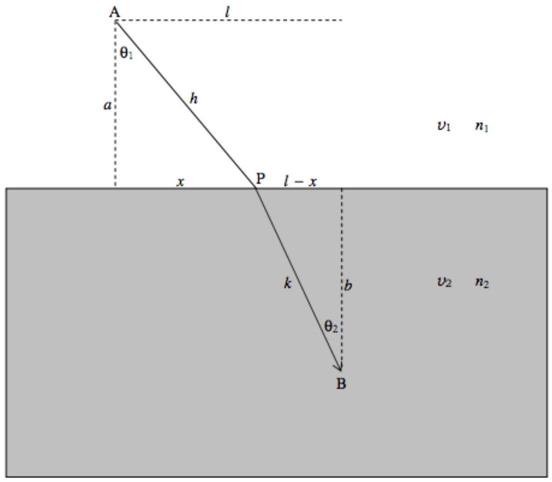
Here is a graph of s versus x (i.e. s as a function of the position of the point P from which the light is reflected). The drawing above is drawn for $a = \frac{1}{2}b$, and, in the graph below, s and x are in units of b.



From the graph, or by differentiation of s with respect to x (do it!), it is seen that the path length is least when $x = \dfrac{1}{2}b$, i.e. when the angle of reflection is equal to the angle of incidence.

Of all possible paths the light might conceivably take, the path that it actually takes is the one with the shortest path.

Now let us look at refraction at an interface.



In the drawing, light is travelling from A to B, first in a medium of refractive index n_1 (speed = v_1) and then in a medium of refractive index n_2 (speed = v_2), via the point P.

The time taken to get from A to B is

$$t=\frac{\sqrt{a^2+x^2}}{v_1}+\frac{\sqrt{b^2+(l-x)^2}}{v_2} \tag{1.1.3}$$

Given that $v_1=c/n_1$ and $v_2=c/n_2$, this can be written

$$ct=n_1\sqrt{a^2+x^2}+n_2\sqrt{b^2+(l-x)^2}\,. \tag{1.1.4}$$

If we vary the position of P, the time taken varies as

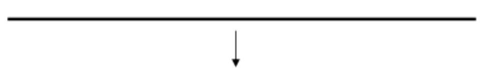
$$c\frac{dt}{dx}=\frac{n_1x}{\sqrt{a^2+x^2}}-\frac{n_2(l-x)}{\sqrt{b^2+(l-x)}^2}=n_1\sin\theta_1-n_2\sin\theta_2. \tag{1.1.5}$$

The route actually taken is such that for any small deviation dx from the route actually taken, the corresponding variation dt in the time taken is zero. That is to say, the derivative is zero, or $n_1\sin\theta_1=n_2\sin\theta_2$.

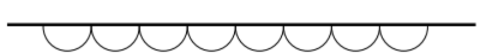
In both cases, reflection and refraction, the route taken is such that the time taken is least. This is an example of *Fermat's Principle of Least Action*. I am not sure that this is an *explanation* of *why* reflection and refraction happen the way they do as much as an interesting *description* of what happens in nature. A further example of the principle, from Classical Mechanics, is *Hamilton's Variational Principle*. The *action* that a mechanical system takes in going from one state to another is $\int Ldt$, where L is the Lagrangian at time t. The route that any mechanical system takes in going from one state to another is such that any small deviation from this route results in no change in the action (which usually means that the action is a minimum). Again, I am not sure that this *explains why* mechanical systems behave as they do. It is more a useful *description* of how mechanical events unfold.

Now let us looks at Huygens' Construction. Imagine that you are following the progress of a wavefront, and you can see the wavefront at some instant of time, and that you want to know what happens next. Huygens' construction supposes that any point in the wavefront can be regarded as point source for a new disturbance.

For example, in the drawing below, we have wavefront moving from top to bottom.

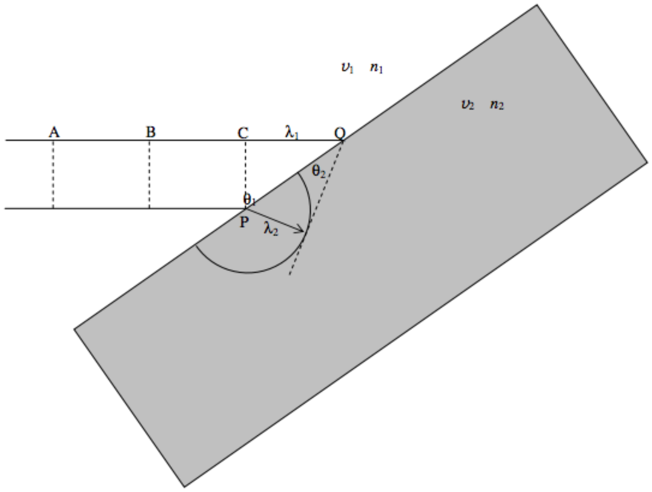


Now we suppose that every point on the wavefront is a point source for a new disturbance:



The tangent to these little wavelets is the new wavefront:

Now let us look at refraction at an interface between two media.



The vertical dashed lines at A, B, C represent wavefronts, separated by the wavelength λ_1. The angle of incidence is θ_1. At some instant of time, the lower ray reaches the point P, and it starts to generate a new wavelet. At a time P later, where P is the period of the electromagnetic oscillations, the upper ray has reached the point Q, where CQ = λ_1, while the wavelet generated at P has attained a radius $\lambda 2$. Here $\lambda_1 \lambda_2 = v_1/v_2 = n_2/n_1$.

The new wavelet generated at Q hasn't started yet, or at least is just about to start. The new wavefront is constructed by drawing the tangent from Q to the wavelet generated at P. From the geometry of the drawing we see that

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{\lambda_1/\mathrm{PQ}}{\lambda_2/\mathrm{PQ}} = \frac{\lambda_1}{\lambda_2} = \frac{n_1}{n_2},$$

and so Snell's Law is derived from the Huygens Construction. I leave it to the reader to decide whether this *explains* what happens, or merely *describes* what happens. Perhaps in science we never "explain" nature - we just "describe" it.

CHAPTER OVERVIEW

2: Reflection and Transmission at Boundaries and the Fresnel Equations

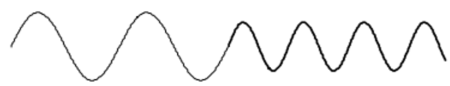
When a ray of light encounters an interface between two media of different refractive indices, some of it is reflected and some is transmitted. This chapter will concern itself with how much is reflected and how much is transmitted. (Unless the media are completely transparent, some of the light will also be *absorbed* - and presumably degraded as heat - but this chapter will concern itself only with what happens at the interface, and not in its passage through either medium.) We shall do this at three levels: Normal incidence; incidence at the Brewster angle (we'll explain what is meant by this); incidence at an arbitrary angle.

Topic hierarchy

Thumbnail: Reflection at a surface. (Public Domain; Benbuchler).

2.1: Waves in a Stretched String

Before discussing the reflection of light, it will be useful to discuss the following problem. Consider two ropes, one thin and one thick, connected together, and a sinusoidal wave moving from left to right along the ropes:



The speed c of waves in a rope under tension is $c=\sqrt{F/\mu}$, where F is the tension, and μ is the mass per unit length, so the speed and the wavelength are less in the thicker rope. We'll call the speed in the left hand rope c_1 and the speed in the right hand rope c_2. At the boundary $(x=0)$, some of the wave is transmitted, and some is reflected. (I haven't drawn the reflected part in the drawing). We wish to find how much is transmitted and how much is reflected. I'll call the amplitudes of the incident, transmitted and reflected waves 1, T and R respectively, and I'll suppose that the wave is a sinusoidal wave of angular frequency ω. The equations to the incident, transmitted and reflected waves are as follows:

$$y=\cos\omega\left(t-\frac{x}{c_1}\right) \tag{2.1.1}$$

$$y=T\cos\omega\left(t-\frac{x}{c_2}\right) \tag{2.1.2}$$

$$y=R\cos\omega\left(t+\frac{x}{c_1}\right) \tag{2.1.3}$$

To the right of the boundary, the displacement as a function of x and t is

$$y=T\cos\omega\left(t-\frac{x}{c_2}\right) \tag{2.1.4}$$

and to the left of the boundary the displacement is

$$y=\cos\omega\left(t-\frac{x}{c_1}\right)+R\cos\omega\left(t+\frac{x}{c_1}\right) \tag{2.1.5}$$

At the boundary $(x=0)$, unless the rope breaks these two displacements must be equal, and therefore

$$T=1+R. \tag{2.1.6}$$

The x-derivatives (i.e. the slopes) of the ropes are:

To the right of the boundary

$$\frac{\partial y}{\partial x}=\frac{T}{c_2}\sin\omega\left(t-\frac{x}{c_1}\right). \tag{2.1.7}$$

and to the left of the boundary

$$\frac{\partial y}{\partial x}=\frac{A}{c_1}\sin\omega\left(t-\frac{x}{c_1}\right)-\frac{AR}{c_1}\sin\omega\left(t+\frac{x}{c_1}\right). \tag{2.1.8}$$

Unless there is a kink in the rope at the boundary, these are equal at $x=0$, and therefore

$$\frac{T}{c_2}=\frac{1}{c_1}-\frac{R}{c_1}. \tag{2.1.9}$$

Combining these with Equation 2.1.6, we obtain

$$T=\frac{2c_2}{c_2+c_1}\text{ and }R=\frac{c_2-c_1}{c_2+c_1}. \tag{2.1.10}$$

We see that if $c_2 < c_1$, R is negative; that is, there is a phase change at reflection. If $c_2 = c_1$ (i.e. if there is only one sort of rope) there is no reflection (because there is no boundary!).

> In the above analysis, we considered a simple sine wave. However, any function, even a nonperiodic function, can be represented by a sum (perhaps an infinite sum) of sinusoidal waves, so the same result will be obtained for any function.

One hopes that energy is conserved, so let's see. The energy in a wave is proportional to the square of its amplitude and, in the case of a vibrating rope, to the mass per unit length. And the rate of transmission of energy is equal to this times the speed. Thus the rate of transmission of energy is proportional to $A^2\mu lc$. However, $c = \sqrt{F/\mu}$, so that the power is proportional to A^2/c. Thus the incident, transmitted and reflected powers are in the ratio

$$1 : \frac{4c_1c_2}{(c_1+c_2)^2} : \frac{(c_1-c_2)^2}{(c_1+c_2)^2}. \tag{2.1.11}$$

We see that the sum of the transmitted and reflected powers is equal to the incident power, and all's well with the world.

2.2: Light Incident Normally at a Boundary

The result described by Equation (3) for the transmitted and reflected amplitudes is an inevitable consequence of the continuity of displacement and gradient of a wave at a boundary, and is not particularly restricted to waves in a rope. It should be equally applicable to electromagnetic waves moving from one medium to another at normal incidence, and indeed it is verified by measurement. Thus, as with the ropes, the amplitudes of the incident, transmitted and reflected waves are in the ratio

$$1:\frac{2c_2}{c_2+c_1}:\frac{c_2-c_1}{c_2+c_1}. \tag{2.2.1}$$

One hopes that energy is conserved, so let's see. The energy stored per unit volume in an electric field in an isotropic medium is $\frac{1}{2}\epsilon E^2$. The rate of transmission of energy per unit area (i.e. the flux density) is this times the speed of propagation. But $\epsilon=\frac{1}{\mu_0 c^2}$.

(We suppose in the present context that both media are nonmagnetic, so both have permeability μ_0.) Thus we see that the rate of propagation of energy per unit area is proportional to the square of the amplitude and inversely proportional to the speed.

Thus the incident, transmitted and reflected powers are in the ratio

$$1:\frac{4c_1c_2}{(c_1+c_2)^2}:\frac{(c_1-c_2)^2}{(c_1+c_2)^2}. \tag{2.2.2}$$

As with the two ropes, the sum of the transmitted and reflected flux densities is equal to the incident flux density, and, once again, all's well with the world.

It may at first glance be surprising that the rate of transmission of energy is *inversely* proportional to the speed. In the case of the ropes, the "slow" rope has a larger mass per unit length. In the case of the electromagnetic field, the "slow" medium has a larger permittivity, so the electric field is having to work the harder.

The speed of light in a medium is inversely proportional to the refractive index, so the amplitude ratios can be expressed as

$$1:\frac{2n_1}{n_1+n_2}:\frac{n_1-n_2}{n_1+n_2}. \tag{2.2.3}$$

We see that there is a phase change on reflection from an optically denser medium.

The flux density ratios can be written as

$$1:\frac{4n_1n_2}{(n_2+n_1)^2}:\frac{(n_2-n_1)^2}{(n_2+n_1)^2}. \tag{2.2.4}$$

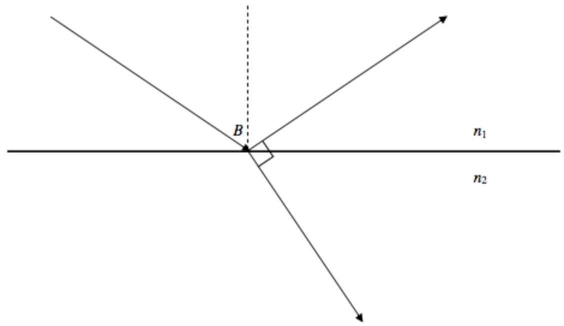
If light is going from air $(n_1=1)$ to glass $(n_2=1.5)$, the transmitted amplitude will be 80 percent of the incident amplitude, and the reflected amplitude will be 20 percent of the incident amplitude. The transmitted flux density will be 96 percent of the incident flux density, and the reflected flux density will be 4 percent of the incident flux density.

If $n_2=n_1$ there will be no reflection at the boundary; in effect there is no boundary. The larva of the midge *Chaoborus*, known as the Phantom Midge, is an aquatic creature whose body has a refractive index equal to the refractive index of water. The picture below shows a photograph of one of them in the water:

Larva of the Phantom Midge, Chaoborus sp. (If you don't believe me, look it up on the Web.)

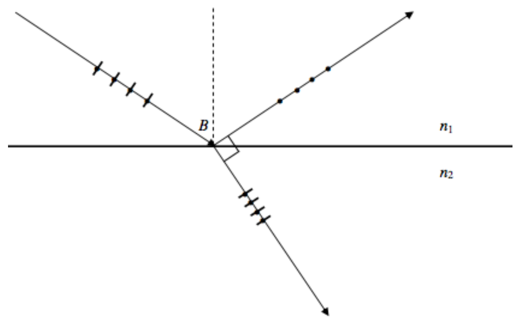
2.3: Light Incident at the Brewster Angle

If a ray of light is incident at an interface between two media in such a manner that the reflected and transmitted rays are at right angles to each other, the angle of incidence, B, is called the Brewster angle.



A moment's thought will show that, if the refractive indices are n_1 and n_2. $\tan B = n_2/n_1$. For example, at an air $(n_1 = 1)$ to glass $(n_2 = 1.5)$ interface the Brewster angle is 56 degrees.

