CP2: Electromagnetism, Optics and Circuit Theory

Toby Adkins

January 7, 2016

Contents

1. *Electromagnetism*

This chapter aims to cover the basics of Electromagnetic theory including:

- Maxwell's Equations
- Coulomb's Law, Dipoles and Electric Field Energy
- Conductors, Image Charges and Capacitors
- Biot-Savart Law and Magnetic Field Energy
- Motion under Magnetic and Electric Fields
- Electromagnetic Induction
- Electromagnetic Waves

This chapter is extensively mathematical in its approach, unlike high-school electromagnetism, and so the reader should be familiar with vector calculus (including the Divergence and Stokes' Theorems), basic integral calculus and small expansion techniques (e.g. the Taylor Series Expansion). Note that throughout this document, we will be using ϕ for the scalar potential to differentiate it from the volume element ∂V or the voltage V.

1.1 Maxwell's Equations

The following sections list Maxwell's Equations. Note that they are shown both in integral and differential form; each type is useful for different applications.

1.1.1 Gauss' Electric Law

Gauss' Electric law states that:

$$\oint_{\partial V} \underline{E} \cdot d\underline{S} = \frac{1}{\varepsilon_0} \int_V \rho \cdot dV \tag{1.1}$$

$$\nabla \cdot \underline{E} = \frac{\rho}{\varepsilon_0} \tag{1.2}$$

where ∂V is the boundary of V. We can use spherical/cylindrical symmetry to evaluate the electric field.

Gauss' law comes from considering the electric flux in a region. Consider a collection of charges q_i, and (1.12):

$$\begin{aligned} \underline{E} &= \frac{\underline{F}}{q} = \frac{q_i}{4\pi\varepsilon_0} \cdot \frac{1}{r_i^2} \cdot \underline{\hat{r}} \\ d\phi_i &= \underline{E_i} \cdot d\underline{S} \\ &= \frac{q_i}{4\pi\varepsilon_0} \cdot \frac{1}{r_i^2} \cdot \underline{\hat{r}} \cdot d\underline{S} \\ &= \frac{q_i}{4\pi\varepsilon_0} \cdot d\Omega \end{aligned}$$

Subsequently:

$$\oint_{\partial V} \underline{E} \cdot d\underline{S} = \frac{q}{\varepsilon_0} \int \frac{d\Omega}{4\pi}$$

And hence the result follows.

Physically, this describes how charges generates an electric field, and that electric field lines begin and end on charge; electric monopoles exist. We can also consider Gauss's law in differential form to derive Laplace's Equation:

$$\begin{aligned} \nabla \cdot \underline{E} &= -\nabla \cdot (\nabla\phi) \\ \nabla^2 \phi &= -\frac{\rho}{\varepsilon_o} \end{aligned}$$

This is know as Poisson's Equation. In regions with zero charge density, this reduces to Laplace's Equation:

$$\nabla^2 \phi = 0 \tag{1.3}$$

1.1.2 Gauss' Magnetic Law

Gauss' Magnetic Law states that:

$$\oint_{\partial V} \underline{B} \cdot d\underline{S} = 0 \tag{1.4}$$

$$\nabla \cdot \underline{B} = 0 \tag{1.5}$$

This states that magnetic fields are not divergent; magnetic field lines form closed lines around current sources. There are no magnetic monopoles.

1.1.3 Faraday's Law

Faraday's Law states that:

$$\oint_{\partial S} \underline{E} \cdot d\underline{l} = -\frac{\partial}{\partial t} \oint_S \underline{B} \cdot d\underline{S} \tag{1.6}$$

$$\nabla \times \underline{E} = -\frac{\partial \underline{B}}{\partial t} \tag{1.7}$$

This is the electric analogue of Ampere's law; that time varying magnetic fields create electric fields. Note that this is not the full form of the law. See Section 1.6 for details.

1.1.4 Ampere's Law

Ampere's law in it's original form states that:

$$\oint_{\partial S} \underline{B} \cdot d\underline{l} = \mu_o \oint_S \underline{J} \cdot d\underline{S} \tag{1.8}$$

$$\nabla \times B = \mu_o \cdot \underline{J} \tag{1.9}$$

However, Ampere's law in this form is inconsistent with the conservation of charge. This can be seen by considering $\nabla \cdot (\nabla \times B)$.

$$\nabla \cdot (\nabla \times B) = \mu_o \nabla \cdot \underline{J}$$
$$0 = \mu_o \nabla \cdot \underline{J}$$

Conservation of charge tells us the following:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \cdot \underline{v}) = 0$$
$$\therefore \nabla \cdot \underline{J} = -\frac{\partial \rho}{\partial t}$$

