Third Level Module - Optics

Although ‘optics’ has been taught in several parts of your curriculum to date - *e.g.* geometrical optics as part of the laboratory work, diffraction as an illustration of the importance of Fourier transformation in Physical Mathematics, this set of lectures is the first course to specialise in optics. The course will not attempt to cover the entire wide range of topics included in the recommended textbooks; rather it will concentrate on a few of the key topics and thus enable you to read further in the subject if/when interested. The course will assume little previous knowledge of optics but will only give reminders of the material which has already been covered in earlier years. Please note that some of the topics in Electromagnetism are relevant to this course, *e.g.*the radiation pattern from a Hertzian dipole, reflection and transmission coefficients of an electromagnetic wave at a boundary, *etc.*

Optics Syllabus

(Not necessarily covered in the order below)

- Introduction
- Basic Geometric Optics
- Polarisation effects
    1. Types of polarised light, methods of production and detection
    2. Polaroid, Nicol prism
- Birefringence
    1. Quarter- and half-wave plates
    2. Second harmonic generation
- Optical activity
- Diffraction
    1. Fraunhofer and Fresnel

2. Far field diffraction as a Fourier Transform (link with Physical Mathematics) - one dimension
3. Diffraction from two dimensional object - rectangular aperture and circular aperture.
4. Resolution limits

- Convolution (link with Physical Mathematics)
- The diffraction grating
- The Zone-plate
- Interferometry
  1. Michelson,
  2. Twyman-Green,
  3. Fourier Transform Spectroscopy
- Thin Films
  1. anti-reflection coatings
  2. multi-layer reflecting coatings, cold mirrors
  3. flatness of surfaces
- Fabry-Perot interferometer
- Optical Processing
  1. Regularity detection/ suppression
  2. Edge enhancement
  3. Half-tone suppression
- Photographic properties (Possibly but unlikely)
  1. Basis of light sensitive materials
  2. Exposure and optical density
  3. Optical transmission and linearisation by double photographic process
- Holography

Recommended textbooks:-

'Optics' by E. Hecht, published by Addison and Wesley World Student Series
'Geometrical and Physical Optics' by R.S.Longhurst, published by Longman
Other textbooks are
'Optics' by F G Smith and J H Thomson, published by Wiley.
and useful
'Insight into Optics' by Heavens and Ditchburn, published by Wiley

1 Basic Geometric Optics

1.1 Ray optics

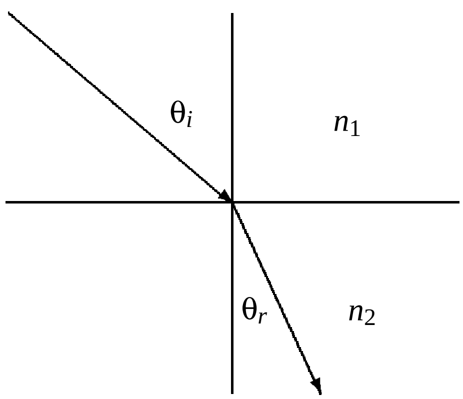
Straight line propagation of light except at the surface between two transparent media (or in a non-isotropic material); the Law of Reflection and Snell's Law of Refraction apply at the boundary.

Law of Refraction;

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

Refractive index of vacuum, $n_{vac} = 1$

Refractive index of air, $n_{air} = 1.00029$, (*i.e.* ≈ 1)



1.2 Fermat's Principle

Fermat's Principle of Least Time: - Fermat established the basic principle of why light takes one path rather than another. He postulated that in travelling from A to B, light chooses the most economical route, *i.e.* the time taken for the path is a minimum.

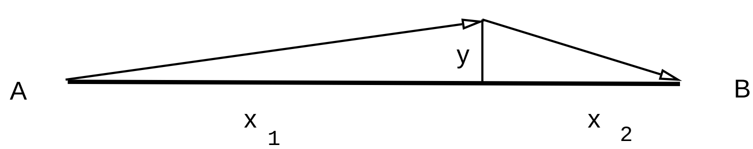
In fact, this is now expressed more accurately in the form:-

"The actual path taken by a ray of light is the one for which the Optical Path Length (**OPL**) is stationary with respect to variations of that path."

The definition of the optical path length is the product of the refractive index of the medium multiplied by the geometrical path length.

Example 1

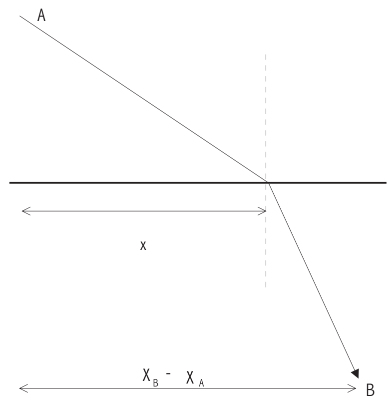
Straight line travel from A to B in a homogeneous medium is obvious but here is a derivation.



$$OPL = \sqrt{(x_1^2 + y^2)} + \sqrt{(x_2^2 + y^2)}$$
$$d(OPL)/dy = y/\sqrt{(x_1^2 + y^2)} + y/\sqrt{(x_2^2 + y^2)}$$
$$d(OPL)/dy = 0$$

only if $y = 0$, *i.e.* straight line travel.

Example 2, Refraction



Let the co-ordinates of the source of the light be (x_A, y_A) in the first medium and those of point B in the second medium be (x_B, y_B). If the light enters the second medium at a horizontal distance x (treated as a variable) from the source point, the optical path length between A and B is

$$OPL = n_1(y_A^2 + x^2)^{\frac{1}{2}} + n_2(y_B^2 + (x_B - x_A - x)^2)^{\frac{1}{2}}$$
$$\frac{d(OPL)}{dx} = n_1 \frac{1}{2} \frac{2x}{(y_A^2 + x^2)^{\frac{1}{2}}} + n_2 \frac{1}{2} \frac{(x_B - x_A - x)(-1)}{(y_B^2 + (x_B - x_A - x)^2)^{\frac{1}{2}}}$$
$$\frac{d(OPL)}{dx} = n_1 \sin\theta_1 - n_2 \sin\theta_2 = 0$$

i.e.

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

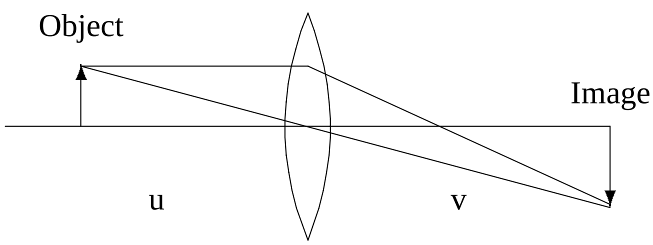
1.3 Converging and diverging lenses.

All converging lenses are thicker at the middle than at the edge. For a perfect lens, *i.e.* one free from aberrations, rays travelling parallel to the axis are brought to a focus at a point on the axis one focal length from the lens – at the rear principal focus. Similarly, rays leaving the front principal focus travel parallel to the axis after passing through the lens.

Diverging lenses are thinner at the middle than at the edge. For a perfect lens, rays travelling parallel to the axis diverge after passing through the lens; they seem to originate from a point one focal length in front of the lens. The image formed by a diverging lens is often “virtual”.

1.4 Image Formation

- Rays pass straight through the centre of a thin lens.
- Rays travelling parallel to the axis pass are deviated and pass through the focal point.
- To trace any ray, draw a construction line parallel to it on the object side of the lens and through the centre of the lens. Extend this construction line to cut the rear focal plane. The required ray also passes through the rear focal plane at this point.



Imaging properties of the lens

$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$$

The magnification is $M = \frac{v}{u}$

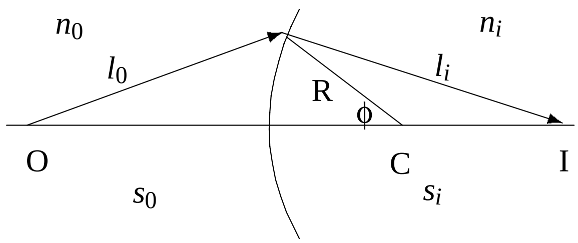
1.5 Lens Aberrations

- (a) Aperture aberration. It has been shown by James Clerk Maxwell that no lens can give perfect imagery except for one particular pair of image and object points. For other pairs of image and object points, rays from the outer parts of the lens are brought to a focus at a different distance from the lens than rays close to the axis. This aberration is characteristic of lenses with spherical surfaces (hence the alternative name) but is present for all surface shapes. The defect can be reduced by reducing the lens aperture but this also reduces the brightness of the image formed by the lens.
- (b) Chromatic aberration. The refractive index of all transparent materials is a function of the wavelength of light - dispersion. Blue light is brought to a focus nearer to the lens than is red light. It can be corrected by forming an achromatic doublet – a combination of a strong, converging lens made of a glass of low dispersion and a weak, diverging lens made of a glass of high dispersion. The doublet is designed so that the red and the blue are brought to the same focus; the other colours will still be brought to a focus somewhat closer to the lens.

1.6 Imaging by a Lens

(a) Spherical surface

See figure below



Optical Path Length from O via P to I is $OPL = n_o l_0 + n_i l_i$

Triangle OPC gives l_0^2, *i.e.* $l_0^2 = (s_0 + R)^2 + R^2 - 2(s_0 + R)R\cos(\phi)$

Triangle CPI gives l_i^2, *i.e.* $l_i^2 = (s_i - R)^2 + R^2 + 2(s_i - R)R\cos(\phi)$

To find the relationship between l_0, l_i *etc.* for the ray from O to P to I, apply Fermat's Principle, *viz.* $\Delta(OPL)/\Delta(\phi) = 0$.

$$\Delta(OPL)/\Delta\phi = (n_o/l_0)R(s_0 + R)\sin(\phi) - (n_i/l_i)R(s_i - R)\sin(\phi)$$

$$n_0/l_0 + n_i/l_i = (1/R)(n_i s_i/l_i - n_o s_0/l_0) \quad (1)$$

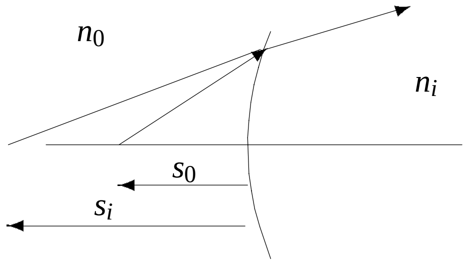
N.B. The OPL is a function of ϕ, therefore the rays are brought to a focus at different points on the axis depending on ϕ. This is aperture/ spherical aberration.
N.B. It is only in Britain that this is named *spherical aberration*; it is more correctly named "aperture aberration". It is reduced in photography by using 'stops'.

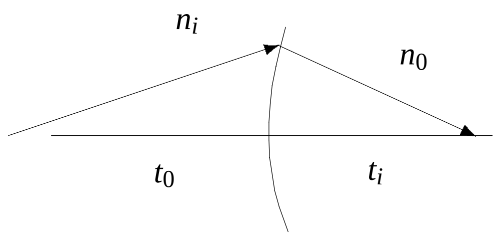
(b) Paraxial Rays Approximation

This corresponds to first order theory, where the rays are considered to be close to the axis of the lens, *i.e.* ϕ is small and $\cos(\phi) \approx 1$. Thus $s_0 \approx l_0$ and $s_i \approx l_i$; equation 1 above gives

$$n_0/s_0 + n_i/s_i = (n_i - n_0)/R \quad (2)$$

Now consider moving the object closer to the single surface so that s_0 decreases and s_i becomes negative, *i.e.* the single surface gives a virtual image, **see figure below**.