If a ray of unpolarized light is incident at the Brewster angle, the reflected ray is totally plane-polarized. The is no component of the oscillating electric field that is in the plane



of the paper and at right angles to the direction of propagation of the reflected ray. The transmitted ray, having lost some of the component of the electric field at right angles to the plane of the paper (i.e. the dots) is partially plane polarized.

2.4: Electric and Magnetic Fields at a Boundary

We next want to discuss the reflection and transmission for an arbitrary angle of incidence. Before we can do this it is well to remind ourselves (and this is just a reminder - we don't go into the theory and definitions here) from electromagnetic theory how electric and magnetic fields behave at a boundary between two media.

If an electric field is incident normally at the boundary between two media, E is larger in the medium with the smaller permittivity, whereas D is continuous. Likewise, if a magnetic field is incident normally at the boundary between two media, H is smaller in the medium with the higher permeability, whereas B is continuous.

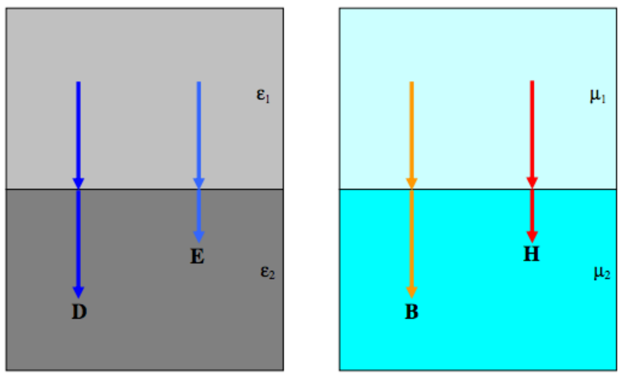
That is:

D_{perp} and B_{perp} are continuous across a boundary.

E_{perp} is inversely proportional to ϵ.

H_{perp} is inversely proportional to μ.

See the drawing below.

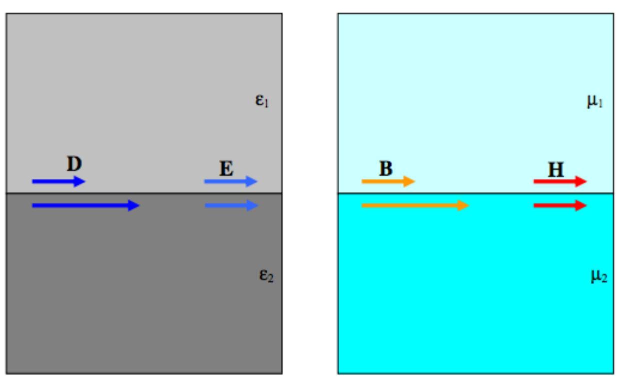


For fields parallel to a boundary, however, the situation is:

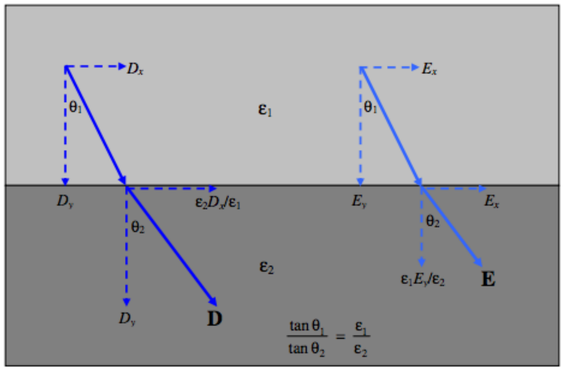
E_{tang} and H_{tang} are continuous across a boundary.

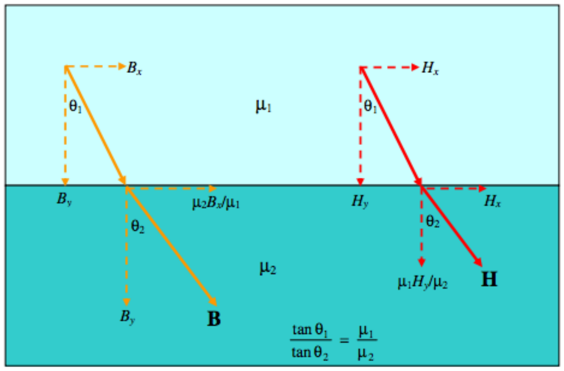
D_{tang} is proportional to ϵ.

B_{tang} is proportional to μ.



These things are assumed known from courses in electromagnetism. It may be asked what happens if a field is neither perpendicular to nor tangential to a boundary. We do not especially need to know that in discussing reflection of light at a boundary, because we shall be resolving any fields into their perpendicular and tangential components, but it is a reasonable question to ask, so for completeness the answers are given in the drawings below.


D_x
θ_1
ε_1
E_x
θ_1
D_y
$\varepsilon_2 D_x/\varepsilon_1$
E_y
E_x
θ_2
θ_2
ε_2
$\varepsilon_1 E_y/\varepsilon_2$
$\mathbf{E}$
D_y
$\mathbf{D}$
$\frac{\tan\theta_1}{\tan\theta_2} = \frac{\varepsilon_1}{\varepsilon_2}$


B_x
θ_1
μ_1
H_x
θ_1
B_y
$\mu_2 B_x/\mu_1$
H_y
H_x
θ_2
θ_2
μ_2
$\mu_1 H_y/\mu_2$
$\mathbf{H}$
B_y
$\mathbf{B}$
$\frac{\tan\theta_1}{\tan\theta_2} = \frac{\mu_1}{\mu_2}$

2.5: Impedance

We need to remind ourselves of one other thing from electromagnetic theory before we can proceed, namely the meaning of *impedance* in the context of electromagnetic wave propagation. The impedance Z is merely the ratio E/H of the electric to the magnetic field. The SI units of E and H are V/m and A/m respectively, so the SI units of Z are V/A, or ohms, Ω. We are now going to see if we can express the impedance in terms of the permittivity and permeability of the medium in which an electromagnetic wave is travelling.

Maxwell's equations are

$$\nabla\cdot\mathbf{D}=\rho \tag{2.5.1}$$

$$\nabla\cdot\mathbf{B}=0. \tag{2.5.2}$$

$$\nabla\times\mathbf{H}=\dot{\mathbf{D}}+\mathbf{J}. \tag{2.5.3}$$

$$\nabla\times\mathbf{E}=-\dot{\mathbf{B}}. \tag{2.5.4}$$

In an isotropic, homogeneous, nonconducting, uncharged medium (such as glass, for example), the equations become:

$$\nabla\cdot\mathbf{E}=0 \tag{2.5.5}$$

$$\nabla\cdot\mathbf{H}=0 \tag{2.5.6}$$

$$\nabla\times\mathbf{H}=\epsilon\dot{\mathbf{E}}. \tag{2.5.7}$$

$$\nabla\times\mathbf{H}=-\mu\dot{\mathbf{H}}. \tag{2.5.8}$$

If you eliminate $\mathbf{H}$ from these equations, you get

$$\nabla^2\mathbf{E}=\epsilon\mu\ddot{\mathbf{E}}, \tag{2.5.9}$$

which describes an electric wave of speed

$$\frac{1}{\sqrt{\epsilon\mu}}. \tag{2.5.10}$$

In free space, this becomes

$$\frac{1}{\sqrt{\epsilon_0\mu_0}} \tag{2.5.11}$$

which is $2.998\times10^8\ m\ s^{-1}$.

The ratio of the speeds in two media is

$$\frac{v_1}{v_2}=\frac{n_2}{n_1}=\sqrt{\frac{\epsilon_2\mu_2}{\epsilon_1\mu_1}}, \tag{2.5.12}$$

and if, as is often the case, the two permeabilities are equal (to μ_0), then

$$\frac{v_1}{v_2}=\frac{n_2}{n_1}=\sqrt{\frac{\epsilon_2}{\epsilon_1}}. \tag{2.5.13}$$

In particular, if you compare one medium with a vacuum, you get: $n=\sqrt{\frac{\epsilon}{\epsilon_0}}$.

Light is a high-frequency electromagnetic wave. When a dielectric medium is subject to a high frequency field, the polarization (and hence D) cannot keep up with the electric field E. D lags behind E. This can be described mathematically by ascribing a complex value to the permittivity. The amount of lag depends, unsurprisingly, on the frequency - i.e. on the color - and so the permittivity and hence the refractive index depends on the wavelength of the light. This is *dispersion*.

If instead you eliminate $\mathbf{E}$ from Maxwell's equations, you get

$$\nabla^2\mathbf{H}=\epsilon\mu\ddot{\mathbf{H}}. \tag{2.5.14}$$

This is a magnetic wave of the same speed.

If you eliminate the time between 2.5.9 and 2.5.14, you find that $\frac{E}{H}=\sqrt{\frac{\mu}{\epsilon}}$, which, in free space, has the value $\sqrt{\frac{\mu_0}{\epsilon_0}}=377\Omega$, which is the impedance of free space. In an appropriate context I may use the symbol Z_0 to denote the impedance of free space, and the symbol Z to denote the impedance of some other medium.

The ratio of the impedances in two media is

$$\frac{Z_1}{Z_2}=\sqrt{\frac{\epsilon_2\mu_1}{\epsilon_1\mu_2}}, \tag{2.5.15}$$

and if, as is often the case, the two permeabilities are equal (to μ_0), then

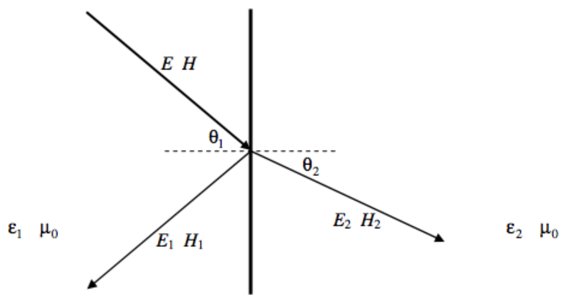
$$\frac{Z_1}{Z_2}=\sqrt{\frac{\epsilon_2}{\epsilon_1}}=\frac{n_2}{n_1}=\frac{v_1}{v_2}. \tag{2.5.16}$$

We shall be using this result in what follows.

2.6: Incidence at an Arbitrary Angle.

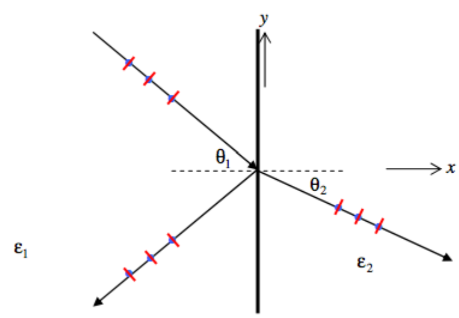
In Section 4 (Incidence at the Brewster Angle) it became clear that the reflection of light polarized in the plane of incidence was different from the reflection of plane polarized light polarized at right angles to the plane incidence. Therefore it makes sense, in this section, to consider the two planes of polarization separately. I shall suppose that both media are isotropic (i.e. not birefringent).

In the following discussion, we'll suppose that light is travelling from a medium of permittivity ϵ_1 to a medium of greater permittivity ϵ_2. Both permeabilities are equal, and close to μ_0. The electric and magnetic fields of the *incident* wave will be denoted by E and H. The electric and magnetic fields of the *reflected* wave will be denoted by E_1 and H_1. The electric and magnetic fields of the transmitted wave will be denoted by E_2 and H_2. (And in case you are wondering, by H I mean H, and by B I mean B.)



Coordinates: x to the right. y upwards z towards you.

We'll start by supposing that the **incident light is plane polarized with the electric field perpendicular (senkrecht) to the plane of incidence.** That is, the electric field has only a z-component. The oscillating electric field E is indicated by blue dots, and the magnetic field H by red dashes in the drawing below.



The boundary conditions are: For the tangential (z) component of $\mathbf{E}$

$$E+E_1=E_2 \tag{9}$$

For the tangential (y) component of $\mathbf{H}$

$$(H-H_1)\cos\theta_1=H_2\cos\theta_2. \tag{10}$$

That is,

$$\frac{(E-E_1)}{Z_1}\cos\theta_1=\frac{E_2}{Z_2}\cos\theta_2, \tag{11}$$

or

$$n_1(E-E_1)\cos\theta_1=n_2E_2\cos\theta_2. \tag{12}$$

Eliminate E_2 between Equations 9 and 12:

Reflected amplitude:

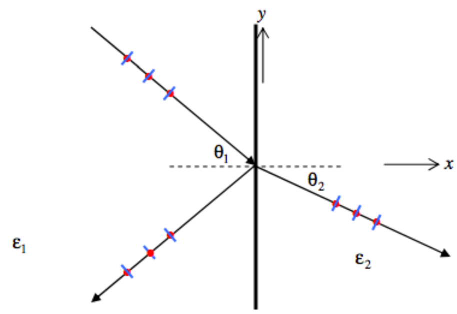
$$\frac{E_1}{E} = \frac{n_1 \cos\theta_1 - n_2 \cos\theta_2}{n_1 \cos\theta_1 + n_2 \cos\theta_2}. \tag{13}$$

Use Equation 9:

Transmitted amplitude

$$\frac{E_2}{E} = \frac{2n_2 \cos\theta_1}{n_1 \cos\theta_1 + n_2 \cos\theta_2}. \tag{14}$$

Now we'll supposing that the **incident light is plane polarized with the electric field parallel to the plane of incidence.** This, it is the magnetic field that has only a z-component. The oscillating electric field E is indicated by blue dashes, and the magnetic field H by red dots in the drawing below.



The boundary conditions are:

For the tangential (z) component of $\mathbf{H}$

$$H + H_1 = H_2$$

That is:

$$\frac{E + E_1}{Z_1} = \frac{E_2}{Z_2} \text{ or } n_1(E + E_1) = n_2 E_2. \tag{15}$$

For the tangential (y) component of $\mathbf{E}$

$$(E - E_1)\cos\theta_1 = E_2 \cos\theta_2. \tag{16}$$

Eliminate E_2 between Equations 15 and 16:

Reflected amplitude:

$$\frac{E_1}{E} = \frac{n_2 \cos\theta_1 - n_1 \cos\theta_2}{n_2 \cos\theta_1 + n_1 \cos\theta_2}. \tag{17}$$

Use Equation 15:

Transmitted amplitude:

$$\frac{E_2}{E} = \frac{2n_1 \cos\theta_1}{n_2 \cos\theta_1 + n_1 \cos\theta_2}. \tag{18}$$

These are the *Fresnel Equations*, gathered together below:

Perpendicular (Senkrecht)

Reflected amplitude:

$$\frac{E_1}{E} = \frac{n_2 \cos\theta_1 - n_2 \cos\theta_2}{n_2 \cos\theta_1 + n_2 \cos\theta_2}.$$

Transmitted amplitude:

$$\frac{E_2}{E} = \frac{2n_1 \cos\theta_1}{n_1 \cos\theta_1 + n_2 \cos\theta_2}.$$

Parallel

Reflected amplitude:

$$\frac{E_1}{E} = \frac{n_2 \cos\theta_1 - n_1 \cos\theta_2}{n_2 \cos\theta_1 + n_1 \cos\theta_2}.$$

Transmitted amplitude:

$$\frac{E_2}{E} = \frac{2n_1 \cos\theta_1}{n_2 \cos\theta_1 + n_1 \cos\theta_2}.$$

They evidently depend only on the *ratio* of the refractive indices (i.e. the refractive index of one medium relative to that of the other). If we write $n = n_2/n_1$, the equations become

Perpendicular (Senkrecht)

Reflected amplitude:

$$\frac{E_1}{E} = \frac{\cos\theta_1 - n\cos\theta_2}{\cos\theta_1 + n\cos\theta_2}.$$

Transmitted amplitude:

$$\frac{E_2}{E} = \frac{2\cos\theta_1}{\cos\theta_1 + n\cos\theta_2}.$$

Parallel

Reflected amplitude:

$$\frac{E_1}{E} = \frac{n\cos\theta_1 - \cos\theta_2}{n\cos\theta_1 + \cos\theta_2}.$$

Transmitted amplitude:

$$\frac{E_2}{E} = \frac{2\cos\theta_1}{n\cos\theta_1 + \cos\theta_2}.$$

For normal incidence, the ratios for the senkrecht component become $\frac{1-n}{1+n}$ and $\frac{2}{1+n}$ as expected. The ratios for the parallel component, however, become $\frac{n-1}{n+1}$ and $\frac{2}{1+n}$, apparently predicting no phase change at external reflection for the parallel component. This is only apparent, however, and the explanation for the apparent anomaly is given on pp. 20-24.