From Gauss' Electric Law (1.2):

$$\nabla \cdot \underline{E} = \frac{\rho}{\varepsilon_o}$$
$$\rho = \varepsilon_o \nabla \cdot \underline{E}$$
$$\nabla \cdot \underline{J} = -\frac{\partial}{\partial t}(\varepsilon_o \nabla \cdot \underline{E})$$

Thus, to make the law consistent with the conservation of charge, we require that:

$$\nabla \cdot (\nabla \times B) = \mu_o \nabla \cdot \underline{J} + \frac{\partial \rho}{\partial t}\mu_0$$
$$= \mu_o \nabla \cdot \underline{J} + \nabla \cdot \mu_o \varepsilon_o \frac{\partial \underline{E}}{\partial t}$$
$$\rightarrow \nabla \times \underline{B} = \mu_o \underline{J} + \mu_o \varepsilon_o \frac{\partial \underline{E}}{\partial t}$$

This can also be shown by considering the charge on a parallel plate capacitor. Consider Figure 1.1:

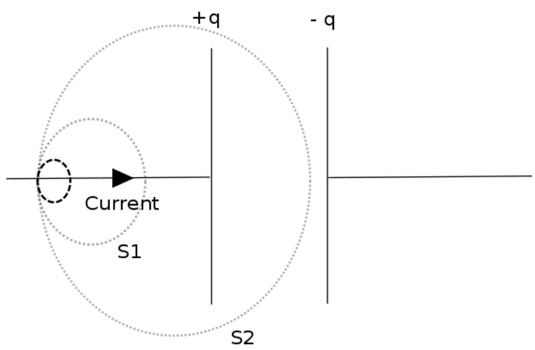

Figure 1.1: Parallel Plate Capacitor with Gaussian Surfaces

Using Stokes' Theorem on (1.9):

$$\oint_S (\nabla \times \underline{B}) \cdot d\underline{S} = \mu_o \oint_S \underline{J} \cdot d\underline{S}$$

Now consider the current that runs through the two surfaces S_1 and S_2:

$$S_1 : I_{enclosed} = I$$
$$S_2 : I_{enclosed} = 0$$

This means that the result for the magnetic field can change as the Gaussian surface that we choose is completely arbitrary. This means that it is inconsistent with charge conservation. From Section (2.1.2):

$$E = \frac{q}{\varepsilon_o A}$$
$$I = \frac{\partial q}{\partial t} = \varepsilon_o A \frac{\partial E}{\partial t}$$

And hence the previous result follows. The final form of Ampere's Law (known as the "Ampere-Maxwell Equation") is as follows:

$$\oint_{\partial S} \underline{B} \cdot d\underline{l} = \mu_o I_{enclosed} + \mu_o \varepsilon_o \frac{\partial}{\partial t} \oint_S \underline{E} \cdot d\underline{S} \quad (1.10)$$

$$\nabla \times \underline{B} = \mu_0 \varepsilon_o \frac{\partial E}{\partial t} + \mu_o \underline{J} \quad (1.11)$$

This can be interpreted as the mathematical statement of the fact that electric currents and time varying electric fields generate magnetic fields.

1.2 Coulomb's Law, Dipoles and Electric Field Energy

This section looks at the basics of electrostatics; that is, electromagnetic theory applied to static charges. Note that we can apply the Principle of Superposition in this case; that the total potential/electric field is the sum of all of the present potentials/electric fields.

1.2.1 Coulomb's Law

Coulomb's Law states that the force between two charges q_i and q_j is given by:

$$\underline{F} = \frac{1}{4\pi\varepsilon_o} \cdot \frac{q_i q_j}{r_i^2} \cdot \underline{\hat{r}} \tag{1.12}$$

This leads to the definition of the electric field as:

$$\underline{E} = \lim_{q \to 0} \frac{\underline{F}}{q} \tag{1.13}$$

Every electric field can be written as the gradient of some scalar potential, provided that the field is conservative.

$$\phi = \frac{W}{q}$$
$$\nabla\phi = \lim_{q \to 0} \frac{\nabla W}{q} = -\frac{\underline{F}}{q}$$

where W is the work done. This leads to one of the most important electrostatics results:

$$\underline{E} = -\nabla\phi \tag{1.14}$$

This means that no electric field can exist without a potential $\longleftrightarrow$ in regions of zero potential, there is no electric field, assuming that the zero of potential is sufficiently well defined.