Introduce a second surface of radius of curvature R_2 with the same sign of curvature; the rays arriving at the second surface seem to come from a distance s_i to the left of the first surface corresponding to a distance t_0 to the left of the second surface but from a refractive index of n_i, **see figure above**.

$$t_0 = -s_i$$

The new equation 2 is

$$n_i/t_0 + n_0/t_i = (n_0 - n_i)/R_2 \tag{3}$$

Add equation 2, with R written as R_1, to equation 3

$$n_0(1/s_0 + 1/t_i) = (n_i - n_0)(1/R_1 - 1/R_2)$$

or:

$$1/s_0 + 1/t_i = (n_i/n_0 - 1)(1/R_1 - 1/R_2)$$

This is one form of the **Lensmakers Formula**; note that the left-hand side is the focal length of the double surface system, *i.e.* the lens and

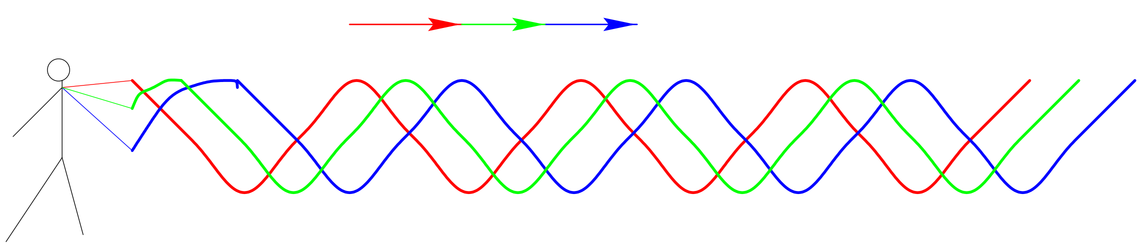
$$1/f = (n_i/n_0 - 1)(1/R_1 - 1/R_2)$$

Note that a normal convex lens will have R_1 positive and R_2 negative.

Example: a plano-convex lens of refractive index = 1.5; if the object is on the plane side, $R_1 = \infty$, $R_2 = -50mm$, say, and $f = 100mm$.
If the object is on the convex side, $R_1 = 50mm$, $R_2 = \infty$ and $f = 100mm$.

2 Waves

2.1 Travelling waves



Transverse displacement of string as a function of x at three equally spaced times. The end of the rope executes SHM with period T.

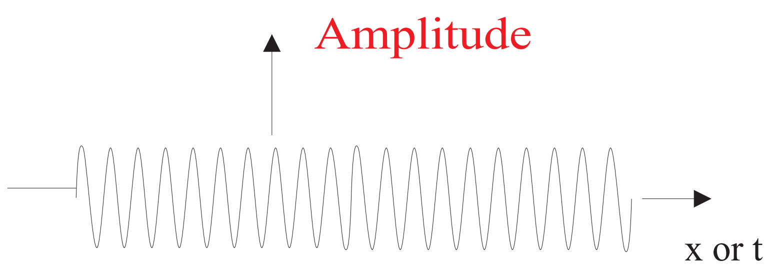
If the string is long enough, after a long time the travelling wave can be described by the equation
$y = A\cos(2\pi\frac{x}{\lambda} - 2\pi\frac{t}{T})$ or $y = A\cos(kx - \omega t)$
or

$$y = Ae^{i(kx-\omega t)}$$

Remember that the disturbance on the string must be finite in both length and time and this equation implies that it is not limited in x or in t.

The travelling wave moves forward by a distance λ in a time T, *i.e.* the velocity of the wave is $v = \frac{\lambda}{T} = \frac{\omega}{k}$.

A more realistic snapshot of a wave packet would be to assume that the string can be started and stopped instantaneously.



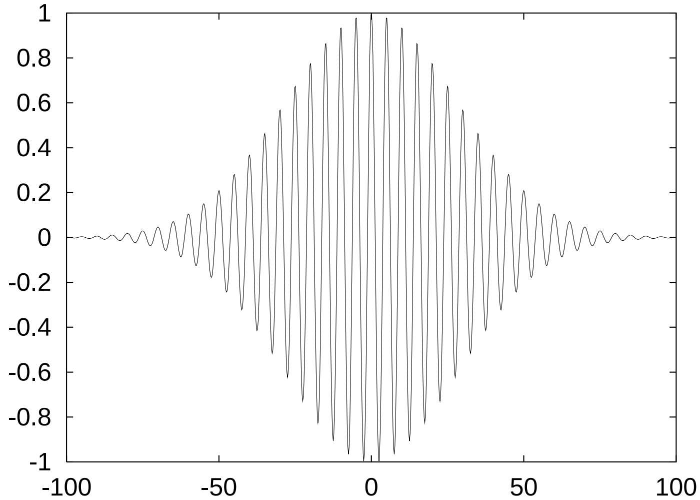
This simple wave packet can be represented as the product of an infinite sinusoidal wave of frequency ω_0 and a top hat function. The distribution of the frequency components of the wave packet are obtained by taking the Fourier Transfrom of the product. The frequency distribution is a delta-function at the frequency of the sinusoidal wave convolved with a sinc-function.

Note that a continuous spectrum of frequencies is required to define a single wave group or wave packet or photon.

2.2 Wave Packet

The wave packet from a light source is often described as a product of a travelling 'cosine wave', $y = A\cos(kx - \omega t)$ modulated by a 'Gaussian envelope'.

The figure below represents a snapshot (at $t = t_0$) of a wave packet from a light source.



The wavepacket would have a similar form if drawn as the magnitude of **E** v t at fixed $x = x_0$ or **E** v x at fixed $t = t_0$. At fixed $x = x_0$, the Fourier frequency spectrum is therefore the convolution of a delta function at frequency ω and the Fourier transform of the Gaussian distribution, *i.e.* another Gaussian distribution. The width of this frequency spectrum is inversely proportional to the time for which the wave packet can be considered as a sinusoidal disturbance.

In optics, the term *coherence time* is used - the period over which the wave-packet can be considered to be (approximately) sinusoidal. The corresponding length - $\Delta x = c \times \Delta t$ - is the coherence length.

White light has a frequency range from about 0.4×10^{15} Hz to about 0.7×10^{15} Hz (400nm - 700nm). This bandwidth of about 0.3×10^{15} Hz gives a very small coherence time - *large frequency range* $\Longrightarrow$ *small coherence time* (the bandwidth theorem) - of about 3×10^{-15} seconds, *i.e.* wave packets only a few wavelengths long.

In the case of light from a quasi-monochromatic source, *e.g.* a sodium discharge tube, the coherence length is a few millimetres and the number of oscillations is approximately 10^4. For the red cadmium line, the coherence length is about 50 cm, the number of oscillations is about 10^6 and the linewidth is about 10^{-12} m.

For a He-Ne laser, the linewidth can be about 10^{-15} m and the coherence length up to several metres.

2.3 Vector Wave Theory

A satisfactory theory of light must include a wave-like description and the work of James Clerk Maxwell showed that the periodically varying quantities are electric and magnetic fields. The directions of the electric and magnetic fields are orthogonal. The **E** and **B** are linked by Maxwell's equations and each satisfy the wave equation

$$\nabla^2 \mathbf{E} = \varepsilon_0 \mu_0 \frac{\partial^2 \mathbf{E}}{\partial t^2}$$

$$\nabla^2 \mathbf{B} = \varepsilon_0 \mu_0 \frac{\partial^2 \mathbf{B}}{\partial t^2}$$

$$E_y = cB_z$$

The Poynting vector describes the flow of energy,

$$\mathbf{S} = c^2 \varepsilon_0 \mathbf{E} \times \mathbf{B}$$

$$c = \frac{1}{\sqrt{(\varepsilon_0 \mu_0)}}, \quad v = \frac{1}{\sqrt{(\varepsilon \mu)}}, \quad n = \frac{c}{v} = \sqrt{(\frac{\varepsilon \mu}{\varepsilon_0 \mu_0})}$$

Transparent substances are usually non-magnetic, *i.e.* $\mu = \mu_0$ and $n \approx \sqrt{\frac{\varepsilon}{\varepsilon_0}}$.
ε is frequency dependent and therefore so is the refractive index.

Speed of light is defined to be 2.99792458×10^8 ms^{-1}.

2.4 Interaction of em-waves with Matter

Simple description When an electro-magnetic wave travels through a medium, the electric field causes the electron cloud of each molecule to oscillate. This oscillating charge re-radiates and causes some scattering of the light. The re-radiated light is of maximum amplitude at right angles to the axis of the oscillation and zero along this axis. *See the polar diagram for the radiation from a Hertzian dipole.*

2.5 Huygens' Principle

Every point on a primary wavefront acts as a source of spherical wavelets; the new wavefront at some later time is the envelope of these secondary waves. This fits with the simple description of the interaction of radiation with matter and we consider the secondary wavelets in terms of the radiation from Hertzian dipoles.

Note:- Huygens assumed that the secondary wavelets only propagate in the forwards direction. In the full theory, the obliquity factor $C(\mathbf{r})$ is zero for backward spreading secondary waves.

See p.104, Fig 4.27 of 'Optics' by Hecht for example.

See p.63, Fig 6.3 of 'Insight into Optics', for example, for diagrams of reflection and refraction in terms of Huygens' Principle.

See also **Fig 5.6 of 'Optics' by Hecht** for a diagram of the wavefronts of a light wave passing through a lens.

2.6 Brewster's Angle

When a light beam is reflected from a reflecting surface at an angle of incidence such that the angle between the reflected beam and the refracted beam is 90°, the reflected beam is linearly polarised. We could show this experimentally by arranging for a second reflection at the same angle of incidence on another plate. If the second plate is parallel to the first, the beam is reflected, but if the second plate is rotated through 90°, no reflection takes place. (*See The Experiment of Malus*)

$$\theta_i + \theta_r = 90^o$$

$$n_1 \sin\theta_i = n_2 \sin\theta_r = n_2 \cos\theta_i$$

$$tan\theta_i = n_2/n_1$$

The reason for this is that light propagates by interacting with the charges of the medium through which it is travelling; note that it can also travel through a vacuum. At the boundary between the two media, the reflected light and the refracted are generated by re-radiation from the forced oscillation of charges in the second medium. Oscillating charges do not radiate along the axis of oscillation.

Application:- Use Polaroid sun glasses to cut out the *linearly polarised* glare of the sunlight being reflected off still water at around Brewster's angle; you will then be able to see through the surface to the fish or whatever beneath.

3 Polarisation effects

3.1 Types of polarised light

- *Unpolarised* Light from 'normal' light sources is unpolarised - street lamps, desk lamps, torches, *etc*. The plane of oscillation of the electric field vector changes randomly, more frequently than every 10^{-16} seconds. Every plane of oscillation is present over any useful period of time.

- *Linear or plane polarised* The electric field vector oscillates in one plane only, *e.g.*

$$E_y = E_{0_y} \exp i(\omega t - kx)$$

- *Circularly polarised* Consider the following two light waves being present, travelling along the x-axis -

$$E_y = E_{0_y} \exp i(\omega t - kx) \text{ and } E_z = E_{0_z} \exp i(\omega t - kx + \frac{\pi}{2})$$

  If $E_{0_y} = E_{0_z}$, the tip of the resulting electric field vector will trace out a circle with angular frequency ω at any position on the x-axis.