It will be noted that n, θ_1, θ_2 are also related by Snell's law: $\sin\theta_1 = n\sin\theta_2$, so that we can eliminate n from Fresnel's equations in order to express them in terms of the angles of incidence and refraction only. If this is done we obtain:

Perpendicular (Senkrecht):

Reflected amplitude:

$$\frac{E_1}{E} = -\frac{\sin(\theta_1 - \theta_2)}{\sin(\theta_1 + \theta_2)}.$$

Transmitted amplitude:

$$\frac{E_2}{E} = \frac{2\sin\theta_2\cos\theta_1}{\sin(\theta_1+\theta_2)} = \frac{2}{1+\frac{\tan\theta_1}{\tan\theta_2}}.$$

Parallel

Reflected amplitude:

$$\frac{E_1}{E} = -\frac{\tan(\theta_1 - \theta_2)}{\tan(\theta_1 + \theta_2)}.$$

Transmitted amplitude:

$$\frac{E_2}{E} = \frac{2\sin\theta_2\cos\theta_1}{\sin(\theta_1+\theta_2)\cos(\theta_1 - \theta_2)}.$$

In perhaps the most useful form of all, we could eliminate θ_2 from the Fresnel equations and hence obtain them as functions of θ_1 and n only. This will enable us easily to calculate the reflected and transmitted amplitudes in terms of the angle of incidence. Thus:

Perpendicular (Senkrecht)

Reflected amplitude:

$$\frac{E_1}{E} = -\frac{(n^2 - \sin^2\theta_1)^{\frac{1}{2}} - \cos\theta_1}{(n^2 - \sin^2\theta_1)^{\frac{1}{2}} + \cos\theta_1}$$

Transmitted amplitude:

$$\frac{E_2}{E} = -\frac{2\cos\theta_1}{(n^2 - \sin^2\theta_1)^{\frac{1}{2}} + \cos\theta_1}\ .$$

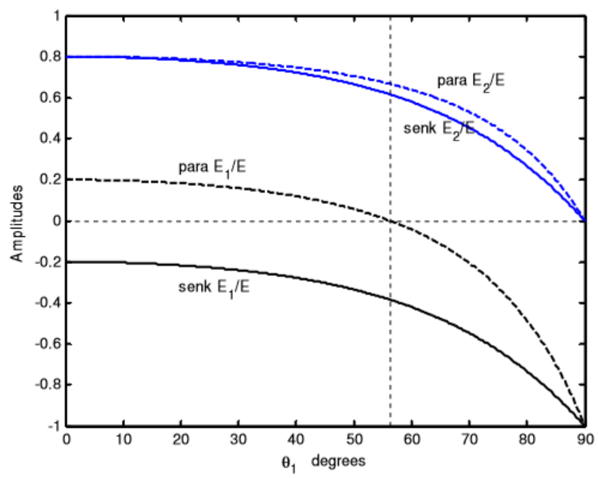
Parallel

Reflected amplitude:

$$\frac{E_1}{E} = \frac{n^2\cos\theta_1 - (n^2 - \sin^2\theta_1)^{\frac{1}{2}}}{n^2\cos\theta_1 + (n^2 - \sin^2\theta_1)^{\frac{1}{2}}}\ .$$

Transmitted amplitude:

$$\frac{E_2}{E} = \frac{2n\cos\theta_1}{n^2\cos\theta_1 + (n^2 - \sin^2\theta_1)^{\frac{1}{2}}}\ .$$



Black curves are the amplitudes of the reflected waves.

Blue curves are the amplitudes of the transmitted waves.

Continuous curves are for senkrecht (perpendicular) waves.

Dashed curves are for parallel waves.

Negative values show where there is a 180º phase shift on reflection. Notice that, at the Brewster angle (about 56º), none of the parallel component is reflected.

At 90º (grazing incidence) no light is transmitted; it is all reflected, but with a phase change (negative amplitude).

Energy considerations

Recall that for the parallel component, the incident, reflected and transmitted amplitudes are in the ratio

$$E \quad : \quad E_1 \quad : \quad E_2 = 1 \quad : \quad -\frac{(n^2 - \sin^2\theta_1)^{\frac{1}{2}} - \cos\theta_1}{(n^2 - \sin^2\theta_1)^{\frac{1}{2}} + \cos\theta_1} \quad : \quad \frac{2\cos\theta_1}{(n^2 - \sin^2\theta_1)^{\frac{1}{2}} + \cos\theta_1}$$

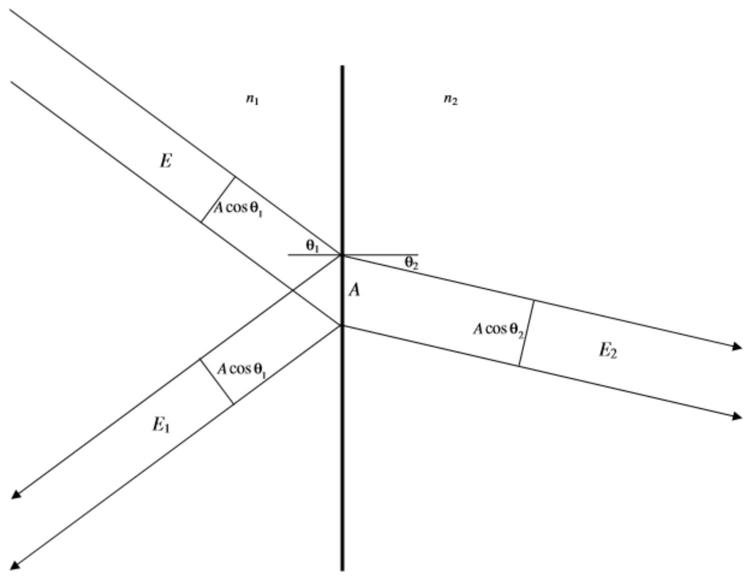
and for the senkrecht component they are in the ratio

$$E : E_1 : E_2 = 1 : \frac{n^2\cos\theta_1 - (n^2 - \sin^2\theta_1)^{\frac{1}{2}}}{n^2\cos\theta_1 + (n^2-\sin^2\theta_1)^{\frac{1}{2}}} : \frac{2\cos\theta_1}{n^2\cos\theta_1 + (n^2-\sin^2\theta_1)^{\frac{1}{2}}}$$

(Here $n = \frac{n_2}{n_1}$.)

Suppose that the incident light strikes the interface in an area A. That means that the incident and reflected light are each in beams of cross-sectional area $A\cos\theta_1$, and the transmitted light is in a beam of cross-sectional area A_2. We are going to calculate the ratio $P : P_1 : P_2$ of the rate of transmission of energy (power) in each beam; and if we do our algebra correctly, we should find that $P_1 + P_2 = P$.

Recall that the energy per unit volume in an electric field is proportional to ϵE^2, where ϵ, the permittivity, is proportional to the square of the refractive index. The power transmitted by each beam is proportional to the energy per unit volume, times the speed of transmission (which is inversely proportional to the refractive index), and to the crosssection area of the beam.



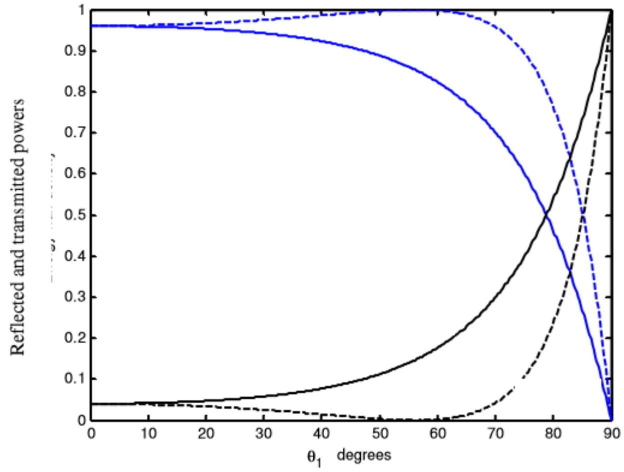
Therefore, for the parallel component and for the senkrecht component,

$$P : P_1 : P_2 = n_1E^2\cos\theta_1 : n_1E_1^2\cos\theta_1 : n_2E_2^2\cos\theta_2$$

Normalizing this expression so that $P = 1$, we obtain

$$P : P_1 : P_2 = 1 : \left(\frac{E_1}{E}\right)^2 : n\left(\frac{E_2}{E}\right)^2\frac{\cos\theta_2}{\cos\theta_1} \quad .$$

These are shown below for n = 1.5, and indeed $P_1 + P_2 = P$ for each component, and energy is conserved.



Notice that at grazing incidence we have total external reflection.

Black curves are the reflection coefficients of the reflected waves.

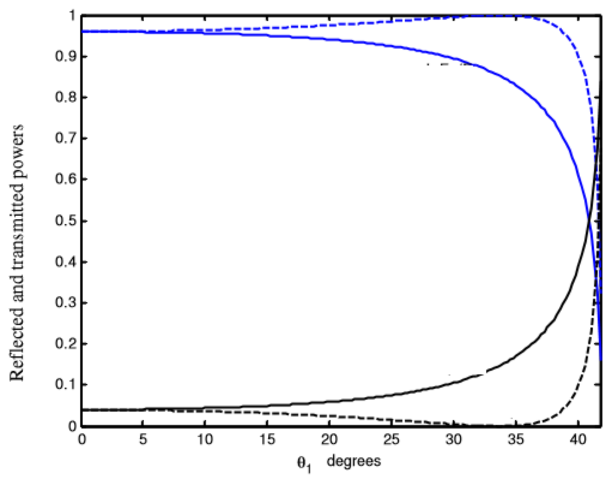
Blue curves are the transmission coefficients of the transmitted waves.

Continuous curves are for senkrecht (perpendicular) waves.

Dashed curves are for parallel waves.

At the Brewster angle no parallel waves are reflected.

For light going from $n_1 = 1.5$ to $n_2 = 1$:



Black curves are the reflection coefficients of the reflected waves.

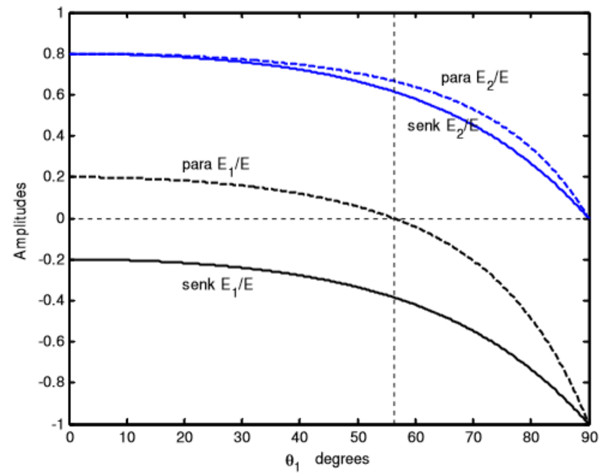
Blue curves are the transmission coefficients of the transmitted waves.

Continuous curves are for senkrecht (perpendicular) waves.

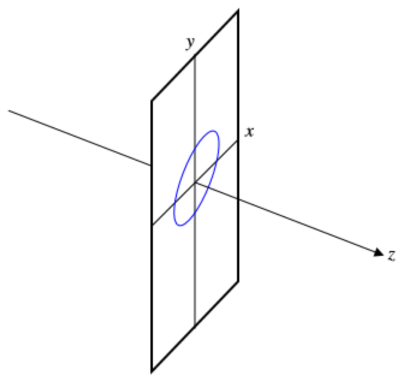
Dashed curves are for parallel waves.

For angles of incidence greater than 42 degrees (the critical angle for total internal reflection) all light is reflected. The phase of this totally reflected light is something that we have not yet discussed.

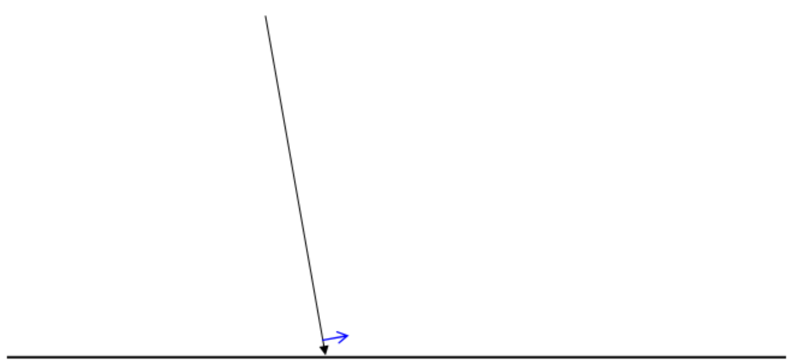
I return now to external reflection and to the graphs, repeated below, which show the reflected and transmitted amplitudes of the parallel and senkrecht components. The blue curves show the transmitted amplitudes, and there is no problem with them. The amplitudes are all positive, meaning that the transmitted waves have no phase change at the boundary. My students pointed out an apparent paradox with the dashed black curve, which is the reflected amplitude of the parallel component. It is positive, indicating (apparently) no phase change, even at normal incidence - and yet we know that there must be a phase change for reflected light at normal incidence. My students demanded (and rightly so) an explanation. The apparent anomaly was also noted on p.15. Following the diagram is the solution that I offer.



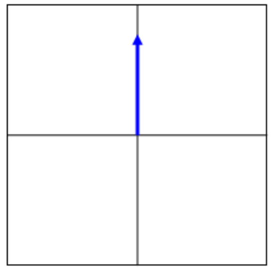
When we describe the state of polarization of light, whether, linear, circular or elliptical, we refer for convenience and of necessity to a coordinate system in which the z-axis is in the direction of the ray, and the xy-plane is perpendicular to it. The observer is supposed to be on the positive z-axis looking towards the source of light:



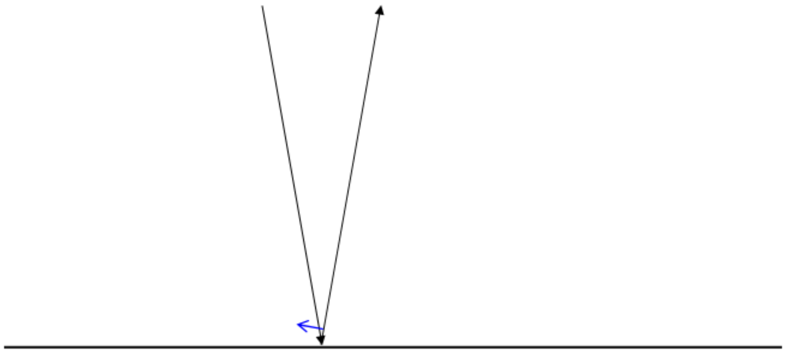
Consider a ray coming down at a steep angle to a water surface. Suppose at some instant of time the electric vector just above the surface is as shown by the little blue arrow below.



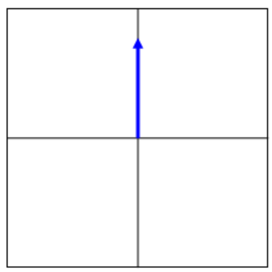
What does our observer (who is underneath the water) see, and how does he describe the state of polarization? This is what he sees:



Now the light is reflected, the observer changes his position, and he looks down on the water from above.



And this is what he sees:



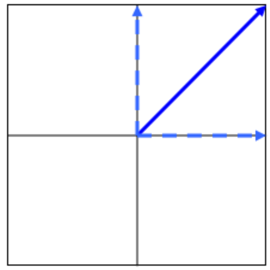
And so there has been no phase change.

Or has there?

One might say that there has been a phase change, but it looks as though there hasn't been. In effect, before and after, we are referring the situation to two reference frames, one of which is the mirror image of the other.

You will see that this apparent paradox does not arise with the senkrecht component.

We have hitherto considered the reflection and transmission of light that was initially plane polarized either parallel to the plane of incidence, or perpendicular (senkrecht) to it. Suppose that the incident light is plane polarized in a direction 45º to the parallel and senkrecht planes. We can resolve it into parallel and senkrecht components, each of amplitude $\frac{E}{\sqrt{2}}$. We suppose that the angle of incidence is θ_1, and the angle of refraction, which is easily calculated from Snell's Law, is θ_2. And so there has been no phase change.



After reflection, the amplitudes of the parallel component will be $\frac{E}{\sqrt{2}} \times \frac{\tan(\theta_1 - \theta_2)}{\tan(\theta_1 + \theta_2)}$.

and the amplitude of the senkrecht component will be $-\frac{E}{\sqrt{2}} \times \frac{\sin(\theta_1 - \theta_2)}{\sin(\theta_1 + \theta_2)}$.

From these we can calculate the resultant amplitude of the reflected wave as well as its polarization direction (which is quite different from the plane of polarization of the incident wave.)

The transmitted light will have a parallel component of amplitude

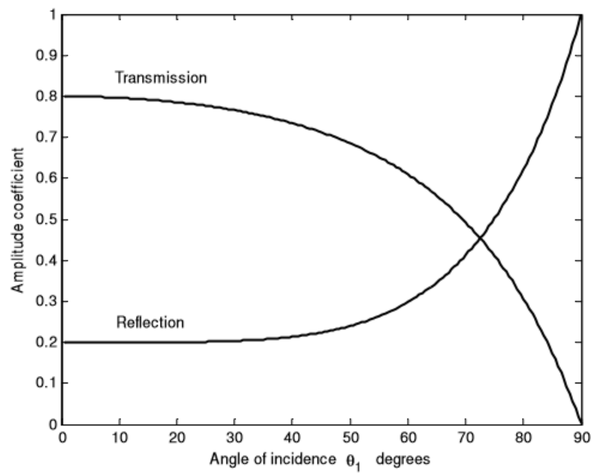
$$\frac{E}{\sqrt{2}} \times \frac{2\sin\theta_2\cos\theta_1}{\sin(\theta_1 + \theta_2)\cos(\theta_1 - \theta_2)}$$

and a senkrecht component of amplitude

$$\frac{E}{\sqrt{2}} \times \frac{2}{1 + \frac{\tan\theta_1}{\tan\theta_2}} .$$

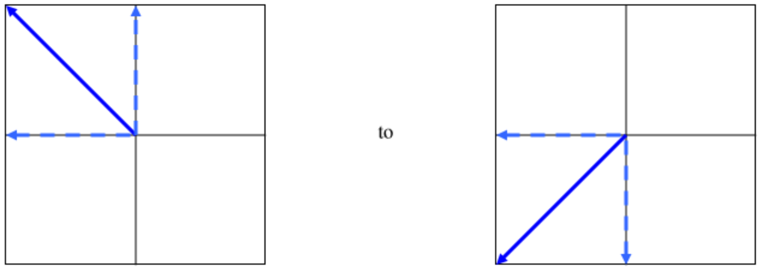
From these we can calculate the resultant amplitude of the transmitted wave as well as its polarization direction (which, as for the reflected wave, is in a different plane from the plane of polarization of the incident wave.)

We show here the magnitudes (without regard to sign) of the amplitude reflection and transmission coefficients, and the polarization directions for the reflected and transmitted wave, as a function of angle of incidence θ_1, assuming $n = \frac{n_2}{n_1} = 1.5$.



At grazing incidence θ_1 = 90º, all the light is reflected. Although it has no particular significance, we note that, for n = 1.5, the reflection and transmission amplitude coefficients are equal (to 0.4544) for an angle of incidence equal to 72º.464. Except for normal and grazing incidence, the reflection and transmission amplitude coefficients do not add exactly to one. While there is a requirement for energy to be conserved, there is no similar requirement for the amplitudes.

As the angle of incidence goes from zero (normal incidence) to 90º (grazing incidence), the plane of polarization of the reflected wave goes from



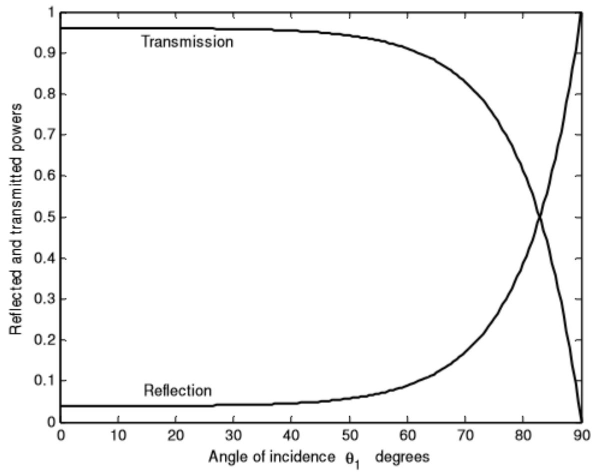
Note that, for normal incidence, the reflected wave has a phase change for the senkrecht component, but (apparently) not for the parallel component, as explained above.

The plane of polarization of the transmitted moves slightly from the initial 45º to 56º.6 (the Brewster angle) at grazing incidence, although this has little significance since no light is transmitted at grazing incidence.

As described on 17-18, if the incident, reflected and transmitted amplitudes are in the ratio , and the corresponding powers are in the ratio , then

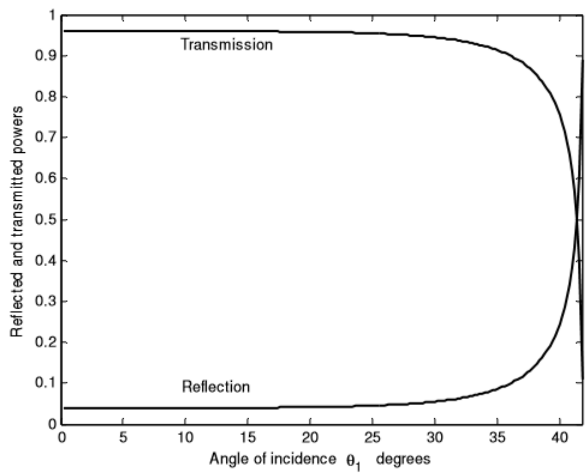
$$P \ : \ P_1 \ : \ P_2 = 1 \ : \ \left(\frac{E_1}{E}\right)^2 \ : \ n\left(\frac{E_2}{E}\right)^2 \frac{\cos\theta_2}{\cos\theta_1}$$

These are shown below for n = 1.5.



Recall that in these calculations, it has been assumed that the incident light is plane polarized at 45º to the parallel and senkrecht planes, so that the parallel and senkrecht amplitude components of the incident light are equal. Completely unpolarized incident light also has equal parallel and senkrecht amplitude components, so that the above graph also shows the reflection and transmission coefficients for unpolarized incident light. For n = 1.5, the reflection and transmission coefficients are equal for an angle of incidence of 82º.82. For any angle of incidence less than 60º, very much more light is transmitted than reflected., but, in the limit as $\theta_1 \to 90^\circ$, all the light is reflected.

We show below the reflection and transmission coefficients of internal reflection for angles of incidence from zero to the critical angle, which, for n = 1.5, is 41º.8. This is achieved merely by replacing 1.5 with $\frac{2}{3}$ in the calculations.



CHAPTER OVERVIEW

3: The Cornu Spiral

Thumbnail: A double-end Euler spiral. The curve continues to converge to the points marked, as t tends to positive or negative infinity. (CC BY-SA 3.0; AdiJapan).

3.1: Cornu's Spiral

If a parallel beam of light from a distant source encounters an obstacle, the shadow of the obstacle is not a simple geometric shadow but is, rather, a diffraction pattern. For example, it is well known that the diffraction pattern formed by a slit looks like the function shown in Figure 1.

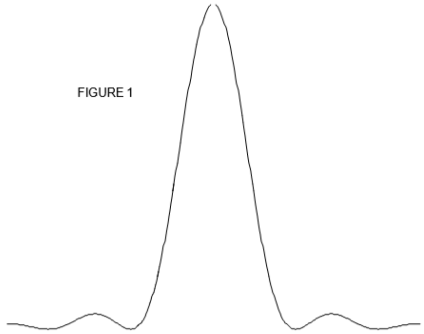


FIGURE 1

Such diffraction is called *Fraunhofer diffraction.*

If, however, the source of light is not distant, but is close to the diffracting obstacle so that the incident waves are not plane waves, the diffraction pattern will look somewhat different. Such diffraction is called *Fresnel diffraction*, and its theory is, unsurprisingly, a little more difficult than the theory for Fraunhofer diffraction.

If the source of light is a point source, so that the incident wavefronts are spherical, the detailed quantitative theory is not at all easy. If the incident wavefronts, however, are cylindrical (say from a linear source) the analysis, which is two dimensional, is a little more tractable. *Cornu's spiral* is a graphical device that enables us to compute and predict the Fresnel diffraction pattern from various simple obstacles.

> "Cornu", by the way, is French for "horned", and can also mean "spiral" - i.e. like the horns of a bighorn sheep or of an ibex. Because of this I wondered, when I first heard about Cornu's spiral, whether it should really be called a "cornu spiral", rather than Cornu's spiral. However, it is correctly named Cornu's spiral after a real nineteenth century French scientist, Marie Alfred Cornu. The mathematical properties of the spiral had been examined by various mathematicians (for example, Euler) before Cornu, but it has acquired the name of Cornu because of its application by Cornu to the theory of Fresnel diffraction.

Let us look, in Figure III.2, at the geometry of a cylindrical wavefront from a linear source at O.

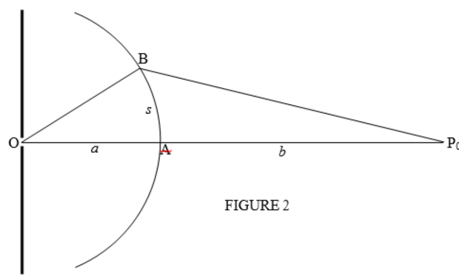


FIGURE 2

Introduce a dimensionless variable v by

$$v=\sqrt{\frac{2(a+b)}{ab\lambda}}\,s. \tag{1}$$

where λ is the wavelength of the light.

Theory shows that the intensity (square of the amplitude) of the radiation received at the point P_0 from the portion AB of arc-length s of the wavefront is proportional to

$$\left(\int_0^v \cos\frac{1}{2}\pi u^2 du\right)^2+\left(\int_0^v \sin\frac{1}{2}\pi u^2 du\right)^2. \tag{3}$$

Here

$$C=\int_0^v \cos\frac{1}{2}\pi u^2 du \quad \text{and} \quad S=\int_0^v \sin\frac{1}{2}\pi u^2 du \tag{4}$$

are the *Fresnel integrals*.

The derivation of Equation (3) may be somewhat heavy-going, and we shall relegate it to Appendix A at the end of this chapter. For the time being, we shall accept Equation (3) as being correct, and we shall see how to use it to construct the Cornu spiral and how the spiral can be used to compute the forms of the shadows produced by various obstacles. In Equations (4), u is just a dummy variable. C and S are functions of v, which is proportional to s. They must be integrated numerically, and I have provided a brief table of them in Appendix B.

The Cornu spiral is a graph of S versus C. Figure III.3 shows such a graph. Better ones exist in the literature, but this one will suffice to show how it is used. However, I shall shortly suggest that, while it is fun to use the spiral, for precise work it is preferable to compute the forms of the shadows numerically rather than graphically. The spiral was useful in the days before high-speed computers, but today one can compute the Fresnel integrals instantaneously, and hence we can compute the forms of the shadow, using the spiral perhaps to guide us. A word of warning, though. The rapid and accurate computation of the Fresnel integrals requires some care in programming, for the integrand changes rapidly with the variable u. In preparing the graphs and tables in this note, I found that Simpson's Rule was inadequate - it worked provided I used a large number of intervals, but this slowed down the computation. I was able to get better and faster results with Gaussian quadrature.

The dimensionless variable v (which is proportional to s - see Equation 1) is measured along the spiral. I have drawn dots on the spiral for every 0.1 increment in v. I haven't labelled the numerical values of v beside the dots, but feel free to do so if you wish. Note that, as $v \to \pm\infty$, C and $S \to \pm\frac{1}{2}$. The intensity at P is proportional to the square of the distance between these two limiting points $(\frac{1}{2}, \frac{1}{2})$ and $(-\frac{1}{2}, -\frac{1}{2})$. The distance between these points is $\sqrt{2}$, and the square of the distance is 2. Thus, with no obstacle between the source and the point P, the amplitude of the radiation at P is $\sqrt{2}$ (arbitrary units), and the intensity at P is 2 (arbitrary units).