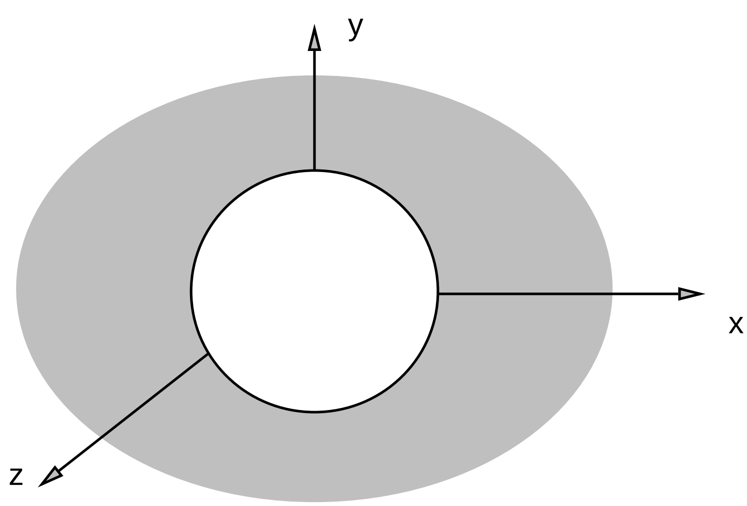
- *Elliptically polarised* If $E_{0_y} \neq E_{0_z}$, the tip of the resulting electric field vector will trace out an ellipse with angular frequency ω at any position on the x-axis.

4 The Mechanical Oscillator Model of Dielectrics

Consider what happens when light travels into a dielectric:-

- The electro-magnetic waves interact with the charges in the dielectric.
- The charges oscillate because of the action of the $\vec{E}$.
- The charges re-radiate - Huygen's Principle.
- The speed of travel depends on the refractive index, which in turn depends on the (inverse) of the force keeping the charges in position.
- Many substances are isotropic, *i.e.* the charges are bound in the same way in all directions $\Longrightarrow$ light propagates at the same speed in all directions.
- Some are not $\Longrightarrow$ interesting properties, *e.g.* dichroism.

One can derive an analytical expression for the refractive index n(ω) in terms of what happens within the dielectric at the atomic level. A simple, classical treatment can be found in *'Optics' by Hecht on pages 39 - 42 inc. and 230 - 231 inc.*.

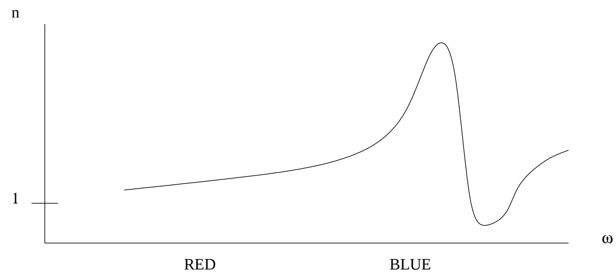


The treatment assumes that the charges in the dielectric are bound in place by an elastic restoring force proportional to the displacement from the equilibrium position; the force will be $m \times \omega_0^2 \times displacement$, where ω_0 is the natural frequency of oscillation.

A linearly, polarised electro-magnetic wave, *e.g.* $E_{y_x} = E_{0_y} \exp i(\omega t - kx)$ will cause the charge to oscillate along the y axis. If $\omega < \omega_0$, the charge oscillates in phase with the electro-magnetic wave but if $\omega > \omega_0$ then the oscillation is 180^o out of phase.

The oscillating charge $\Longrightarrow$ an electric dipole, the magnitude of which varies sinusoidally with time and therefore re-radiates. If the frequency of the electro-magnetic wave coincides with the natural frequency of the bound charges, absorption will take place.

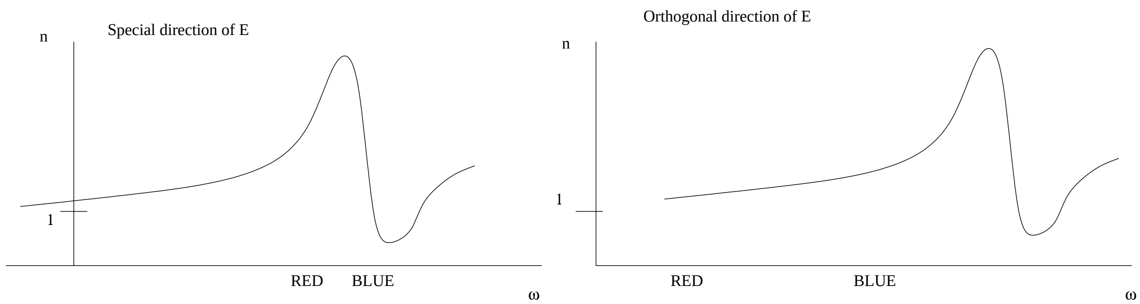
Colour of a dielectric:-



In the case shown above, the high frequency end of the visible spectrum is absorbed; the transmitted colour is red.

Dichroism:-

If the springs are not all the same strength, then absorption can take place for one orientation of the $\vec{E}$ of the electro-magnetic wave but need not for the orthogonal directions. This is the property of Polaroid sheet.



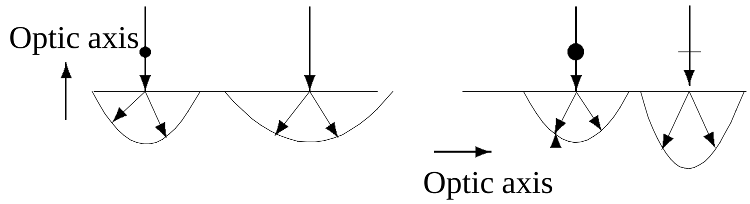
Demonstration Polaroid sheets, quarter- and half-wavelength plates. Stress-induced birefringence. Calcite crystal.

4.1 Uniaxial Crystals

If the springs in the x- and y-directions are of the same strength but different from the springs in the z-direction, then the z-direction will be special; the special direction is named the **optic axis** of the crystal.

Light waves
$E_x = E_{0x}e^{i(\omega t-kz)}$ and $E_y = E_{0y}e^{i(\omega t-kz)}$
travelling along the optic axis (the z-axis in this case) will behave in the same way but
$E_y = E_{0y}e^{i(\omega t-kx)}$ and $E_z = E_{0z}e^{i(\omega t-kx)}$
will behave differently; the refractive index experienced by the second (of the latter pair) will be different from that experienced by the first.

The diagrams below show the Huygen's wavelets spreading out from the surface.



5 Birefringence

Consider a uniaxial crystal cut as a thin sheet of thickness d, and with its optic axis parallel to the surfaces; the speed with which linearly polarised light will travel through the sheet will depend upon the orientation of the E-vector of the electro-magnetic wave; a phase difference will be introduced between the light with E parallel to the optic axis and the light with E perpendicular to the optic axis.

5.1 The Quarter Wave Plate

If the phase difference is $2\pi/4$, the plate is called a quarter wave plate; the condition for this is:-

$$(2\pi/\lambda)(n_o - n_e)d = 2\pi/4$$

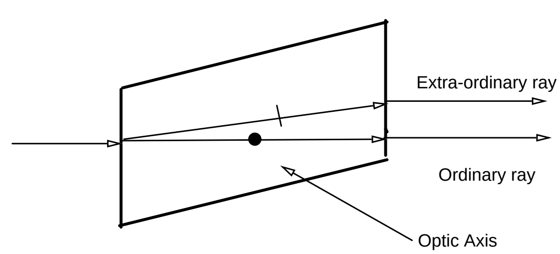
λ is the wavelength in vacuum, n_o is the refractive index of the light with E-vector perpendicular to the optic axis and n_e the ref. index for light with E parallel to the optic axis.

If linearly polarised light with E-vector at 45^o to the optic axis travels through such a plate, it will emerge as circularly polarised light.

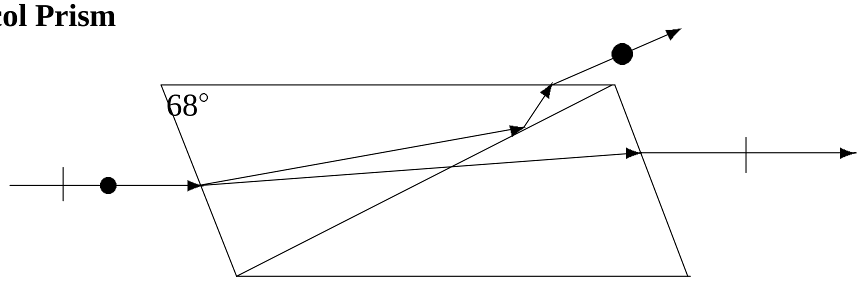
5.2 The Half Wave Plate

Two of the above plates will introduce a phase difference of π. Linearly polarised light with E at angle θ to the optic axis will emerge from such a plate with the plane of polarisation rotated through angle 2θ such that E is now at angle $-\theta$ to the optic axis.

5.3 Calcite Crystal

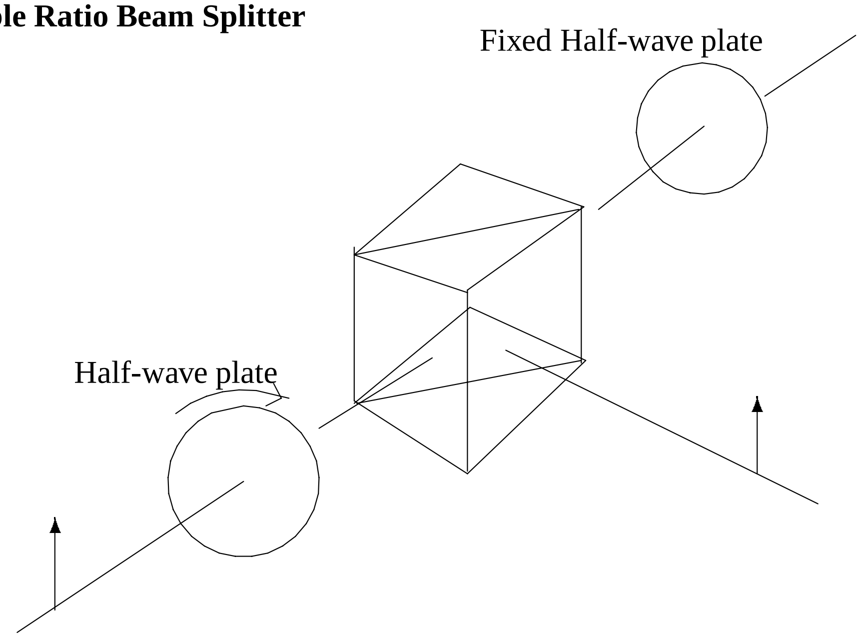


Nicol Prism



A calcite crystal is polished to the above shape, cut along the diagonal as shown and re-assembled by cementing the two parts together with canada balsam. This cement is transparent and has a refractive index mid-way between the ordinary and the extra-ordinary rays.

Variable Ratio Beam Splitter



Variable Ratio Beam Splitter

The variable ratio beam splitter could be used in a holography experiment - *i.e.* one in which one beam would be used as a reference beam, the other would be reflected from an object and then the two beams recombined to form interference fringes.

1. Start with plane polarised (vertical) light.
2. Pass through the first half-wave plate with the axis of the plate at angle θ.
3. Plane polarised light with plane at angle $2 \times \theta$ emerges.
4. Vertical component reflected at mid-plane of bi-prism.
5. Horizontal component transmitted.

6. At this stage, the two planes of polarisation are orthogonal, one being vertical and the other being horizontal. If these are to be combined to form interference fringes, they must have the same plane of polarisation.

7. Use a half-wave plate with axis at 45^o to rotate the plane of polarisation of the transmitted beam through 2×45^o.

6 Second Harmonic Generation

In the above (qualitative) description of the interaction of the electro-magnetic wave with matter, it was assumed that the response of the charges (mainly the outer, loosely bound valence electrons) was linear, *i.e.* the restoring force, $F(x)$ was

$$F(x) = -\frac{\partial V(x)}{\partial x} = -m\omega_0^2 x$$

If the crystal has non-linear properties, the restoring force on the charges has additional terms and the displacement of the charges is not described by simple harmonic motion, *e.g.*

$$F(x) = -\frac{\partial V(x)}{\partial x} = -(m\omega_0^2 x + mDx^2 + ...)$$

If $D > 0$, a positive displacement x results in a larger restoring force than a negative displacement. The re-radiated em wave contains a series of frequencies, the fundamental and strongest, ω_0, the second harmonic with frequency $2\omega_0$, *etc.* Note that the non-linearity is greater for large amplitude/ intensity em waves.