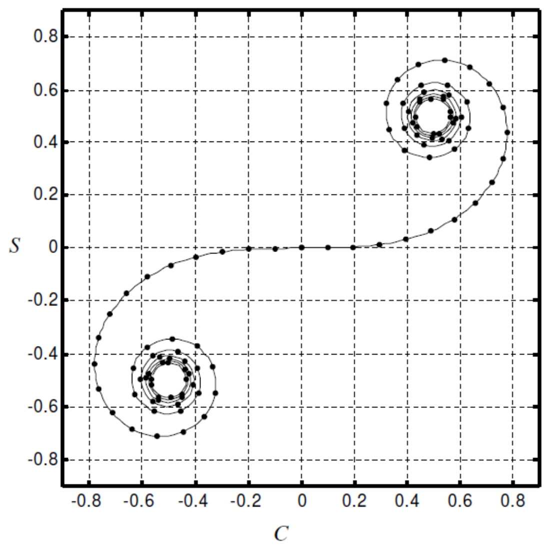
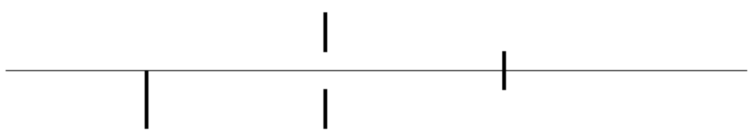


FIGURE 3

In what follows, we are going to put three obstacles in front of the light source and we are going to compute the Fresnel diffraction pattern (i.e. the structure of the shadow.) The three obstacles will be a single straight edge, a slit between two straight edges, and an opaque strip:



All the time recall that the distance s along the wavefront is linearly related to the distance v along the Cornu spiral.

We'll start with the **single straight edge.**

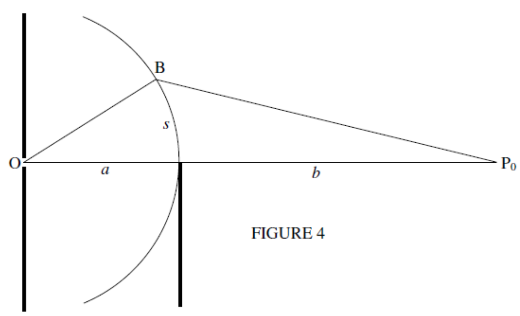


FIGURE 4

At the point P_0 we see all of the upper part of the wavefront. That is, we see along the Cornu spiral from $v = 0$ to where the spiral converges at $C = \frac{1}{2}$, $S = \frac{1}{2}$. The amplitude of the radiation at P is proportional to the distance between these two points, which is $\frac{1}{\sqrt{2}}$, and the intensity at P is proportional to the square of this, which is $\frac{1}{2}$, which is one quarter of the intensity when the light was unobstructed by any obstacle.

Now let us see what the intensity is at a point P some way above the axis (Figure 5).

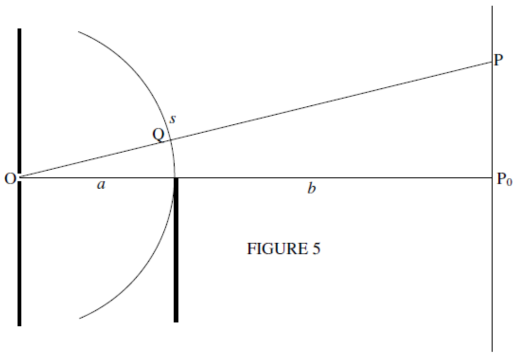


FIGURE 5

The distance s is must now be measured not from the edge of the obstacle, but from the point Q. At P we see more of the wavefront than we did at P_0. We see all of s above Q, as well as some negative values of s below Q. The amplitude at P, then, corresponds to the length of the chord in Figure III.6, in which the negative v is related to the negative s by Equation (1). We see that, as we move P upwards in Figure 5, We take in more and more negative s, and more and more negative v in the Cornu spiral.

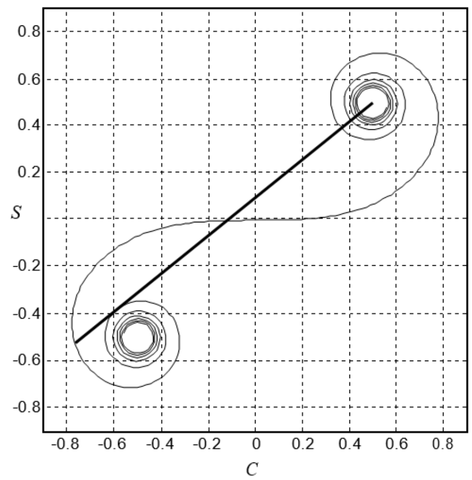


FIGURE 6

Thus, as P moves upwards in Figure 5, we keep the upper end of the chord in Figure 6 fixed and we move the lower end around the spiral. The length of the chord is proportional to the amplitude of the light received at P, and its square is proportional to its intensity.

We can use a ruler and the spiral to determine the intensity as a function of v and hence of s, and this would have been an appropriate procedure before the advent of high-speed computers. To delineate the intensity as a function of v by computer, as we move along the spiral, for each value of v we calculate C and S and then calculate the intensity from the square of the length of the chord, which is

$$\left(\frac{1}{2}-C\right)^{2}+\left(\frac{1}{2}-S\right)^{2}. \tag{3.1.1}$$

This is what I did for Figure 7 except that I divided this expression by two, so that an intensity of one represents the intensity at P_0 in the absence of any obstacle. A reader who tries to duplicate this will soon appreciate the value of programming a fast and accurate method of evaluating the Fresnel integrals.

The portion to the right of $v = 0$ is within the geometric shadow.

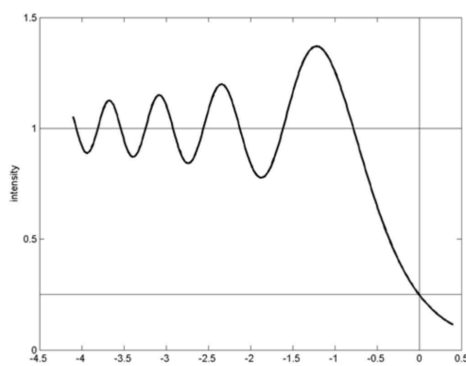


FIGURE 7

Now we'll look at what happens when the obstacle is a **slit between two straight edges**. We'll suppose that the width of the slit is Δs, corresponding to a distance along the spiral $\Delta v = \sqrt{\frac{2(a+b)}{ab\lambda}}\,\Delta s$. In the calculations that I have done below, I have taken Δv to be 4.0. The point P (see Figure 8) is receiving energy from the part of the wavefront between $s - \Delta s$ and s, corresponding to a chord on the spiral spanning a distance Δv along the spiral. As the point P moves upward along the screen, so the chord slides along the spiral (see Figure 9), keeping Δv constant

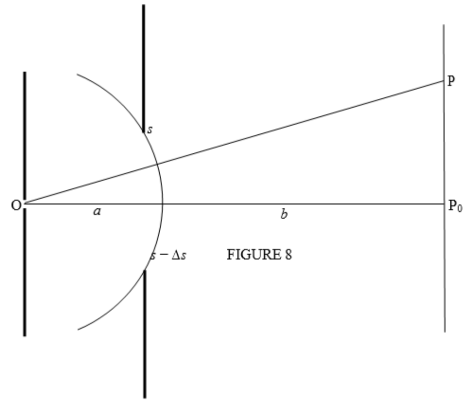


FIGURE 8

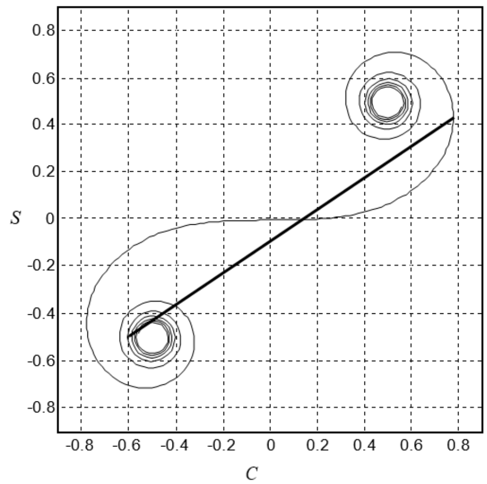


FIGURE 9

For each position of the chord, we need to calculate the Fresnel integrals C_u, S_u of the upper end of the chord and the Fresnel integrals C_l,S_l of the lower end of the chord and then calculate the square of the length of the chord (and then divide by two, so that an intensity of 1 is the intensity when the light is unobstructed). That is, we calculate

$$\frac{1}{2}\left[(C_u - C_l)^2 + (S_u - S_1)^2\right]$$

I got the result shown in Figure 10, using a slit width corresponding to $\Delta v = 4$.

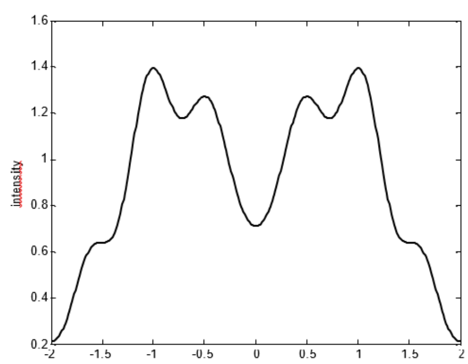


FIGURE 10

The positions $v = \pm 2$ correspond to the edge of he geometric shadow. The intensity has not fallen to zero there - some light spills over into the geometric shadow.

The details of the diffraction pattern are very sensitive to the value of Δv. That is to say to Δs. That is to say to the slit width. Figure 11, for example, shows the same calculation but for $\Delta v = 3.9$ rather than 0.4.

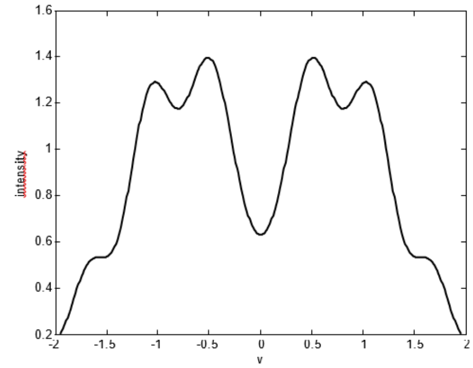


FIGURE 11

As the slit width is changed, sometimes there will be a dip at $v = 0$, and sometimes a maximum. Generally, a large Δv results in a more complicated pattern, and a smaller Δv results in a simpler pattern. As Δv becomes smaller, the pattern approaches the familiar Fraunhofer diffraction pattern for a slit, as in Figure 1.

Now let us choose as the obstacle a single opaque strip. I'll make the width of the strip equal to the width of the slit in the example of Figure 10, which corresponds to a distance along the spiral of $\Delta v = 4$. Instead of sliding the chord of Figure III.10 along the spiral, we have to slide the two complementary chords shown in Figure 12. We have to calculate the same Fresnel integrals C_u, S_u, C_l, S_l as before, but this time the resultant of the two, added as vectors, and normalized so that the unobstructed intensity is 1, is $\frac{1}{2}\left[(1-C_u+C_l)^2 + (1-S_u+S_1)^2\right]$. I obtain the result shown in Figure 13.

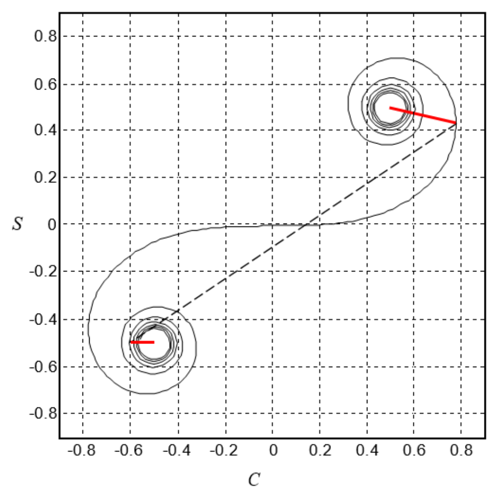
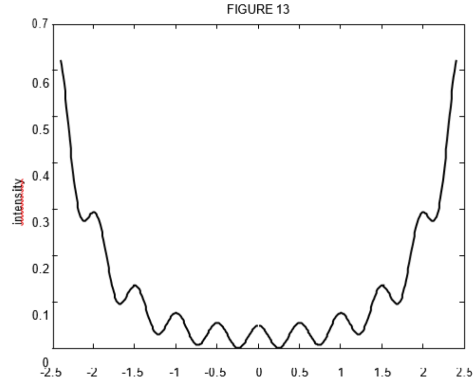


FIGURE 12

FIGURE 13



The key to doing these calculations successfully is to have an efficient, fast and accurate routine for calculating the Fresnel integrals. In each of these graphs each of the Fresnel integrals (sine and cosine) was calculated by numerical integration about 400 times. I found Simpson's Rule was inadequate, so I used Gaussian Quadrature.

APPENDIX A

In the above notes, I have described what the Fresnel integrals and the Cornu spiral are, and how to use them in some simple cases. I have not shown why it is that the diffraction patterns can be generated by the Fresnel integrals, or how to derive Equation (3). I hope sometime to derive this and explain the rationale behind the theory in this Appendix at some later date. I'm afraid I can't say when I expect to get round to doing this. Could be this year, next year, sometime, never...

APPENDIX B

The Fresnel Integrals

v	C	S
0.10	0.1000	0.0005
0.20	0.1999	0.0042
0.30	0.2994	0.0141
0.40	0.3975	0.0334
0.50	0.4923	0.0647
0.60	0.5811	0.1105
0.70	0.6597	0.1721
0.80	0.7228	0.2493
0.90	0.7648	0.3398
1.00	0.7799	0.4383
1.10	0.7638	0.5365
1.20	0.7154	0.6234
1.30	0.6385	0.6863
1.40	0.5431	0.7135
1.50	0.4453	0.6975
1.60	0.3655	0.6389
1.70	0.3238	0.5492
1.80	0.3336	0.4509
1.90	0.3945	0.3733
2.00	0.4883	0.3434
2.10	0.5816	0.3743
2.20	0.6363	0.4557
2.30	0.6266	0.5532
2.40	0.5550	0.6197
2.50	0.4574	0.6192
2.60	0.3889	0.5500
2.70	0.3925	0.4529
2.80	0.4675	0.3915
2.90	0.5624	0.4101
3.00	0.6057	0.4963
3.10	0.5616	0.5818
3.20	0.4663	0.5933
3.30	0.4057	0.5193
3.40	0.4385	0.4296
3.50	0.5326	0.4152
3.60	0.5879	0.4923
3.70	0.5419	0.5750
3.80	0.4481	0.5656
3.90	0.4223	0.4752

4.00	0.4984	0.4205
4.10	0.5737	0.4758
4.20	0.5417	0.5632
4.30	0.4494	0.5540
4.40	0.4383	0.4623
4.50	0.5260	0.4343
4.60	0.5672	0.5162
4.70	0.4914	0.5671
4.80	0.4338	0.4967
4.90	0.5002	0.4351
5.00	0.5636	0.4992
5.10	0.4998	0.5624

CHAPTER OVERVIEW

4: Stokes Parameters for Describing Polarized Light

Thumbnail: The Poincaré sphere is the parametrisation of the last three Stokes' parameters in spherical coordinates. (Public Domain; Inductiveload).

4.1: Polarized Light and the Stokes Parameters

Suppose that we wish to characterize a beam of parallel monochromatic light. A description of it should include the following.

* Its wavelength or frequency. Its wavelength depends upon the refractive index of the material in which it is travelling, whereas its frequency does not. Therefore, if the wavelength is given, the medium must be specified. It may not always be realized, but most tables of wavelengths of spectrum lines in the visible region of the spectrum are given for *air* and not for a vacuum. [Actually for something called “Standard Air” - details of which may be found in http://orca.phys.uvic.ca/~tatum/stellatm/atm7.pdf] Specifying the frequency rather than the wavelength removes possible ambiguity. Spectroscopists often quote the *wavenumber in vacuo*, which is the reciprocal of the vacuum wavelength.

* Its flux density in W m^{-2}. This is related to the electric field strength of the electromagnetic wave, in a manner that will be discussed later in the chapter.

* Its state of *polarization*. In this chapter, polarized light will in general be taken to mean elliptically polarized light, which includes circularly and linearly (plane) polarized light as special cases. The state of polarization can be described by specifying

* the eccentricity of the polarization ellipse

* the orientation of the polarization ellipse

* the chirality (handedness) of the polarization ellipse

* whether the polarization is *total* or *partial*, and, if partial, the *degree of polarization.*

Up to and including Equation (A15) (page 8) we shall assume that the polarization is *total*. We shall look at *partial* polarization after that.

Polarized light is generally described by supposing that, at some point in space, the tip of the vector that represents the strength of the electric field describes a Lissajous ellipse (Figure IV.1).

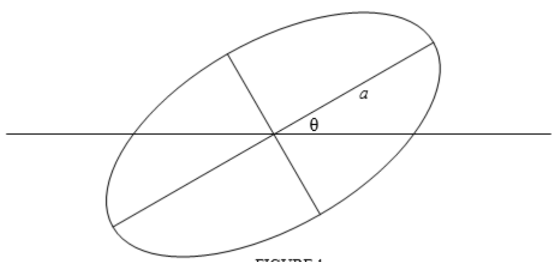


FIGURE 1

In the drawing the semi major axis *a* represents the greatest value of the electric field strength, in volts per metre, during a cycle, and the semi minor axis *b* represent the least value of the electric field strength during the cycle. If you prefer, you could use symbols such as E_{max} and E_{min} instead of a and b.

In order to describe the ellipse, we need to describe its size, its shape, its orientation and its chirality or handedness (i.e., whether the vector is rotating clockwise or counterclockwise).

The natural way of doing this is to give the length a of the semi major axis (in volts per metre), the eccentricity of the ($e = \sqrt{1 - \frac{b^2}{a^2}}$) , the angle θ that the major axis makes with the horizontal, and perhaps one of the words "clockwise" or "counterclockwise". It will be necessary, however, to make clear whether you, the observer, are looking towards the source of light, or are looking in the direction of travel of the light. Not everyone uses the same convention in this matter, and the onus is on the writer to make clear which convention he or she is using. In this chapter I shall assume that we are looking towards the source of the light. In Figure IV.1, I have drawn the ellipse with $\frac{b}{a} = \frac{1}{2}$ $(e = \frac{\sqrt{3}}{2} = 0.8660$) and $\theta = 30^\circ$.

[Since I wrote the above paragraph, I received in December 2015 a memorandum from the International Astronomical Union stating that there has long been an IAU convention that position angle is to be reckoned positive in the counterclockwise direction for an observer looking towards the source of light. This is in fact the convention that I use in these notes. The IAU memorandum, however, pointed out that some scientists who investigate the polarization of the Cosmic Background Radiation have been using the opposite convention, and consequently the IAU reiterates its recommendation that all astronomers, including those working on the CBR, use the above convention. This is a good example of what I meant in the previous paragraph. I would emphasize that, even although there is an IAU convention - one which I strongly support - it is incumbent upon YOU, to make certain, if you wish your readers to understand you, to make it unambiguously clear, whenever you write about polarization, as to what convention you are using. And don't just say "the IAU convention". Say that angles are reckoned positive if increasing counterclockwise when you are facing towards the source of light. I hope that referees and editors will enforce this!]

We noted above that the flux density of the beam is related to the electric field strength of the electromagnetic wave. In this paragraph and the next we explore this relation. Suppose, for) example, that the light is *plane polarized*, and that the maximum value of the electric field is $\hat{E}$ volts. Its mean square value during a cycle is $\overline{E^{\,2}} = \frac{1}{2}\hat{E}^{\,2}$. The energy per unit volume is $\frac{1}{2}\,\epsilon\,\overline{E^{\,2}} = \frac{1}{4}\,\epsilon\,\hat{E}^{\,2}\ \mathrm{J\ m^{-3}}$, where ϵ is the permittivity of the medium in which the radiation is travelling. If it is moving at speed v, the flux density of the beam is $\frac{1}{4}\,v\epsilon\,\hat{E}^{\,2}\ \ \mathrm{W\ m^{-2}}$. The speed of an electromagnetic wave in a medium of permittivity ϵ and permeability μ is given by $v = \frac{1}{\sqrt{\epsilon\mu}}$, so this expression becomes $\frac{1}{4}\sqrt{\frac{\epsilon}{\mu}}\hat{E}^{\,2} = \frac{\hat{E}^{\,2}}{4Z}$, where $Z = \sqrt{\frac{\mu}{\epsilon}}$ is the impedance (in the sense used in electromagnetic theory) of the medium. For most transparent media, μ is very close to μ_0. the permeability of free space. This is not the case for the permittivity, which usually ranges from 1 up to a few tens of times ϵ_0. For a vacuum, the impedance has a value of about 377Ω.

If the light is *elliptically polarized*, the expression for the flux density will be $\frac{a^2+b^2}{4Z}$, where a and b are the electric fields described in earlier paragraphs. That the $\hat{E}^{\,2}$ for plane polarized light can be replaced by $a^2 + b^2$ for elliptically polarized light should become apparent later while discussing the director circle property of an ellipse.

While these parameters may be the obvious ones to use in describing the state of polarization, the fact is that none of them is directly measurable. What we can measure relatively easily is the intensity of the light when viewed through a polarizing filter oriented at various angles. What we can measure are four parameters known as the *Stokes parameter*s, which we shall describe shortly. We can measure the Stokes parameters, and it will then be our task to determine from these the eccentricity, orientation and chirality of the polarization ellipse, and the degree of polarization.

Before describing them, a word about notation.

The traditional symbols used to describe the Stokes parameters are **IQUV**. These may seem somewhat haphazard, so some modern authors prefer a more systematic $S_1,\ S_2,\ S_3,\ S_4$ while some prefer $S_0,\ S_1,\ S_2,\ S_3$. If you use the modern S notation, I would (strongly) recommend $S_0,\ S_1,\ S_2,\ S_3$ over $S_1,\ S_2,\ S_3,\ S_4$. In these notes, however, I shall be old-fashioned and I shall use **IQUV** , which at least has the advantage of avoiding the ambiguity over the two possible S notations, and you will not have to worry which version I am using.

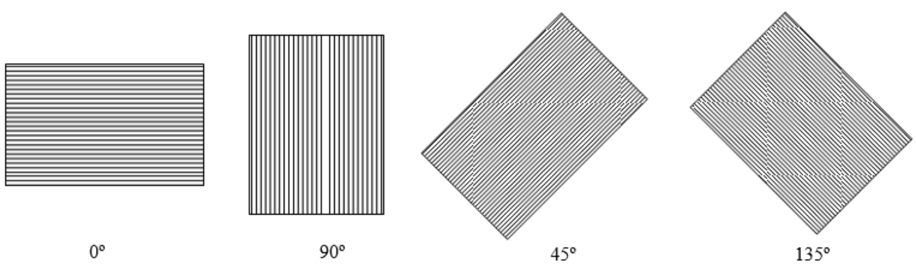


FIGURE 2

In the figure the lines represent the component of the electric field passed by the filter. The lengths of the long organic molecules embedded within the filter are perpendicular to this transmission direction. Light (i.e. an oscillating electromagnetic field) that is oscillating parallel to the lengths of these molecules is strongly absorbed, because of the highly anisotropic polarizability of these molecules.

Perhaps we can measure the intensity of the light after passage through the filter at each of these angles, and also without the filter, and somehow determine from these measurements the shape and orientation of the polarization ellipse.

The Stokes parameters are named after a nineteenth century British physicist, Sir George Stokes, and may be referred to as Stokes's parameters, Stokes' parameters or the Stokes parameters, but not, of course, as Stoke's parameters.

Let us imagine that we have in our hand a flux meter, and that it can measure the flux density, in W m^{-2} of our parallel beam of monochromatic light. While we would prefer to use the symbol F for flux density, in fact the flux density of the unobstructed light is the first of the Stokes parameters, for which the traditional symbol is **I** (and whose modern symbol is S_0 or S_1, depending on which book you are reading.)

Now let us suppose that we measure the flux density of the light after passage through a polarizing filter oriented at various angles as suggested in Figure IV.2. The second and third Stokes parameters, then, are defined by

$$\mathbf{Q} = F_0 - F_{90} \tag{1}$$

and

$$\mathbf{U} = F_{45} - F_{135} \tag{2}$$

Unless you are fortunate or rich, it is unlikely that your little flux meter will accurately measure the flux densities in absolute SI units in W m^{-2}. Therefore those of us of more modest means will just have to be content with dimensionless Stokes parameters - measured in units so that the unobstructed flux density is 1. We define the dimensionless Stokes parameters (for which I use a different font) by

$$Q = \frac{F_0 - F_{90}}{F} = \frac{\mathbf{Q}}{\mathbf{I}} \tag{3}$$

$$U = \frac{F_{45} - F_{135}}{F} = \frac{\mathbf{U}}{\mathbf{I}} \tag{4}$$

Thus, for the dimensioned Stokes parameters in W m^{-2} (which we may not easily be able to measure), I use **IQUV.** For the dimensionless Stokes parameters, I use QUV. (There is no need for a dimensionless I, because it is 1.)

It is possible to determine the eccentricity e and the inclination θ of the polarization ellipse from Q and U. Here I give the relations without derivation. I shall give a derivation in an Appendix to this chapter. For the time being, then, here are the relations:

$$Q = \frac{e^2 \cos 2\theta}{2 - e^2} \tag{5}$$

$$U = \frac{e^2 \sin 2\theta}{2 - e^2} \tag{6}$$

Perhaps of more interest are the converses of these:

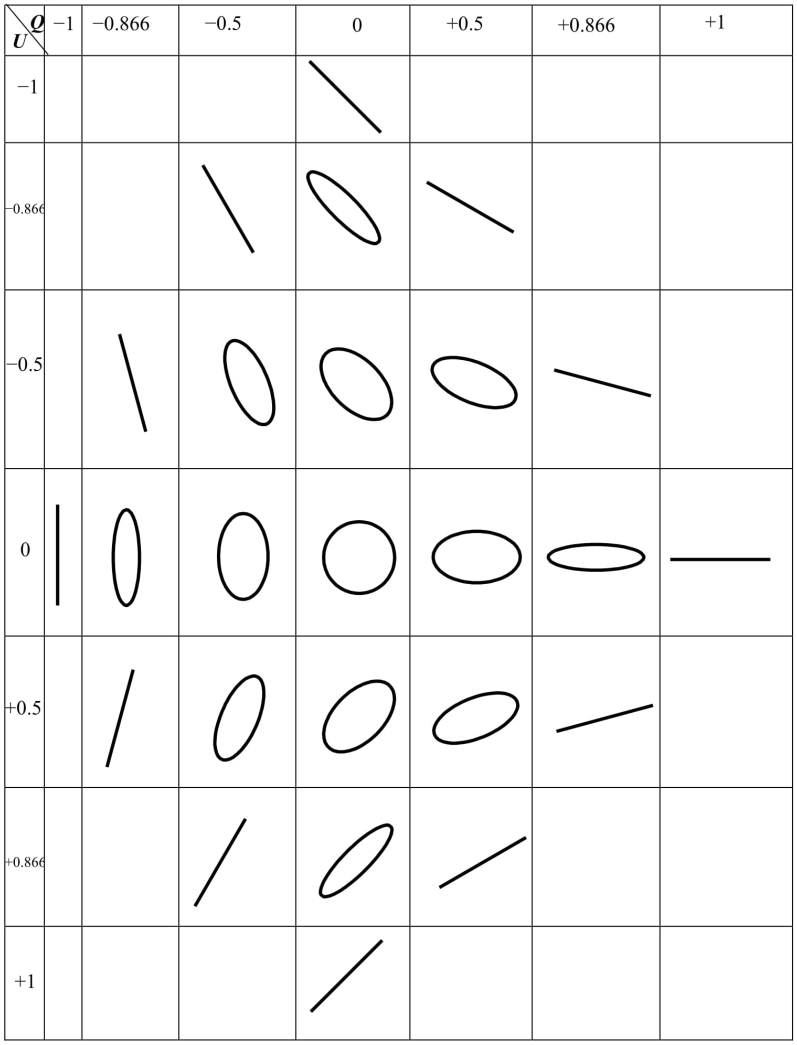
$$e^2 = \frac{2\sqrt{Q^2 + U^2}}{1 + \sqrt{Q^2 + U^2}} \tag{7}$$

$$\tan 2\theta = \frac{U}{Q} \tag{8}$$

In solving Equation (8) for θ, it is necessary to know the signs of U and Q separately, in order to avoid an ambiguity of quadrant. Provision of the $\arctan 2$ function in a calculator or computer greatly facilitates this.

The table below shows a sample of polarization ellipses for various combinations of Q and U. For reasons that will become apparent during the derivation of the formulas in the Appendix, all of the ellipses are drawn such that $a^2 + b^2$ is the same for each. This ensures that the flux density is the same for each.

7



Thus far we have dealt with the Stokes parameters **I** (related to the flux density of the light), and Q and U (related to the shape and orientation of the polarization ellipse). Now we have to describe the Stokes parameter **V**, and how it is related to the chirality (handedness) of the ellipse. In this account, when I use the words "clockwise" and "counterclockwise" I shall assume that we are looking *towards* the source of light.

If we really want to know the polarity, we need to have a good research grant and to be in possession of a filter that passes only circularly polarized light. A linear polarizer in conjunction with a quarter-wave plate will do it. I shall take it that the filter passes only light that is circularly polarized in the clockwise sense. Suppose the flux density after passage through such a filter is F_C. The Stokes **V** parameter is defined as

$$\mathbf{V} = 2F_C - I, \tag{9}$$

or, in dimensionless form,

$$V = \frac{2F_C}{F} - 1. \tag{10}$$

It will be observed that this parameter (like the others) ranges from −1 (if $F_C = 0$) to +1 (if $F_C = 1$), and hence that negative V implies counterclockwise polarization, and positive V implies clockwise polarization. We shall also show in the Appendix, that (*subject to an important condition* - see below), V is related to the eccentricity by

$$V^2 = \frac{4(1-e^2)}{(2-e^2)^2}. \tag{11}$$

This means that $V=0$ implies $e=1$, and hence linear polarization (for which there is no chirality). Also, $V^2=1$ implies $e=1$, and hence circular polarization. Conversely

$$e^2 = \frac{2(-1+V^2+\sqrt{1-V^2})}{V^2} \tag{12}$$

Thus one can determine both the chirality and the eccentricity (but not θ) from V alone. Figure IV.3 shows the relation between $|V|$ and e.