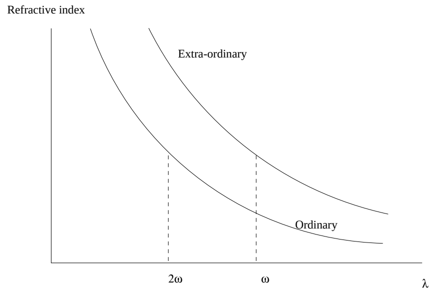
As the em wave travels through the non-linear crystal, the fundamental and the second harmonic waves will be generated at each position. The second harmonic wave will have a significant amplitude/ intensity only if the velocity of the second harmonic and the fundamental are the same - *see below*, *i.e.* the refractive indices must be the same for both the fundamental and the second harmonic. Otherwise the second harmonic waves will not add constructively.

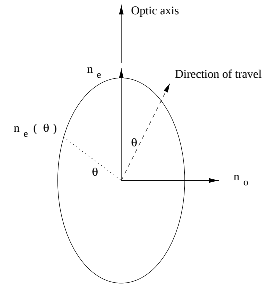
The fundamental wave, $Be^{i(\omega t - nkx)}$, travels through the medium generating the second harmonic continuously.

At $x = 0$, the electric field (the driving force) is $Ae^{i\omega t}$. This causes the charges to oscillate and generates a second harmonic wave $Be^{i(\omega' t - n' k' x)}$

At $x = x_P$, the driving force is $Ae^{i(\omega(t - \frac{nkx_P}{\omega}))}$
This generates another second harmonic wave starting at x_P and it is important that all of these second harmonic waves are in phase in the crystal and when leaving the crystal.

All media are dispersive and so the velocity of the primary wave and the second harminic wave cannot be the same for transmission through homogeneous media. However, this condition can be satisfied for some non-linear, birefringent crystals by choosing the angle of propagation through the crstal.





For a propagation direction at angle θ to the optic axis, the refractive index, $n_e(\theta)$, is

$$\frac{1}{(n_e(\theta))^2} = \frac{\cos^2\theta}{(n_0)^2} + \frac{\sin^2\theta}{(n_e)^2}$$

Can arrange for $n_e^{\omega}(\theta) = n_0^{2\omega}$

$$\frac{\cos^2\theta}{(n_0^{\omega})^2} + \frac{\sin^2\theta}{(n_e^{\omega})^2} = \frac{1}{(n_e^{\omega}(\theta))^2} = \frac{1}{(n_0^{2\omega})^2}$$

The fundamental beam passes through the crystal as an extra-ordinary ray at angle θ to the optic axis and generates the second harmonic as an ordinary ray in the same direction and with the same velocity. Normally, a ray would only generate another ray with the same polarisation state but there is coupling between the two polarisation states corresponding to the ordinary and extra-ordinary directions in the non-linear crystal.

Solve the above for the angle θ.

7 Optical Activity

For some uniaxial crystals, a linearly polarised plane wave travelling along the optic axis of the crystal is unchanged; for other crystals, the orientation of the $\vec{E}$ is rotated continuously as the wave travels through the crystal. This is the phenomenon of *optical activity*.

The degree of rotation is proportional to the path length in the crystal; the **specific rotatory power** gives the rotation per unit length.

Some crystals rotate the $\vec{E}$ clockwise, *i.e.* an observer watching the light coming towards him/her ‘sees’ the plane of the $\vec{E}$ oscillating in a new plane, one which is rotated clockwise relative to the initial plane of the $\vec{E}$; other crystals rotate the plane anti-clockwise.

Demonstration Plane polarised laser beam propogating through a long tube of syrup.

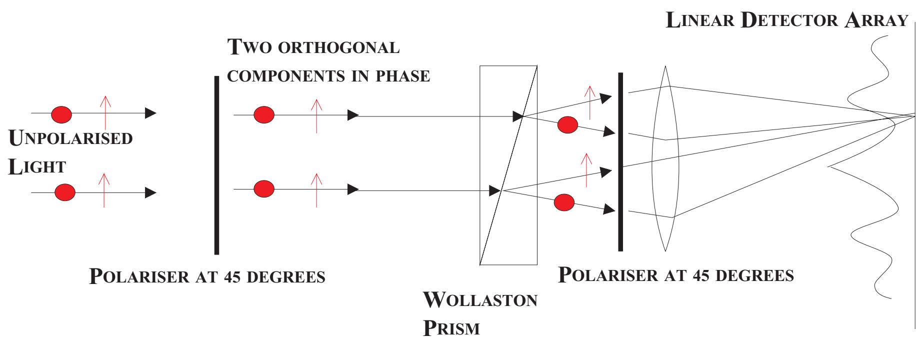
Quartz exists in both forms and 1 mm thickness of quartz rotates the plane of the $\vec{E}$ of yellow Na light by 21.7^o. The characteristic property of optically active crystals is that they have a helical arrangement of molecules along the optic axis of the crystal. Liquids can also cause the plane of the E-vector of linearly polarised light to be rotated; the molecules have individual helix shapes and although randomly oriented, the light experiences an average helical effect because the handedness of a helix is the same from both ends. Note that in liquids there is no optic axis.

Optical activity is most easily described in terms of the crystals and liquids which exhibit this phenomena having two refractive indices: - one for right-handed circularly polarised light and a different one for left-handed circularly polarised light.

The angle through whichthe plane of oscillation is turned is

$$\alpha = (\frac{\pi}{\lambda})(n_R - n_L)z$$

8 Fourier Transfrom Spectrometer (*with no moving parts*)



The birefringence in the Wollaston prism introduces a path difference between the beams with the two orthogonal polarisations. The effective path difference between these beams varies across the prism and the beams are superimposed across the linear array of detectors after changing the beams back to the same polarisation state. The detectors record the same information as in the 'Twyman-Green' Fourier Transform Spectrometer (*see later*) but the information is recorded without using any moving mirrors.

For one component of of the plane polarised beam, the optical path through the Wollaston Prism is

$$n_e \frac{y}{D} L + n_o (L - \frac{y}{D} L)$$

and the OPL for the other component is

$$n_o \frac{y}{D} L + n_e (L - \frac{y}{D} L)$$

The difference in OPL of the light reaching the linear array of detectors is

$$\Delta = 2(n_e - n_o) L (\frac{y}{D} - 1)$$

i.e. a function of y.

Suppose that the light has intensity $B(k)dk$ in the range k to $k + dk$ in each of the orthogonal modes of oscillation.

If now an optical path difference, Δ, is present this will introduce a phase difference of $(\frac{2\pi}{\lambda})\Delta = k\Delta$. The resultant intensity falling on the photo-detector is

$$dI = 2B(k)dk(1 + \cos(k\Delta))$$

The total intensity, considering all wave numbers, is

$$I(\Delta) = 2\int_0^\infty B(k)dk + 2\int_0^\infty B(k)\cos(k\Delta)dk$$

The first term is a constant and the second varies as the optical path difference varies. The second term is the Fourier (cosine) transform of the spectral distribution in the emergent beams - *strictly, the response of the photo-detector is folded in with this, but the above assumes that this is independent of wave number.*

If

$$\Phi(\Delta) = 2\int_0^\infty B(k)\cos(k\Delta)dk$$

$$B(k) = 2\int_0^\infty \Phi(\Delta)\cos(k\Delta)d\Delta$$

If one measures $I(\Delta)$ over a wide range of Δ, one can calculate the spectral distribution of the source ($B(k)$) by calculating the Fourier transform of the $I(\Delta)$ function obtained from the distribution of response of the photo-detectors. (Strictly it should be $\Phi(\Delta)$ that is used but the difference is only a constant which transforms to a delta-function at the origin.)

If the source emits a narrow, single wavelength, then the intensity distribution shows a "cosine" periodic variation across the photo-detectors.

In the case of the Twyman-Green Fourier Transform Spectrometer, the phase difference is changed by moving one of the two mirrors. In this recently patented system, there are no moving parts; the phase difference varies across the set of detectors. The Fourier Transform of the information recorded by the linear array gives the distribution of wavelengths in the incident light beam.

8.1 Fourier Transform

Using the above notation, the Fourier Transform of the function $B(k)$ would be written as

$$\Phi(\Delta) = \int_{-\infty}^\infty B(k)\exp(i(k\Delta))dk$$

Note that

$$\exp(i(k\Delta)) = \cos(k\Delta) + i\sin(k\Delta)$$

If the function $B(k)$ is even, *i.e.* $B(k) = B(-k)$, then

$$\int_{-\infty}^\infty B(k)\sin(k\Delta)dk = 0$$

Similarly, because $\cos(k\Delta)$ is even,

$$\int_{-\infty}^{\infty} B(k)\cos(k\Delta)dk = 2\int_{0}^{\infty} B(k)\cos(k\Delta)dk$$

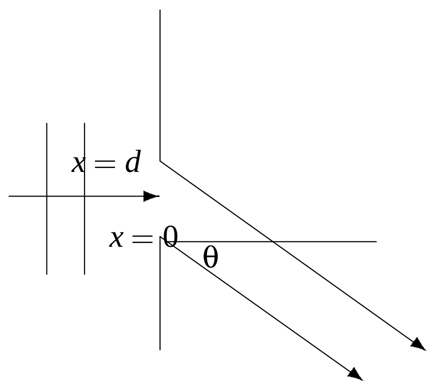
The "cosine" Fourier transform of an even function is the same as the "exponential" Fourier transform.

9 Diffraction

When a beam of light is partially obstructed by an obstacle, some of the light is diverted sideways; this spreading of the light is known as diffraction. There are two types of diffraction, Fresnel, where source or/and screen are at a finite distance from the diffracting aperture and Fraunhofer, where source and screen are effectively at infinity. The experimental set-up is often to have the source at one focal length from a lens and a lens one focal length in front of the screen.

Demonstrations Diffraction pattern of single slit of variable width, two slits, three slits, four slits, etc. Circular aperture.

Consider Fraunhofer diffraction at a single slit first:-



Consider an element of width dx at position x measured from the bottom of the slit. This element will act as a source of secondary wavelets; consider those leaving in direction θ. Let the amplitude of these be a_0; the phase, relative to the wavelets leaving the bottom of the slit, will be δ where

$$\delta = (2\pi/\lambda)(x\sin\theta)$$

This is because of the extra path length travelled by the ray from the element at position x, compared with the path of the ray from the bottom of the slit. The total amplitude leaving the slit at angle θ is

$$A(\theta) = \int_0^d a_0 e^{ikx\sin\theta}dx$$

$$A(\theta) = [(a_0 e^{ikx\sin\theta})/ik\sin\theta]_0^d$$

$$A(\theta) = a_0(e^{ikd\sin\theta} - 1)/(ik\sin\theta)$$

$$A(\theta) = \frac{a_0}{ik\sin\theta}(e^{\frac{1}{2}ikd\sin\theta} - e^{-\frac{1}{2}ikd\sin\theta})e^{\frac{1}{2}ikd\sin\theta}$$

Let $\beta = \frac{1}{2}kd\sin\theta$;

$$A(\theta) = \frac{a_0 d}{\beta}\frac{(e^{i\beta} - e^{-i\beta})}{2i}e^{i\beta}$$

$$A(\theta) = a_0 d(\frac{\sin\beta}{\beta})e^{i\beta}$$

Therefore

$$I(\theta) = A(\theta)^* A(\theta) = I_{inc}(\sin\beta/\beta)^2$$

where $\beta = \frac{1}{2}kd\sin\theta$.