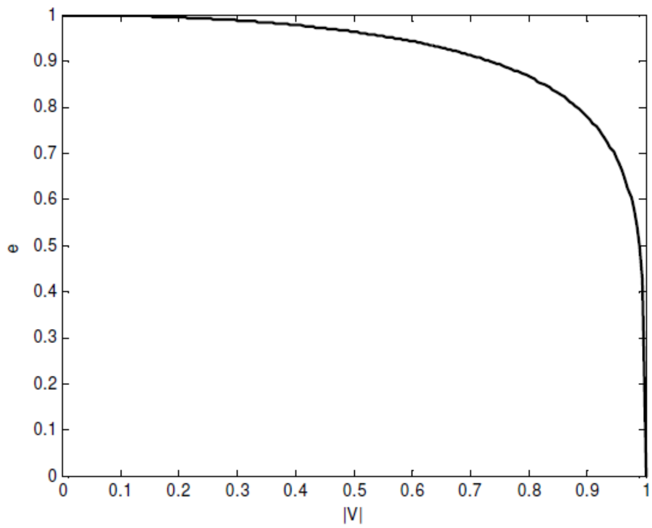


FIGURE 3

This redundancy must mean that Q,U and V are not independent, and indeed it will be observed from equations (5), (6) and (11) that

$$Q^2 + U^2 + V^2 = 1. \tag{13}$$

In terms of the dimensioned Stokes parameters:

$$\mathbf{Q^2 + U^2 + V^2 = I^2.} \tag{14}$$

In one of the S notations, this would conveniently be

$$S_1^2 + S_2^2 + S_3^2 = S_0^2. \tag{15}$$

Just before Equation (11) we referred to an important condition. Equations (11) - (15), and Figure IV.3, are valid only for the case of total elliptical polarization. The case of partial polarization is discussed in what follows. The section on partial polarization should not be thought of as a relatively unimportant afterthought, because most sources of polarized light that one comes across are more likely to be partially polarized rather than totally polarized.

Partial Polarization

Until this point we have assumed that we have been concerned with a single coherent wave with one well-defined polarization state. In practice, we rarely see this, and we more often have to deal with *partially polarized* light. Most of us have a fairly good idea of what is meant by light that is partially plane polarized horizontally. We mean that the light is mostly like this:

but there’s also a little bit of this:

But if that were so with two coherent waves, this would result, if they were in phase, in this:

or if they were not in phase, in this:

In truth, unless we are looking at a coherent light source, such as a laser, partially polarized light might be more like this:

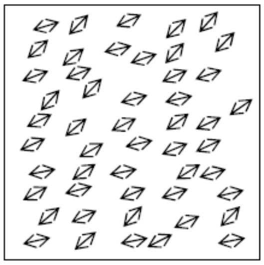


FIGURE 4

This is partially plane polarized at about an angle of 30º, but it is clearly not totally plane polarized. Partially polarized light can be described as the sum of a totally polarized component plus an unpolarized component. Thus we might describe the situation illustrated above by something like this:

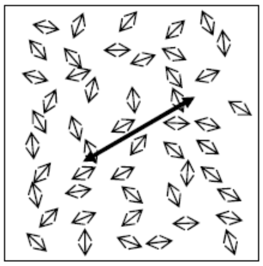


FIGURE 5

Partially elliptical polarized light might be described by a totally elliptically polarized component, plus an unpolarized component:

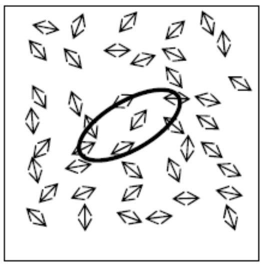


FIGURE 6

If we could somehow separately measure the flux densities of the polarized (p) and unpolarized (u) components, we could define the degree of polarization by

$$p = \frac{F_p}{F_p + F_u} \tag{16}$$

If we know that the light is partially plane (linearly) polarized, as in Figure IV.5 (rather than elliptically polarized as in Figure IV.6), we can measure this rather easily. Place the polarizing filter in front of the source, and rotate it until the transmitted flux density goes through a maximum, $F_{\max}$. and then through a further 90º until it goes through a minimum, $F_{\min}$. This will give you the degree of polarization from.

$$p=\frac{F_{\max}-F_{\min}}{F_{\max}+F_{\min}}. \tag{17}$$

and of course it also gives you the polarization angle. This applies, of course, only to light that you know to be partially linearly polarized. It will not do for partially elliptically polarized light.

Recall that

$$Q=\frac{F_{0}-F_{90}}{F}=\frac{\mathbf{Q}}{\mathbf{I}} \tag{3}$$

and

$$U=\frac{F_{45}-F_{135}}{F}=\frac{\mathbf{U}}{\mathbf{I}} \tag{4}$$

If the source is partially plane polarized, each of the measurements $F_0, F_{90}, F_{45}, F_{135}$ includes a total linear or elliptical component, and an unpolarized component. However, the unpolarized component is the same for each of these four measurements. Consequently Q and U describe the "total" component only. Thus all equations up to and including Equation (8), as well as the table illustrating the shape of the ellipse as a function of Q and U, are still valid for the "total" component.

The parameter V, however, was defined in Equations 9 and 10 by

$$\mathbf{V}=2F_{C}-I, \tag{9}$$

or, in dimensionless form,

$$V=\frac{2F_{C}}{F}-1. \tag{10}$$

F_C and F each contain a "total" and an unpolarized component, so that, unlike Q and U, the "total" component is not separated out.

Recall from Equations (13) and (14) that $\mathbf{I}=\sqrt{\mathbf{Q^2}+\mathbf{U^2}+\mathbf{V^2}}$ and $Q^2+U^2+V^2=1$.

These were derived for totally elliptically (which includes linearly) polarized light. For light that is partially polarized, it applies only to the "total" part, so that, for partially polarized light,

$$\mathrm{p}=\sqrt{Q^{2}+U^{2}+V^{2}}. \tag{18}$$

From Equations (5), (6) and (18) we determine that

$$\mathrm{p}=\sqrt{V^{2}+\frac{e^{4}}{(2-e^{2})^{2}}} \tag{19}$$

Thus from the measurements of $F_0, F_{90}, F_{45}, F_{135}$ and their combinations $IQUV$ we have determined, for partially polarized light, the degree of polarization, and the eccentricity, orientation and chirality of the polarization ellipse.

Equation (18) suggests that that the state of polarization of light can be described by a point in QUV space . This concept is described by the *Poincaré sphere*:

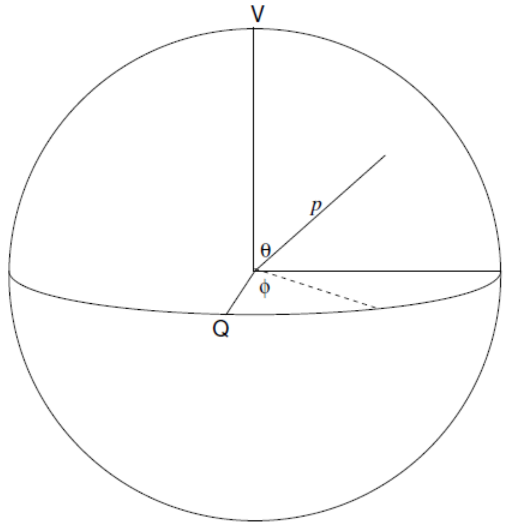
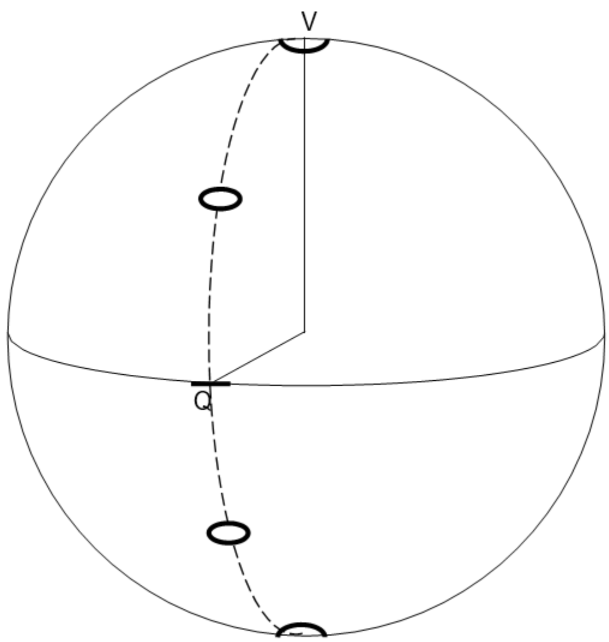


FIGURE 7

In this context I have often seen the notation 2ψ for ϕ and 2χ for 90º − θ. (The θ here, of course, is not the same as the θ of Figure IV.1.

Let us suppose, to begin with, that we have total polarization, so that $p = 1$. The reader is invited to imagine the shape of the polarization ellipse at any point on the surface of the sphere. Recall in particular that $V = 0$ implies linear polarization, and $V = \pm 1$ implies circular polarization. Thus anywhere around the equator of the Poincaré represents linear polarization, and at the poles we have circular polarization.

Let us look along the meridian of longitude with ϕ = 0 (U= 0). As we go from the "north pole" to the "south pole", V goes from +1 (circular) through 0 (linear) to −1 (circular), and Q goes from 0 (circular) through 1 (linear) to 0 (circular). It will be useful (essential) to refer to the table on page 5.



The reader is now invited to think about (while referring to the table on page 5) the situation along the meridian with φ = 90º. And then to try other meridians, eventually covering the sphere with ellipses. This is a little beyond my artistic ability, but I found a very good one by Googling for Poincaré sphere. Choose "Images for poincare sphere". There are some excellent images there. I particularly like the orange-coloured one from University of Arizona. If you click on it, the sphere rotates, and you can see all round the sphere.

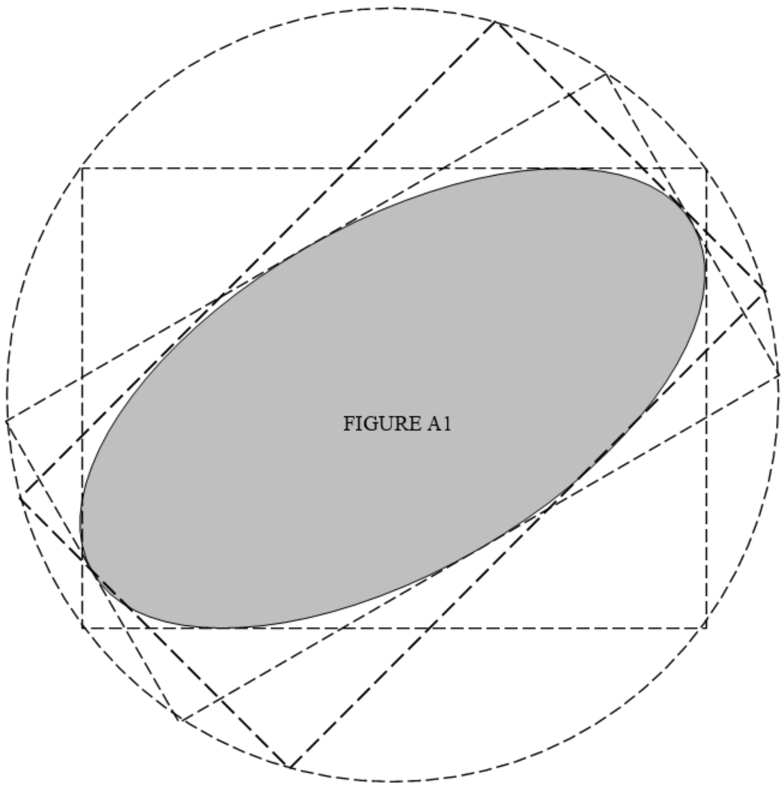
APPENDIX

In the article above I described the Stokes parameters, and I related them to the shape, orientation and chirality of the polarization ellipse, as follows (for total polariazation):

$$Q = \frac{e^2 \cos 2\theta}{2 - e^2} \qquad U = \frac{e^2 \sin 2\theta}{2 - e^2} \qquad V^2 = \frac{4(1 - e^2)}{(2 - e^2)^2}$$

In this Appendix, I derive these relations.

Before starting, let us remind ourselves of an established property of an ellipse of semi major and semi minor axes a and b, namely that the locus of the corners of all circumscribing rectangles to an ellipse is a circle, known as the *director circle*, which is of radius $\sqrt{a^2 + b^2}$. This is illustrated in Figure A1, in which I have drawn three circumscribing rectangles. The semidiagonals of all the circumscribing rectangles are of the same length, namely $\sqrt{a^2 + b^2}$. A proof of this theorem is to be found in http://orca.phys.uvic.ca/~tatum/celmechs/celm2.pdf , Section 2.3, or in many books on the properties of the conic sections.



Recall now the meanings of a and b. They are the semi major and semi minor axes of the ellipse, but they are also the greatest and least values of the electric field during a cycle. Recall also that the energy per unit volume of an electric field is proportional to the square of the electric field strength. When the light is observed direct without the intervention of a polarizing filter, the flux density of the light is proportional, then, to $a^2 + b^2$. That is to say, the Stokes parameter I is proportional to the square of the radius of the director circle.

In what follows, we shall have occasion to refer the polarization ellipse to **three rectangular coordinate systems**.

i. A coordinate system (x, y), in which the axes of coordinates coincide with the axes of the polarization ellipse.

ii. A coordinate system (x_1, y_1), in which the axes of coordinates are horizontal and vertical - or, to more precise, parallel to the transmission axes of the first two filters illustrated in Figure IV.1.

iii. A coordinate system (x_2, y_2), in which the axes of coordinates are parallel to the transmission axes of the last two filters illustrated in Figure IV.1.

The ellipse referred to these three coordinate systems is shown in Figures A2, A3, A4. In each of these drawings, I have drawn a circumscribing rectangle and the director circle. The flux density of the radiation is proportional to the square of the rectangle diagonal, which is the same in all three drawings, and is equal to the diameter of the director circle, namely $2\sqrt{a^2 + b^2}$.

I have also indicated the lengths $a,\ b,\ a_1,\ b_1,\ a_2,\ b_2$ in these drawings. These represent the maximum values of the component of the electric field during a cycle in the directions of the six axes. Indeed, the reader might even prefer an alternative notation:

$$\begin{aligned} a &= \hat{E}_x \\ b &= \hat{E}_y \\ a_1 &= \hat{E}_{x_1} \\ b_1 &= \hat{E}_{y_1} \\ a_2 &= \hat{E}_{x_2} \\ b_2 &= \hat{E}_{y_2} \end{aligned}$$

The first notation is easier for the analysis of the geometry of the ellipse. The second notation reminds us of the physical meaning of the symbols. Indeed the readings of our flux meter are proportional, successively, to $\hat{E}_x{}^2 + \hat{E}_y{}^2,\ \hat{E}_{x_1}{}^2,\ \hat{E}_{y_1}{}^2,\ \hat{E}_{x_2}{}^2,\ \hat{E}_{y_2}{}^2,$ or, in the a, b notation $a^2 + b^2,\ a_1^2,\ b_1^2,\ a_2^2,\ b_2^2,$. The Stokes parameters I, Q, U are proportional successively to $\hat{E}_x{}^2 + \hat{E}_y{}^2,\ \hat{E}_{x_1}{}^2 - \hat{E}_{y_1}{}^2,\ \hat{E}_{x_2}{}^2 - \hat{E}_{y_2}{}^2,$, or in the a, b notation, $a^2 + b^2,\ a_1^2 - b_1^2,\ a_2^2 - b_2^2$.

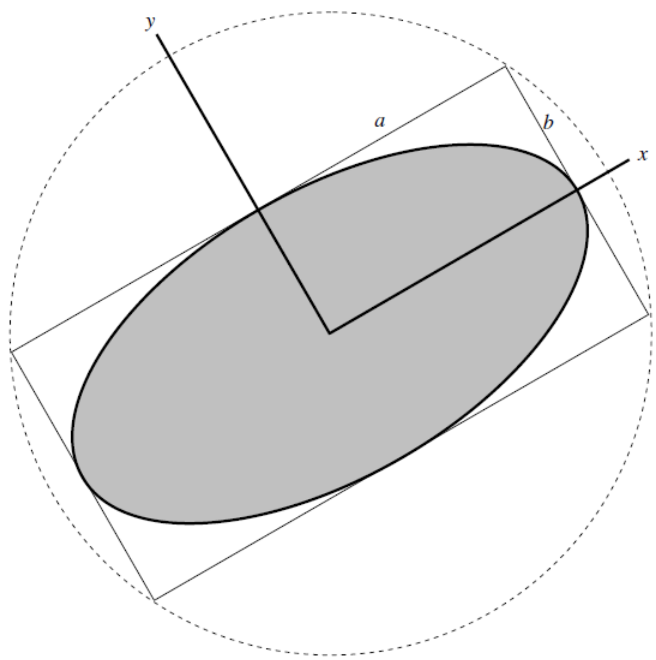


FIGURE A2

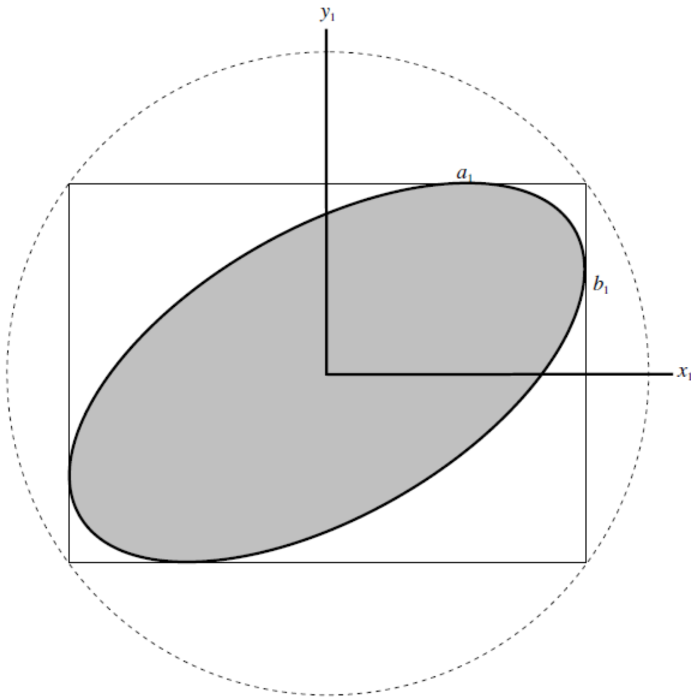


FIGURE A3

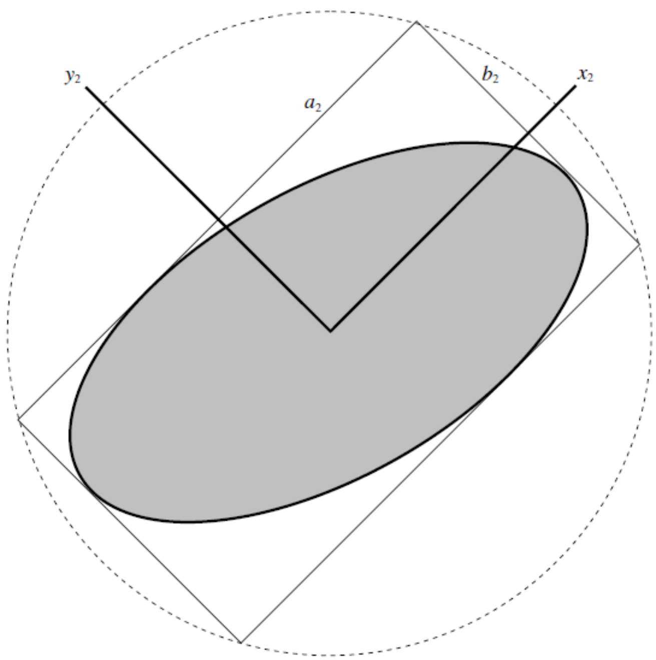


FIGURE A4

Refer to Figure A2. The equation to the ellipse, referred to this coordinate system, is the familiar

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \tag{A1}$$

However, I want to express lengths (electric field strengths) in units such that $a^2 + b^2 = 1$, and, further, I want to write the equation in terms of the eccentricity $e = \sqrt{1 - \frac{b^2}{a^2}}$. In that case, Equation (A1) becomes

$$fx^2 + gy^2 = 1 \tag{A2}$$

where

$$f = 2 - e^2 \text{ and } g = \frac{2-e^2}{1-e^2}. \tag{A3}$$

Now refer to Figure A3. If the major axis of the ellipse makes an angle θ with the horizontal, the coordinate systems are related by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \tag{A4}$$

$c = \cos\theta$ and $s = \sin\theta$

On making use of equations (A2) and (A4), we find that the equation to the ellipse referred to the (x_1, y_1) coordinate system is

$$(fc^2 + gs^2)x_1^2 - 2(g-f)scx_1y_1 + (fs^2 + gc^2)y_1^2 = 1 \tag{A5}$$

We now wish to find $a_1 = \hat{E}_{x_1}$ and $b_1 = \hat{E}_{y_1}$, the maximum horizontal and vertical components of the electric field. The length a_1 can be found as follows. The vertical line $x_1 = a_1$ intersects this ellipse at values of y_1 given by

$$(fs^2 + gc^2)y_1^2 - 2(g-f)sca_1y_1 + (fc^2 + gs^2)a_1^2 - 1 = 0 \tag{A6}$$

But the line $x_1 = a_1$ is to be a vertical tangent to the ellipse, and therefore the quadratic equation (A6) must have two equal real roots, which tells us, after a little algebra, that

$$a_1^2 = \frac{fs^2 + gc^2}{fg}. \tag{A7}$$

A similar analysis starting with the horizontal line $y_1 = b_1$ reveals that

$$b_1^2 = \frac{fc^2 + gs^2}{fg}. \tag{A8}$$

For a check on the correctness of the algebra, it can now be verified that $a_1^2 + b_1^2 = 1$.

The Stokes Q parameter is $a_1^2 - b_1^2$, and, after some algebra and trigonometric identities, it is found that

$$Q = a_1^2 - b_1^2 = \frac{e^2 \cos 2\theta}{2-e^2}, \tag{A9}$$

which is one of the relations that we sought.

Now refer to Figure A3. The x_2, y_2 and x, y coordinate systems are related by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} C & -S \\ S & C \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}, \tag{A10}$$

where

$$S = \sin(45^\circ - \theta) \quad \text{and} \quad C = \cos(45^\circ - \theta). \tag{A11}$$

On making use of Equations (A2) and (A10), we find that the equation to the ellipse referred to the (x_2, y_2) coordinate system is

$$(fC^2 + gS^2)x_2^2 + 2(g-f)SCx_2y_2 + (fS^2 + gC^2)y_2^2 = 1 \tag{A12}$$

To obtain U, we now proceed in a similar fashion to the analysis of Q. We combine this equation with $x_2 = a_2$ and put in the condition that the resulting quadratic equation in y_2 has two equal real roots, to obtain

$$a_2^2 = \frac{fS^2 + gC^2}{fg} \tag{A13}$$

Likewise, by combination with $y_2 = b_2$, we obtain

$$b_2^2 = \frac{fC^2 + gS^2}{fg} \tag{A14}$$

The correctness of the algebra can be checked by verifying that $a_2^2+b_2^2=1$. Then U, which is $a_2^2-b_2^2$, can be calculated with some algebra and trigonometry, to be

$$U=\frac{e^2\sin 2\theta}{2-e^2}. \tag{A15}$$

And this is a good time to remind ourselves of equation (A9)

In our drawings in this chapter, we have taken $b=\frac{1}{2}a, e=\frac{\sqrt{3}}{2}, \theta=30^\circ$ so that $Q=0.3, U=0.5196$.

Now for the *chirality* or *handedness* of the radiation. From measurements of Q and U we have deduced the eccentricity and orientation of the Lissajous ellipse, but we don't yet know whether the tip of the **E**-vector is moving clockwise or counterclockwise (as seen when looking towards the source of light). This is what the Stokes V parameter is going to tell us.

It is well known that a Lissajous ellipse can be generated as the resultant of two simple harmonic linear oscillations at right angles to each other. In order to understand the V *parameter* it is necessary to understand that a Lissajous ellipse can also be generated by two *circular* motions, of different amplitude, and moving in opposite directions. If the semi major and semi minor axes of the Lissajous ellipse are, respectively, a and b, the radii of the circular components are $\frac{1}{2}(a+b$ and $\frac{1}{2}(a-b$ (see Figure 11).

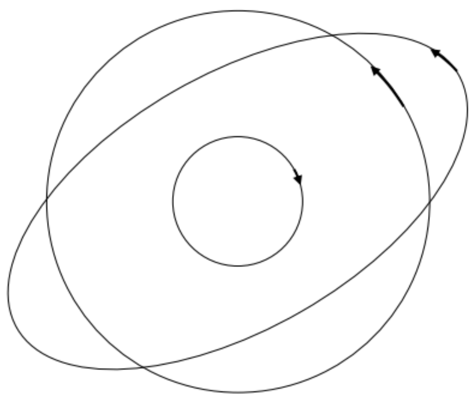


FIGURE A5

To measure V we place in front of the light source a filter that transmits only *circularly* polarized light. We'll suppose that it transmits light that is left-handed (counterclockwise) as seen when looking towards the light source. I.e. it will obstruct the smaller circle of Figure A5 and transmit the large circle.

If the fraction of the flux density passed by the filter is f, the Stokes V parameter is $2f-1$.

Examples:

If the light is lefthand circularly polarized, the filter will transmit all of the light. That is, $f=1, V=1$.

If the light is righthand circularly polarized, the filter will transmit none of the light. That is, $f=0, V=-1$.

If the light is linearly polarized, the filter will transmit half of the light. (Linearly polarized light can be generated by two equal circles moving in opposite directions.) That is, $f=\frac{1}{2}, V=0$.

In Figure A5, $b=\frac{1}{2}$. The radius of the small circle (which is obstructed) is $\frac{1}{4}a$ and the radius of the large circle (which is transmitted) is $\frac{3}{4}a$. The flux density of the unfiltered light is proportional to $a^2+b^2=\frac{5}{4}a^2$. The flux density of the light that is passed is proportional to $\frac{9}{8}a^2$. (The flux density, we recall, is proportional to the square of the director circle. The radius of the director circle of the large circle is $\sqrt{(\frac{3}{4}a)^2+(\frac{3}{4}a^2)}=\sqrt{(\frac{8}{9}a)^2}$. So we have $f=0.9, V=0.8$.

If we were to reverse all of the arrows in Figure A5, it would be the larger circle that would be blocked and the small circle passed. The flux density of the light that is passed is then proportional to $\frac{1}{8}a^2$. So we have $f=0.1, V=-0.8$.

Thus positive V means that the tip of the **E**-vector is moving counterclockwise, and negative V means that it is rotating clockwise.

In general, the radius of the large circle is $\frac{1}{2}(a+b)$ and the radius of its director circle is $\frac{1}{\sqrt{2}}(a+b)$. If this is the circle that is transmitted, the flux density passed is proportional to $\frac{1}{2}(a+b)^2$.

We have, then, $f=\frac{\frac{1}{2}(a+b)^2}{a^2+b^2}, V=\frac{2ab}{a^2+b^2}$. This means, incidentally, that V is proportional to the area of the ellipse. If we take $a^2+b^2=1$, then $V=2ab$. If it is the small circle that is passed, $f=\frac{\frac{1}{2}(a-b)^2}{a^2+b^2}, V=\frac{-2ab}{a^2+b^2}$.

Since the eccentricity of the ellipse is given by $e^2=1-\frac{b^2}{a^2}$, we can express V^2 in terms of the eccentricity, thus

$$V^2=\frac{4(1-e^2)}{(2-e^2)^2}. \tag{A16}$$

This equation is valid for totally polarized light. For partially polarized light, return to the main text.

Index

Index

Glossary

Sample Word 1 | Sample Definition 1

Detailed Licensing

Overview

Title: Physical Optics (Tatum)

Webpages: 29

Applicable Restrictions: Noncommercial

All licenses found:

- Undeclared: 51.7% (15 pages)
- CC BY-NC 4.0: 48.3% (14 pages)

By Page

- Physical Optics (Tatum) - *CC BY-NC 4.0*
  - Front Matter - *Undeclared*
    - TitlePage - *Undeclared*
    - InfoPage - *Undeclared*
    - Table of Contents - *Undeclared*
    - Licensing - *Undeclared*
  - 1: Reflection and Refraction via Fermat's Principle and Huygens' Construction - *CC BY-NC 4.0*
    - 1.1: Reflection and Refraction - *CC BY-NC 4.0*
  - 2: Reflection and Transmission at Boundaries and the Fresnel Equations - *CC BY-NC 4.0*
    - 2.1: Waves in a Stretched String - *CC BY-NC 4.0*
    - 2.2: Light Incident Normally at a Boundary - *CC BY-NC 4.0*
    - 2.3: Light Incident at the Brewster Angle - *CC BY-NC 4.0*
    - 2.4: Electric and Magnetic Fields at a Boundary - *CC BY-NC 4.0*
    - 2.5: Impedance - *CC BY-NC 4.0*
    - 2.6: Incidence at an Arbitrary Angle. - *CC BY-NC 4.0*
  - 3: The Cornu Spiral - *CC BY-NC 4.0*
    - 3.1: Cornu's Spiral - *CC BY-NC 4.0*
  - 4: Stokes Parameters for Describing Polarized Light - *CC BY-NC 4.0*
    - Front Matter - *Undeclared*
      - TitlePage - *Undeclared*
      - InfoPage - *Undeclared*
    - 4.1: Polarized Light and the Stokes Parameters - *CC BY-NC 4.0*
    - Back Matter - *Undeclared*
      - Index - *Undeclared*
  - Back Matter - *Undeclared*
    - Index - *Undeclared*
    - Index - *Undeclared*
    - Glossary - *Undeclared*
    - Detailed Licensing - *Undeclared*