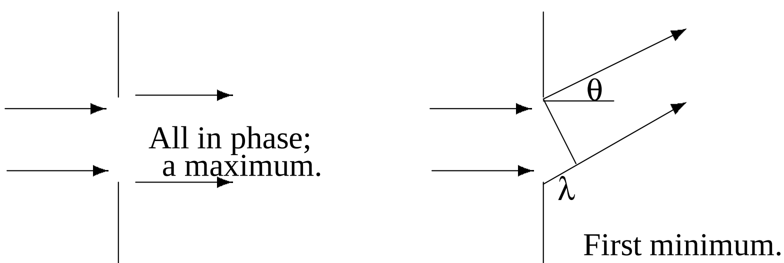
For maxima and minima, differentiate w.r.t. β;

$$\frac{dI}{d\beta} = 2(\sin\beta/\beta)(\cos\beta/\beta - \sin\beta/\beta^2) = 0$$

For minima, the first term is zero, *i.e.* $\sin\beta = 0$ or $\beta = n\pi$, where n=1, 2, 3, 4, 5, ...
N.B. $\sin\beta/\beta$ is not zero for $\beta = 0$; in fact $\beta = 0$ corresponds to a maximum.

For maxima, the second term is zero, *i.e.* $\beta = \tan\beta$. This is most easily solved graphically by plotting function $g = \beta$ and $g = \tan\beta$ together and finding the intercepts. The maxima are found where $\beta \approx (2n+1)(\pi/2)$, where n=1, 2, 3, 4, 5, ...

9.0.1 Understanding without Mathematics



Each element in top half cancels
the corresponding element in the top half.

9.0.2 A Rectangle

In the above, the slit has been considered a "very long" so that only one dimension has been considered. For a finite sized slit or rectangle, an element of size $dxdy$ at position x and y is first considered and the light leaving this element at angles θ_x and θ_y then integrated over the whole rectangle.

$$A(\theta_x, \theta_y) = \int_0^{d_x}\int_0^{d_y} a_0 e^{ikx\sin\theta_x} e^{iky\sin\theta_y} dxdy$$

Obviously this is just two one-dimensional integrations much as above -

$$I(\theta_x, \theta_y) = I_0(\sin\beta_x/\beta_x)^2(\sin\beta_y/\beta_y)^2$$

where $\beta_x = \frac{\pi}{\lambda} d_x \sin\theta_x$ and $\beta_y = \frac{\pi}{\lambda} d_y \sin\theta_y$.

All real slits are of finite size and the above shows that the diffraction pattern is more spread out in the direction of the smaller slit direction, an inverse relationship.

9.0.3 Width of Pattern

The width of the pattern between the zeros on either side of the central maximum is easily calculated. For these zeros, $\beta = \pm\pi$ and therefore $d \sin\theta = \lambda$, for a slit width of d. In many cases the angles are small and $\sin\theta \approx \theta$; then $d \times \theta \approx \lambda$.

This shows that the pattern is wider if the slit is smaller and is wider for larger wavelengths.

9.0.4 Experimental Arrangements

The mathematical analysis above assumes that a plane wavefront is incident on the slit or rectangle. In this case, the diffraction pattern has the inverse relationship between sizes, *i.e.* the pattern is spread out in the 'narrow' direction of the slit or rectangle and is very narrow in the 'long' direction of the slit or rectangle.

One way of achieving the plane wavefront is to have a point source in the focal plane of a lens and to use the light from the lens to illuminate the slit. An easier way is to use a laser beam, or an expanded parallel laser beam to illuminate the aperture.

In many experimental setups, the source used is a line source in the focal plane of the lens above and not a point source; the line source is parallel to the slit. This corresponds to having a set of independent point sources along the y-axis. Experimentally, the diffraction pattern has the predicted distribution in the x-direction and this distribution is the same for all values of y. A ray diagram will show that each point on the line source gives rise to a diffraction pattern at a different distance (in the y-direction) from the axis of the system and in accord with what is seen experimentally.

9.0.5 Fraunhofer Diffraction and the Fourier Transform

In the handout for the Physical Mathematics module, you were shown that the (amplitude) Fraunhofer diffraction pattern of a transmitting screen is the (one dimensional) fourier transform of the transmission or aperture function, $f(x)$, of the screen. The top hat function or single slit was used to illustrate,

$$F\left[f(x)\right] = F(\kappa) = \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{\infty} f(x) e^{i\kappa x} dx$$

In the above, the κ is the spatial frequency with dimensions L^{-1}.

The Fraunhofer diffraction distribution is

$$A(\theta) = \int_0^d a_0 e^{ikx\sin\theta} dx$$

see above, and can be written as

$$A(\theta) = \int_{-\infty}^{\infty} f(x)e^{ivx}dx$$

where $k\sin\theta = \frac{2\pi}{\lambda}\sin\theta$ has been written as v and $f(x)$ is zero except for $0 \le x \le d$.
i.e.

$$F(v) = \int_{-\infty}^{\infty} f(x)e^{-ivx}dx = \int_0^d a_0 e^{-ivx}dx = a_0\frac{1}{-iv}(e^{-ivd}-1) = a_0 d(e^{-iv\frac{d}{2}})(\frac{\sin(\frac{vd}{2})}{\frac{vd}{2}})$$

Now v is the spatial frequency.

The "**lens**" is often described as a fourier transformer.

9.0.6 Shift Invariance

What happens if the slit is shifted sideways, *i.e.* shifted off axis?

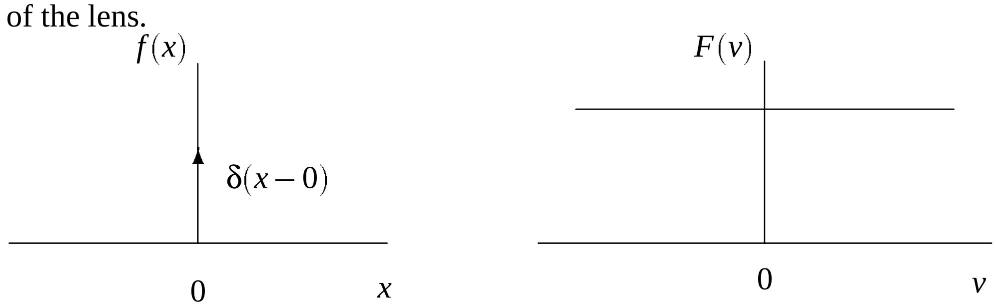
All rays leaving the diffracting rectangle in either position at angle θ to the axis are brought to a focus at the same point on the screen; it could therefore be concluded that there is no difference in the diffraction pattern due to a slit on axis or off axis. In fact, there is no difference in intensity and that is why no care is needed in lining up the slit with the axis of the lens in a diffraction experiment. However, a ray diagram will show that only rays leaving at $\theta = 0$ travel the same distance from the slit on axis as off axis; for other angles, the path difference is $d_{shift}\sin\theta$ and therefore the distribution from a slit on axis differs in phase from that of a slit off axis - a phase gradient difference.

10 More on Fourier Transforms

The lens as a Fourier transformer.

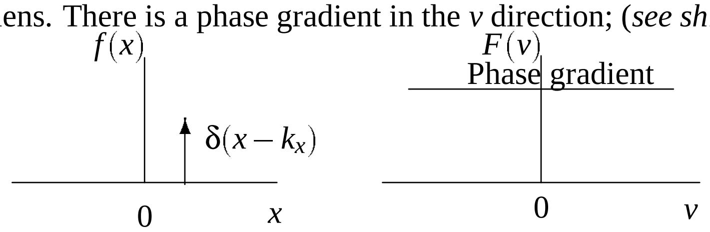
The top hat aperture function or rectangular aperture is not the simplest transmission function to deal with. Here are some other simple ones.

1. A point source at $x = 0$, *i.e.* a δ-function, in the object plane.
   Object plane of $\delta(x-0) \Longrightarrow$ Uniform illumination in the fourier plane over an area defined by the area of the lens.



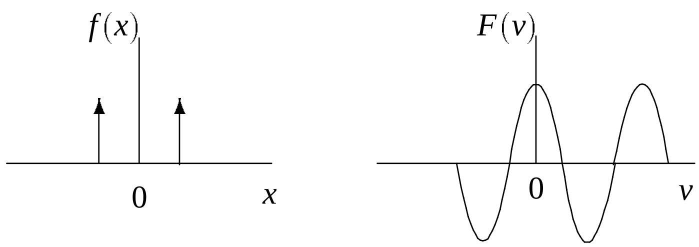
$$F\,[f(x] = C\int_{-\infty}^{\infty}\delta(x-0)e^{-ivx}dx = C$$

2. A point source at $x = d_x$, *i.e.* a δ-function at $x = d_x$ in the object plane.
   Object plane is $\delta(x-d_x) \Longrightarrow$ Uniform illumination in the fourier plane over an area defined by the lens. There is a phase gradient in the v direction; (*see shift invariance* above).



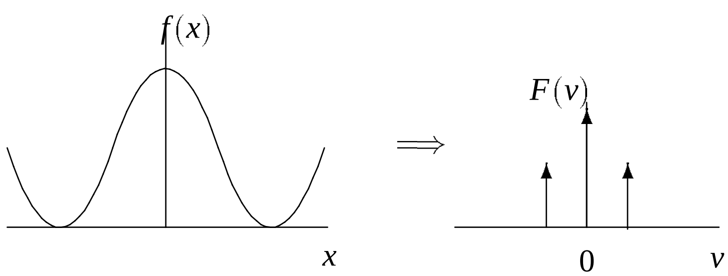
$$F\,[f(x] = C\int_{-\infty}^{\infty}\delta(x-d_x)e^{-ivx}dx = Ce^{-id_x v}$$

3. Two very narrow slits in the object plane at $x = \pm\frac{d_x}{2}$, *i.e.* symmetrically about the origin.



$$F\,[f(x] = C\int_{-\infty}^{\infty}(\delta(x-\frac{d_x}{2})+\delta(x+\frac{d_x}{2}))e^{-ivx}dx = C(e^{-ivd_x/2}-e^{ivd_x/2}) = 2C\cos(\frac{vd_x}{2})$$

4. A "cosine", amplitude transmission function. In fact only a "$(1+\cos\ )$" transmission function is possible, *e.g.* f(x) = $(1+\cos(\frac{d_x x}{2}))$.



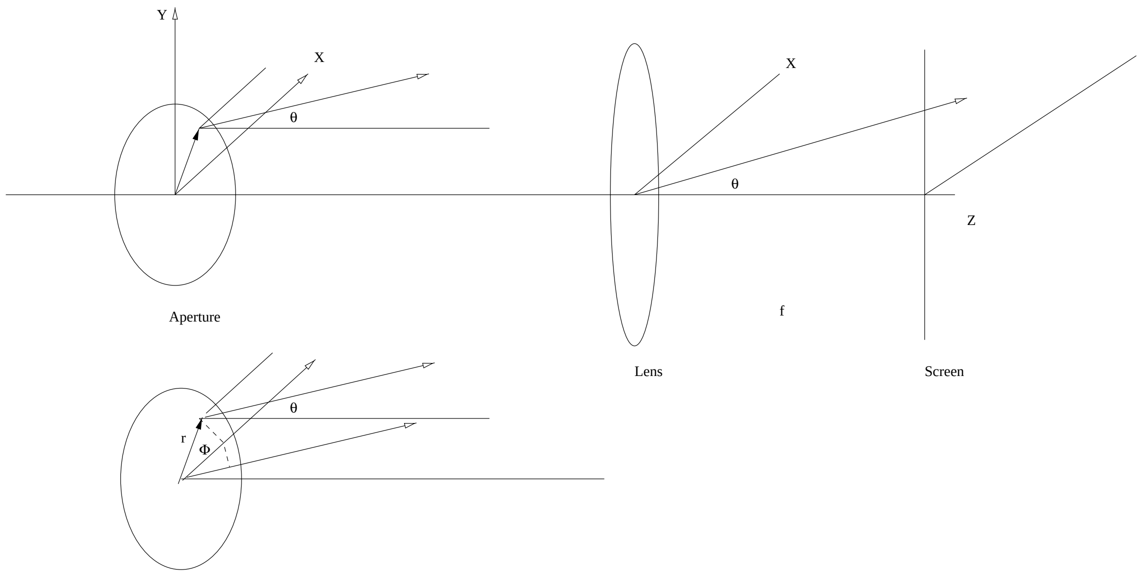
$$F\,[f(x] = C\int_{-\infty}^{\infty}(1+\cos(\frac{d_x x}{2}))e^{-ivx}dx$$

$$F\,[f(x] = C\int_{-\infty}^{\infty}(1+0.5(e^{i\frac{d_x x}{2}}-e^{-i\frac{d_x x}{2}}))e^{-ivx}dx$$

$$F\,[f(x] = \delta(v-0)+(0.5\delta(v-\frac{d_x}{2})+0.5\delta(v+\frac{d_x}{2}))$$

11 Circular aperture - The Airy Pattern

The Fraunhofer diffraction pattern, or Airy pattern, of a circular aperture is of great practical importance because the limiting aperture in most optical instruments is circular. Sir George Biddell Airy, Astronomer Royal *(for England)* first derived an expression for the distribution of intensity.



Consider an element of the circle defined by polar coordinates $r \to r+dr$, $\phi \to \phi + d\phi$. The area of this element is $rdrd\phi$.

The axial symmetry of the system and of course of the expected intensity pattern allows one to consider rays leaving in planes parallel to the xz-plane at angle θ to the axis only. The phase difference between a ray at angle θ from the element above and one from the centre of the aperture is $kr\cos\phi\sin\theta$.

The amplitude on the screen at a point a distance $f \times \tan\theta$ from the axis is given by the double integral over r and over ϕ to cover the whole circle.

$$A(\theta) = B\int_0^R \int_0^{2\pi} re^{ikr\sin\theta\cos\phi} drd\phi$$

where B = constant.
Let $b = kr\sin\theta$ and $db = k\sin\theta dr$; the new upper limit is $\beta_A = kR\sin\theta$.

$$A(\theta) = B\int_0^{\beta_A} \int_0^{2\pi} b/(k\sin\theta)^2 e^{ib\cos\phi} d\phi db$$

The integral w.r.t. ϕ is the *Zeroth-order Bessel function*

$$(J_0(b) = \int_0^{2\pi} e^{ib\cos\phi} d\phi)$$

$$A(\theta) = const/(k\sin\theta)^2 \int_0^{\beta_A} bJ_0 db$$

One of the properties of Bessel functions is that

$$\int_0^{\beta_A} bJ_0 db = \beta_A J_1(\beta_A)$$

where $J_1(\beta_A)$ is the *First-order Bessel function*.

$$A(\theta) = (const/k^2 sin^2\theta)\ \beta_A J_1(\beta_A)$$

$$A(\theta) = const\ R^2\ J_1(\beta_A)/\beta_A$$

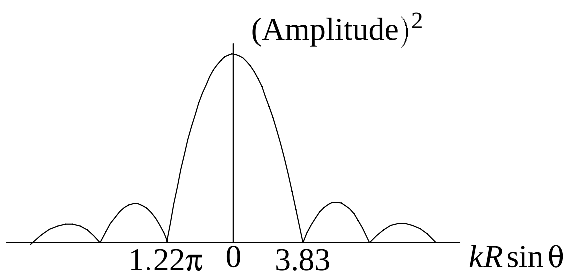
The First-order Bessel function is tabulated in Table 10.1 of Optics by 'Hecht'.

It is convenient to normalise this expression to the amplitude at the centre.

$J_1(\beta_A)/\beta_A \to 0.5$ as $\beta_A \to 0$, therefore $A(\theta) = 2A_0 J_1(\beta_A)/\beta_A$
The first zero of the First-order Bessel function is at $\beta_A = 1.22\pi$, *i.e.*

$$R\sin\theta = 0.62 \times \lambda$$

$$sin\theta = 1.22\lambda/Diam$$



The complete pattern is formed by rotating the above about the central ordinate and consists of a bright central disc (called the Airy disc) surrounded by rings of rapidly decreasing intensity.

The angle at which the first minimum occurs is

$$\sin\theta = 1.22\lambda/diameter$$

Demonstrations

Diffraction pattern of circular aperture
Diffraction pattern of single slit, variable width.
Diffraction pattern of single slit, two slits, three slits and four slits showing principal and secondary maxima.

11.0.7 Resolution limits

Rayleigh Criterion

Equal intensity images are 'resolved' if the main maximum of one falls no closer than the first minimum in the diffraction pattern of the other.

The diffraction pattern in the case of most optical instruments is that of a circular aperture. The angle separation of two point sources viewed by a telescope is therefore $\sin\theta_{min} = 1.22\lambda/diam$, where $diam$ is the size of the limiting circular aperture. Note that because the two sources are incoherent, the intensities add, not the amplitudes.

The resolving power of the system is the inverse of this.

11.1 Convolution

(link with the Physical Mathematics module)
The *convolution* of two functions, $f(x)$ and $g(x)$, is defined to be

$$f(x) * g(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x')g(x-x')dx'$$

One of the most important results is that convolving $f(x)$ with $\delta(x-a)$ has the effect of putting a copy of the shape of f at $x = a$.

$$\delta(x-a) * f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \delta(x'-a)f(x-x')dx' = \frac{1}{2\pi}f(x-a)$$

11.2 The Diffraction Grating

Consider N slits of width d and slit separation a, consider corresponding elements of width dx on each slit at a distance x from the bottom of each slit. The amplitude contribution from these elements is

$$\text{Amplitude} = a_0 e^{ikx\sin\theta}dx + a_0 e^{ik(x+a)\sin\theta}dx........a_0 e^{ik(x+(N-1)a)\sin\theta}dx$$

$$A(\theta) = a_0(1 + e^{i\alpha} + e^{i2\alpha} + e^{i3\alpha}.....e^{i(N-1)\alpha})\int_0^d e^{ikx\sin\theta}dx$$

where $\alpha = (2\pi/\lambda)a\sin\theta$.

$$A(\theta) = a_0 d(1 + e^{i\alpha} + e^{i2\alpha} + e^{i3\alpha}.....e^{i(N-1)\alpha})(\frac{\sin\beta}{\beta})e^{i\beta}$$

Multiply by $e^{i\alpha}$

$$A(\theta)e^{i\alpha} = a_0 d(e^{i\alpha} + e^{i2\alpha} + e^{i3\alpha}.....e^{i(N)\alpha})(\frac{\sin\beta}{\beta})e^{i\beta}$$

Subtract

$$A(\theta)(e^{i\alpha}-1) = a_0 d(e^{iN\alpha}-1)(\frac{\sin\beta}{\beta})e^{i\beta}$$

$$A(\theta)(e^{i\alpha/2}-e^{-i\alpha/2})e^{i\alpha/2} = a_0 d(e^{iN\alpha/2}-e^{-iN\alpha/2})e^{iN\alpha/2}(\frac{\sin\beta}{\beta})e^{i\beta}$$

$$A(\theta)e^{i\alpha/2}\sin(\alpha/2) = a_0 d\sin(N\alpha/2)e^{iN\alpha/2}(\frac{\sin\beta}{\beta})e^{i\beta}$$

$$A(\theta) = a_0 d(\frac{\sin\beta}{\beta})e^{i\beta}(\frac{\sin(N\alpha/2)}{\sin(\alpha/2)})e^{i(N-1)\alpha/2}$$

$$I(\theta) = A^*A = I_0(\frac{\sin\beta}{\beta})^2(\frac{\sin(N\alpha/2)}{\sin(\alpha/2)})^2$$

The form of this is the product of the single slit distribution and the effect of having N slits.

Alternatively

The diffraction grating can be considered as the convolution of a single slit and a comb of δ-functions, $\delta(x-na)$, $n = 0,1,2,3,.....(N-1)$. The fourier transform is the product of the fourier transforms of the individual functions, *i.e.*

$$A(\theta) \propto (\frac{\sin\beta}{\beta})(1+e^{-iva}+e^{-iv2a}.....e^{-iv(N-1)})$$

Now $v = -\frac{2\pi}{\lambda}\sin\theta$ and so the series is the same as above and

$$A(\theta) = 2a_0 d(\frac{\sin\beta}{\beta})(\frac{\sin(N\alpha/2)}{\sin(\alpha/2)})e^{i(N-1)\alpha/2}$$

Note that the origin is effectively in the centre of the first slit and so the factor $\exp(i\beta)$ does not now appear in the expression for $A(\theta)$.

11.2.1 Maxima

The second term will vary faster than the first and the positions of the maxima will be determined by the second term; this has the "value" $0/0$ when $\alpha/2 = m\pi$, m= integer. Let $\alpha/2 = m\pi+\varepsilon$, where ε is small.

$$(\frac{\sin(N\alpha/2)}{\sin(\alpha/2)})^2 = (\frac{\sin(\pi mN + N\varepsilon)}{\sin(\pi m+\varepsilon)})^2 = (\frac{\sin(N\varepsilon)}{\sin\varepsilon})^2 \approx (N\varepsilon/\varepsilon)^2 \approx (N)^2$$

For $\alpha/2 = m\pi$, $I = N^2 I_0$.
N.B. the positions of the maxima depend only on a, not on N or d.

11.2.2 Width of the Principal Maxima

This will be determined by the second term. As above, the maxima are at $\alpha/2 = m\pi$; the nearest minima are at $\alpha/2 = m\pi \pm \pi/N$.

$$(I = I_0 \frac{(\sin(Nm\pi \pm \pi))^2}{(\sin(m\pi \pm \pi/N))^2}) = I_0(\sin\pi)^2/(\sin(\pi/N))^2 = 0$$

Therefore the width of the principal maxima decreases as N increases.

The above can be written as

$$(\alpha/2)_{min} - (\alpha/2)_{max} = (m\pi \pm \pi/N) - m\pi$$

i.e.

$$\sin\theta_{min} - \sin\theta_{max} = \pm(\lambda/Na)$$

again showing that the width decreases as the number of slits increases.

11.2.3 Number of Orders

The principal maxima are at angles which satisfy $a\sin\theta = m\lambda$ where m is the order of the maximum and is an integer. The maximum order is therefore limited because $\sin\theta$ has a maximum value of 1 and m_{max} is the nearest integer below a/λ.

11.2.4 Missing Orders

When the ratio of the slit width and the slit separation is an integer or a ratio of integers, the minimum of the single slit term will exactly coincide with some of the principal maxima of the N-slit term, *e.g.* if the ratio is 5, orders 5, 10, 15 ... will be missing; if the ratio is 7/2, the 14th order will be missing.

11.2.5 Chromatic Resolving Power

Using the Rayleigh Criterion again, if the maximum for wavelength $\lambda + \Delta\lambda$ is at the same angle as the minimum for λ, the two wavelengths will be just resolved.

$$Na\sin\theta = mN(\lambda + \Delta\lambda) = (mN + 1)\lambda$$

$$\lambda/\Delta\lambda = mN$$

Again this makes sense because one would expect that the resolving power should increase as the total number of slits increases and experiment shows that the zeroth-order contains all the colours present in the light source at the $\theta = 0$ angle, *i.e.* no resolving power.
N.B. the resolving power increases as m increases.

The maximum order, *i.e.* maximum value of m, can be increased by almost a factor of two by having the light incident on the grating at grazing angle.

11.2.6 Free Spectral Range

There is only a limited range of orders with no overlap; at some value of m the (m+1)th order of wavelength λ will coincide with the mth order for wavelength $\lambda + \Delta\lambda$. Care must then be taken to interpret the spectra correctly.

$$(m+1)\lambda = m(\lambda + \Delta\lambda)$$

i.e.

$$\Delta\lambda_{fsr} = \lambda/m$$

11.2.7 Blazed Grating

For a normal diffraction grating, the greatest intensity goes into the zeroth-order and there is no resolving power in this order. By using a reflecting grating at normal incidence but with each reflecting strip set at an angle to the plane of the grating, the greatest intensity can be directed to a useful order. By setting the reflecting strips at an angle, the positions of the principal maxima remain unchanged but the single-slit pattern is centred at a position predicted by the law of reflection.

The blazed grating can be considered as the convolution of a set of δ-functions and a single reflecting surface at an angle, *i.e.* the surface introduces a phase gradient.

Instead of a transmission function, the equivalent is the amplitude reflection function (arf). For the blazed grating,
arf = (top hat function) $\times$ (phase gradient) $\otimes$ (Comb)
The Fraunhofer diffraction pattern will have the form
(FT((top hat function) $\times$ (phase gradient))) $\times$ (FT(Comb))
i.e.
(FT(top hat function) $\otimes$ FT(Phase gradient)) $\times$ (FT(Comb))
$(\frac{\sin\beta}{\beta}) \otimes (\delta\text{-function}) \times (\frac{\sin\frac{N\alpha}{2}}{\sin\frac{\alpha}{2}})$

The delta-function reproduces the single slit distribution at a shifted position and the maximum intensity shifts to a useful order.

11.3 The Zone-plate

Consider a plane wave front approaching a circular aperture, the aperture being divided into circular zones such that the distance from the mth zone to a point P on the axis differs from the distance from the (m-1)th zone to the same point by $\lambda/2$. The radius R_m of the mth zone will be given by the Theorem of Pythagoras as:

$$R_m^2 = (f + m\lambda/2)^2 - f^2$$

i.e.

$$R_m^2 = mf\lambda + \frac{1}{4}m^2\lambda^2$$

If every second zone is blanked out, the path differences from the transparent zones to the point P are all an integer number of times λ; the light diffracted at the zones will all arrive at point P in phase and the zone plate will act as a converging lens.

The light intensity reaching the point P will be approximately $(N/2)^2$ times the light intensity if there is no zone structure present, where N is the number of zones. In fact this can be increased to approximately $(N)^2$ if instead of blanking out the alternate zones, these are covered with a layer of material which has an optical thickness of $\lambda/2$.

The zone plate will act as a diverging lens also because another set of diffracted rays will all be in phase (more precisely, differ by λ), *viz.* those rays apparently originating from the virtual focus on the side of the lens from which the original plane wave front is approaching.

N.B. The second term in the above equation is negligible compared with the first and the zone plate exhibits chromatic aberration because the design of the zones depends on the λ of the light used.

11.3.1 Application

Bifocal contact lenses may be based on this principle: these have what appears to be a zone plate formed on the inside surface of the lens and will therefore bring light from a distance to two different points of focus. The eye-lens changes shape to use one focal length or the other to bring the image of an object into sharp focus on the retina; the brain then ‘concentrates’ on this image and ignores the out-of-focus information formed by the other focal length lens.

12 Interferometry

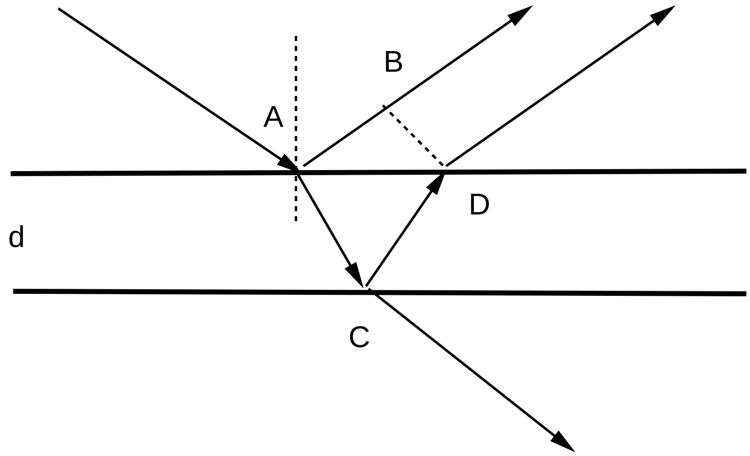
The work on diffraction gratings involved interference following division of wavefront.

12.1 Interference by Amplitude Division

The simplest example is the formation of localised fringes in the plane of a wedge film following reflection at the two surfaces of the wedge.

Demonstration White light fringes from a soap film.

In many arrangements, the two reflecting surfaces are parallel. Consider a thin film of refractive index n_2, the surrounding medium having refractive index of n_1. For a ray entering the film at angle i and refracted at angle r, there will be a path difference between the ray reflected from the top surface and the ray refracted then reflected from the lower surface. The *amplitude division* in the title takes place at A in the diagram.



The optical path difference (OPD) of the two rays shown when they reach the screen is:-

$$OPD = n_2 ACD - n_1 AB$$

$$OPD = 2n_2 AC - n_1 AB$$

$$OPD = 2n_2 AC - n_1 AD \sin i$$

$$OPD = 2n_2(d/\cos r) - 2n_1 d \tan r \sin i$$

Use Snell's Law, $n_2 \sin r = n_1 \sin i$.

$$OPD = 2n_2(d/\cos r) - 2n_2 d \frac{(\sin r)^2}{\cos r}$$

$$OPD = 2n_2 d \frac{(1-(\sin r)^2)}{\cos r}$$

$$OPD = 2n_2 d \cos r$$

N.B. This result *appears* to be independent of the refractive index of the surrounding medium but note that

$$\cos r = \sqrt{(}1 - (\frac{n_1}{n_2})^2 \sin i^2)$$

If $n_2 > n_1$, as has been assumed in the above diagram, there is a phase change of π on reflection at A and this must be added to the above OPD; the **total** OPD is then $2n_2 d \cos r + \frac{\lambda}{2})$.

If the total OPD is $m\lambda$, m integer, then constructive interference takes place; if the total OPD is $(m+1/2)\lambda$, then there is destructive interference. The fringes formed are **fringes of equal inclination** and are circular because they depend only on the angle r. The fringes are formed at infinity, hence the use of the lens to bring infinity closer.

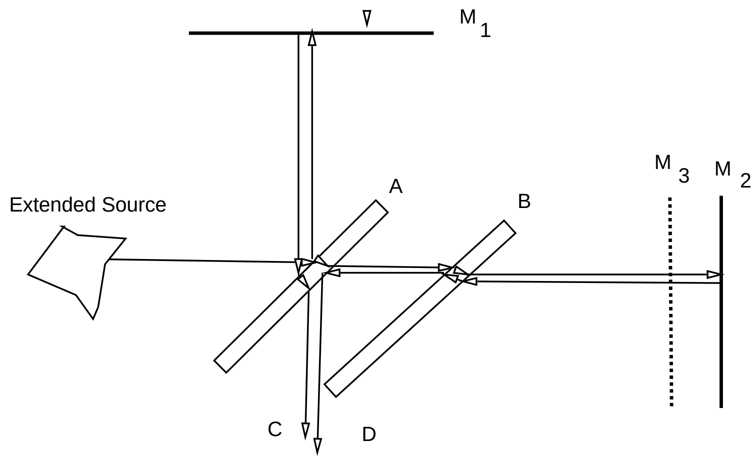
If a point source is used to illuminate the film, fringes are seen only over a small part of the screen; if, however, an extended source is used, fringes are seen over a larger region of the film. The fringes are also much brighter than in the case of a point source.

If the fringes are viewed at normal incidence, the fringes are known as Haidinger fringes. The above theory normally applies to the case of multiple reflections but first it will be applied to an instrument in which only single reflections take place, the Michelson Interferometer.

12.2 The Michelson Interferometer

In the Michelson Interferometer, the two surfaces of the thin film are separated in space.

An extended source is used; the surface at A is semi-silvered so that ≈ 50% of the light is reflected and ≈ 50% is transmitted; there is a compensating plate at B; M_1 and M_2 are reflecting mirrors and M_2 can be moved in the vertical direction. The mirrors are set up so that they are accurately at right angles to each other. The **effective** position of M_1 is at M_3 and therefore thin film interference takes place. The *optical path difference* is $2nd\cos r$, as above, and the fringes are circular and easily seen if monochromatic light is used.



The compensating plate is present to make the distance travelled in glass the same for both beam paths; this is absolutely necessary for there to be any possibility of obtaining *white light fringes*, another requirement for obtaining such fringes being that the path difference is less than a few λ's. This is because the dispersion of the glass will otherwise make it impossible to adjust to near zero path difference for more than one λ and the fringes set up by the other colours will overlap and not be visible.

To obtain white light fringes, the mirrors are usually set at a small wedge angle and the path difference adjusted until the position of the centre of one mirror and the effective position of the other are coincident; switching from monochromatic source to white light source should then show a black vertical central fringe and a spectrum of coloured vertical fringes on either side, blue next to the black and red being on the outside. The central, vertical fringe is black because of the phase change of ≈ π on reflection at the first semi-silvered surface.

The m^{th} order fringe is when the optical path difference is $m\lambda$, *i.e.* when $2nd\cos r = m\lambda$. The maximum order fringe is the one nearest to the centre.

12.2.1 Applications

a) Change of path length. As the difference in path is decreased, the circular fringes shrink towards the centre of the pattern, one disappearing each time d changes by $\lambda/2$; hence d can be measured directly in terms of the wavelength of the monochromatic light. $(d_2 - d_1) = (m_2 - m_1)(\lambda/2)$.

Fractional parts of a fringe can be estimated so that the distance moved can be estimated to $\approx$ 0.1 - 0.05 of a wavelength.

b) Refractive index of a gas. The optical path length can be changed by decreasing the pressure of gas ina container in one of the paths. This allows a measurement of the refractive index as a function of pressure. It has the advantage of measuring (n-1) directly.

c) Length of wavetrains. If the path difference is increased, the visibility of the fringes decreases. This is because the light source emits wave trains of finite length and when the path difference approaches this length the 'sinusoidally' varying light waves no longer overlap with a constant phase difference determined by the path difference. The approximate length of the wave train can be estimated; using the Uncertainty Principle $\Delta E \times \tau \approx (h/2\pi)$, information about the mean width of the excited energy level, ΔE can be obtained.

d) Doublet wavelengths. The difference in wavelengths of a doublet can be measured accurately, *e.g.* the sodium doublet. Let the two wavelengths be λ_1 and λ_2. A path difference can be found such that when the sodium light source is used and therefore both wavelengths are present, the fringes are particularly clear and of high contrast; this is when the set of circular fringes formed by one wavelength exactly coincides near the centre of the pattern with that formed by the other. When the path difference is changed from this value, the fringes become less clear and effectively washout i.e. no fringes are visible; the fringe maxima of one pattern exactly fills in the minima of the other. As the path difference is increased further, the fringes will become clear again as the two fringe patterns are once more in step. The condition for maximum fringe visibility is that

$$2d_1 \cos r = m\lambda_1 = (m+1)\lambda_2$$

$$m = \frac{\lambda_2}{\lambda_1 - \lambda_2}$$

e.g. if one wavelength is 10% larger than the other, the fringe visibility will be a maximum when the distance is such that 10$\times$ the larger wavelength is 11$\times$ the smaller. The next coincidence will take place when

$$2d_2 \cos r = 2m\lambda_1$$

i.e.

$$\Delta d = (1/2)m\lambda_1$$

because near the centre of the pattern $\cos r \approx 1$.
Hence $\Delta d = (1/2)\frac{\lambda_1\lambda_2}{\lambda_1 - \lambda|_2}$; take $\lambda_1\lambda_2 = (589.3nm)^2$ and, by measuring the distance between maximum fringe visibility, calculate $\lambda_1 - \lambda_2$.

12.3 Twyman-Green Interferometer

The extended source is replaced by a point source at the focus of a lens; the waves reflected at the mirrors are therefore plane. The waves reflected are superimposed and brought to a focus by a second lens. If the mirrors are plane and there are no phase inhomogeneities across the cross-section of either path, the resulting intensity will be uniform across the field of view. If however one of the mirrors is not perfect, the reflected wave will not be plane and the resulting pattern will show lines of equal phase difference.

This instrument can also be used to examine how perfect a prism or a lens is - *see Longhurst,* **Fig 8-16** *or Hecht,* **section 9.10.4**.

12.4 Fourier Transform Spectroscopy

A photo-detector is used to record the intensity at the centre of the field of view of a Twyman - Green Interferometer.

Suppose that the light falling on the photo-detector has intensity $B(k)dk$ from each of the mirrors in the range k to $k+dk$, where k is the wave number, $\frac{2\pi}{\lambda}$.

If now an optical path difference, Δ, is present this will introduce a phase difference of $(\frac{2\pi}{\lambda})\Delta = k\Delta$. The resultant intensity falling on the photo-detector is

$$dI = 2B(k)dk(1+\cos(k\Delta))$$

The total intensity, considering all wave numbers, is

$$I(\Delta) = 2\int_0^\infty B(k)dk + 2\int_0^\infty B(k)\cos(k\Delta)dk$$

The first term is a constant and the second varies as the optical path difference varies. The second term is the Fourier (cosine) transform of the spectral distribution in the emergent beams - *strictly, the response of the photo-detector is folded in with this, but the above assumes that this is independent of wave number.*

If

$$\Phi(\Delta) = 2\int_0^\infty B(k)\cos(k\Delta)dk$$
$$B(k) = 2\int_0^\infty \Phi(\Delta)\cos(k\Delta)d\Delta$$

If one measures $I(\Delta)$ over a wide range of Δ, one can calculate the spectral distribution of the source ($B(k)$) by calculating the Fourier transform of the $I(\Delta)$ function - this is now done by taking a number of readings at discrete intervals and calculating the Discrete Fourier Transform. (Strictly it should be $\Phi(\Delta)$ that is used but the difference is only a constant which transforms to a delta-function at the origin.)

The light is incident normally on the mirrors and therefore the path difference is $2nd$, where d is the difference in length between the arms and n is the refractive index. If the source emits a narrow, single wavelength, then the intensity distribution shows a “cosine” periodic variation as d varies.

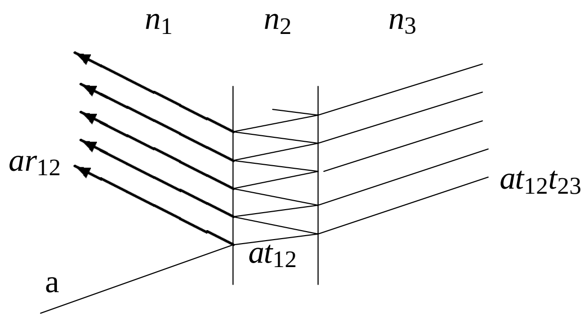
Both Michelson and Rayleigh (late 1890’s) appreciated the principle of the above but the computation required only became available recently.

The advantage of this type of instrument is that all of the light falls on the photo-detector at the same time and this results in the signal-to-noise ratio being better otherwise. This is particularly important when the light intensity is low.

12.5 The Fabry - Perot Interferometer and Etalon

Consider the transmitted rays being brought to a focus by a lens onto a screen. The waves will not all have the same phase because of the extra optical path length; the phase difference between adjacent rays will be $\delta = (2\pi/\lambda)2n_2 d\cos\theta$.

In general -



As a reminder, the *Fresnel's Reflection Coefficients are* -

$$r_{12} = \frac{n_1 - n_2}{n_1 + n_2}$$

$$r_{23} = \frac{n_2 - n_3}{n_2 + n_3}$$

$$t_{12} = 2n_1/(n_1 + n_2)$$

$$t_{23} = 2n_2/(n_2 + n_3)$$

Note that these will not apply if the surfaces are semi-silvered.

The general expressions for the transmitted and reflected sums are -

$$A_T = at_{12}t_{23}(1 + r_{23}r_{21}e^{i\delta} + (r_{23}r_{21})^2 e^{i2\delta} + (r_{23}r_{21})^3 e^{i3\delta} + \ldots\ldots$$

$$A_R = a(r_{12} + r_{23}t_{12}t_{21}e^{i\delta} + r_{23}^2 r_{21}t_{12}t_{21}e^{i2\delta} + \ldots\ldots$$

The thin film of a **Fabry-Perot etalon** has an air gap formed between two accurately parallel semi-silvered glass plates and therefore $n_1 = n_3$, $r_{23} = r_{21}$ and $t_{23} = t_{21}$.

$$A_T = at_{12}t_{21}(1 + r_{21}^2 e^{i\delta} + r_{21}^4 e^{i2\delta} + r_{21}^6 e^{i3\delta} + \ldots\ldots$$

$$r_{21}^2 e^{i\delta} A_T = at_{12}t_{21}(r_{21}^2 e^{i\delta} + r_{21}^4 e^{i2\delta} + r_{21}^6 e^{i3\delta} + \ldots\ldots$$

$$A_T = \frac{at_{12}t_{21}}{1 - r_{21}^2 e^{i\delta}}$$

Treating A_R similarly,

$$A_R = a(r_{12} + r_{21}\frac{t_{12}t_{21}e^{i\delta}}{1 - r_{21}^2 e^{i\delta}})$$

Remember, from tutorials, $t_{12}t_{21} = 1 - r_{12}^2$ and $r_{12} = -r_{21}$.

When $\delta = 2m\pi$,

$$A_R = a(r_{12} + r_{21}) = 0$$
$$A_T = a$$

This condition corresponds to $n_2 d \cos\theta = m\lambda/2$, for $m = 0, 1, 2, 3, 4, 5, \ldots$

Therefore

$$Intensity(\theta) = \frac{a^2(t_{12}t_{21})^2}{(1 - r_{21}^2 e^{i\delta})(1 - r_{21}^2 e^{-i\delta})}$$
$$Intensity(\theta) = \frac{a^2(1 - r^2)^2}{(1 + r^4 - 2r^2\cos\delta)}$$

When $\cos\delta = 1$, the transmitted intensity $I_T(\theta) = a^2$, *i.e.* full transmission because the incident amplitude is a. For this to be true, $\delta = 2m\pi$, where m is integer.

See Longhurst, **Fig 9-2** *or Hecht,* **fig 9.45**

12.5.1 Half-width of Peaks

The peak is at $\delta = 2\pi m$, m integer. The intensity falls to half of the peak intensity when $\delta = 2\pi m + \delta_{\frac{1}{2}}$, *i.e.*

$$\frac{1}{2}a^2 = \frac{a^2(1 - r^2)^2}{(1 + r^4 - 2r^2\cos\delta_{\frac{1}{2}})}$$

Assume that $\delta_{\frac{1}{2}}$ is small, *i.e.* $\cos\delta_{\frac{1}{2}} \approx (1 - \frac{1}{2}\delta_{\frac{1}{2}}^2 \ldots\ldots$

$$\delta_{\frac{1}{2}}^2 = \frac{(1 - r^2)^2}{r^2}$$

The width at half-height is $2\delta_{\frac{1}{2}}$, *i.e.*

$$\gamma = 2\frac{(1 - r^2)}{r}$$

Note that δ can be considered as a function of d, λ or θ. If we consider θ, the expression tells us that the observed circular fringes for a monochromatic light source will be very sharp for $r \approx 1$; if we consider λ, the expression can be used to predict the range of wavelengths that will be strongly transmitted, *etc.* ..

12.5.2 Resolving Power

Assume that the fringes due to two wavelengths λ and $\lambda + \Delta\lambda$ are resolved when the separate intensity curves intersect at half-height - *Taylor's Criterion*, *i.e.* at the same value of θ, This is equivalent to the angle of incidence on the plates being such that the phase difference is $2\pi m$ for one wavelength and $\geq 2\pi m + \gamma$ for the other wavelength.

See Longhurst, **Fig 9-5** *or Hecht,* **fig 9.50**

$$2\pi m = \frac{2\pi}{\lambda+\Delta\lambda} 2nd\cos\theta$$

$$2\pi m+\gamma = \frac{2\pi}{\lambda} 2nd\cos\theta$$

$$m(\lambda+\Delta\lambda) = (m+\frac{\gamma}{2\pi})\lambda$$

$$\frac{\lambda}{\Delta\lambda} = m\frac{2\pi}{\gamma} = m\frac{\pi r}{1-r^2}$$

For large r, the amplitude of the set of rays decreases slowly as the beam is reflected between the surfaces, the width of the maxima decreases and the resolving power increases. It is equivalent to a diffraction grating of larger and larger number of slits.

12.6 Anti-reflection Coatings

Demonstration Anti-reflection coating as on spectacles.

Normally, the surface to be coated is an “air-to-glass” surface. If the refractive index of the coating has a value between that of air and that of glass and such as to satisfy $r_{12} = r_{23}$, the reflection coefficient at each surface will be the same. This is satisfied if $n_2^2 = n_1 n_3$, *see tutorial*.

$$A_R = a(r_{12} + r_{12}t_{12}t_{21}e^{i\delta} + r_{12}^2 r_{21}t_{12}t_{21}e^{i2\delta} + \ldots\ldots$$

$$A_R = ar_{12} + ar_{12}t_{12}t_{21}e^{i\delta}(1-r_{12}^2 e^{i\delta} + \ldots\ldots$$

$$A_R = ar_{12} + ar_{12}t_{12}t_{21}e^{i\delta}(\frac{1}{1+r_{12}^2 e^{i\delta}})$$

If $\delta = \pi$, $e^{i\delta} = -1$

$$A_R = ar_{12} + ar_{12}t_{12}t_{21}(-1)(\frac{1}{1-r_{12}^2})$$

Remember, from tutorials, $t_{12}t_{21} = 1-r_{12}^2$ and $r_{12} = -r_{21}$.

$$A_R = 0$$

More generally, $\delta = (2m+1)\pi$, where $m = 0,1,2,3,\ldots\ldots$ satisfies the requirement.

N.B. Anti-reflection coatings are normally designed for use around the normal incidence and therefore the thickness of the coating is such that $\delta = \pi$,

$$(2\pi/\lambda)2n_2 d = \pi$$

i.e. $n_2 d = \lambda/4$

12.7 Multi-layer Reflecting Coatings

Consider first a single layer coating of refractive index, $n_2 > n_3$, *i.e.* of high refractive index. If the thickness of the layer, d, satisfies $n_2 d = \frac{\lambda}{4}$ and the light is incident at $\theta = 0$, the light reflected from the air-coating boundary will be in phase with the light reflected from the coating-glass boundary, *i.e.* the reflection coefficient will be increased.

For example, the amplitude reflection coefficient for air-glass is about 0.2 and a coating of $n_2 = 2.5$ as above increases this to approximately 0.38.

To improve the coefficient further, multiple double layers are used. Each double layer consists of a high refractive index followed by a low refractive index, each of optical thickness $nd = \frac{\lambda}{4}$. The light reflected from every boundary is in phase when it re-emerges from the surface of the multi-layer coating.