1.2.2 The Dipole

This is one of the more common objects in electrostatics. It consists of two oppositely poled charges separated by a small separation. The product of the vector separation and the magnitude of one of the charges is known as the "electric dipole moment", written as $\underline{p}$. Conventionally, the dipole moment is orientated from the negative charge to the positive charge. Consider the electric field at the point P (where $a \ll r$) from the two point charges, as shown in Figure 1.2:

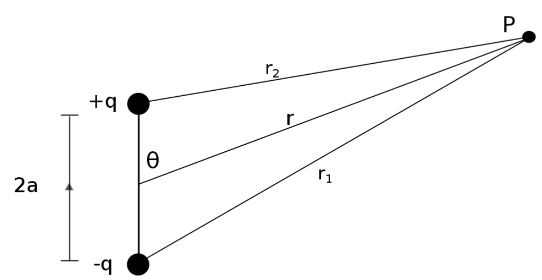

Figure 1.2: Electric Dipole with $\underline{p} = 2\underline{a}q$

By the principle of superposition:

$$\begin{aligned} {r_1}^2 &= r^2 + a^2 - 2ar\cos\theta \\ {r_2}^2 &= r^2 + a^2 + 2ar\cos\theta \end{aligned}$$

$$\begin{aligned} \phi &= \frac{q}{4\pi\varepsilon_o} \cdot \left(\frac{1}{r_1} - \frac{1}{r_2}\right) \\ &= \frac{q}{4\pi\varepsilon_o} \cdot \frac{1}{r} \cdot \left(\frac{1}{\sqrt{1 + \left(\frac{a}{r}\right)^2 - \frac{2a}{r}\cos\theta}} - \frac{1}{\sqrt{1 + \left(\frac{a}{r}\right)^2 + \frac{2a}{r}\cos\theta}}\right) \end{aligned}$$

Taylor expanding using the fact that $a \ll r$:

$$\begin{aligned} \phi &= \frac{q}{4\pi\varepsilon_o} \cdot \frac{1}{r} \cdot \left(1 - \frac{1}{2}\left(\frac{a}{r}\right)^2 + \frac{a}{r}\cos\theta - \left(1 - \frac{1}{2}\left(\frac{a}{r}\right)^2 - \frac{a}{r}\cos\theta\right)\right) \\ &= \frac{q}{4\pi\varepsilon_o} \cdot \left(\frac{2a}{r^2}\cos\theta\right) \end{aligned}$$

And hence we obtain the dipole potential as:

$$\phi_{dipole} = \frac{\underline{p} \cdot \widehat{\underline{r}}}{4\pi\varepsilon_o} \cdot \frac{1}{r^2} \tag{1.15}$$

We can then apply (1.14) in a chose coordinate system (usually spherical) to find the electric field due to the dipole.

Now we consider the force and torque on a dipole in a non-uniform electric field. Suppose that electric field at the positive charge is $\underline{E} + d\underline{E}$ and at the negative charge $\underline{E}$. Suppose that the distance between the two charges is $d\underline{S}$.

$$\begin{aligned} \underline{F} &= \underline{F_1} + \underline{F_2} \\ &= q(\underline{E} + d\underline{E}) - q \cdot \underline{E} \\ &= q \cdot d\underline{E} \end{aligned}$$

$$\begin{aligned} qd\underline{E} &= q(d\underline{S} \cdot \nabla)\underline{E} \\ \underline{p} &= q \cdot d\underline{S} \\ \rightarrow \underline{F}_{dipole} &= \nabla(\underline{p} \cdot \underline{E}) \end{aligned} \tag{1.16}$$

Now for the torque:

$$\begin{aligned} \underline{\tau} &= \underline{r_1} \times \underline{F_1} + \underline{r_2} \times \underline{F_2} \\ &= q(\underline{r_1} \times (\underline{E} + d\underline{E}) - \underline{r_2} \times \underline{E}) \\ &= q((\underline{r_1} - \underline{r_2}) \times \underline{E} + \underline{r_1} \times d\underline{E}) \\ &= q(2\underline{a} \times \underline{E} + \underline{r} \times d\underline{E}) \end{aligned}$$

$$\rightarrow \underline{\tau}_{dipole} = \underline{p} \times \underline{E} + \underline{r} \times \underline{F} \tag{1.17}$$

1.2.3 Electric Field Energy

To bring a charge in from the zero of potential (conventionally located at ∞) to a particular location in the presence of other charges requires work against the already established electric fields. This contributes to the final energy of an assembly of charges.

$$
\begin{aligned}
W &= \int_{r_1}^{r_2} \underline{F} \cdot d\underline{r} \\
&= \frac{q_i q_j}{4\pi\varepsilon_o}\left(\frac{1}{r_2} - \frac{1}{r_1}\right) \\
&= \frac{q_i q_j}{4\pi\varepsilon_o} \cdot \frac{1}{r_2}
\end{aligned}
$$

for $r_2 \to \infty$. Then, consider the potential due to the charge q_i:

$$
\begin{aligned}
\phi_i &= \frac{W_i}{q_j} = \frac{q_i}{4\pi\varepsilon_o} \cdot \frac{1}{r_i} \\
W &= \frac{1}{2} \cdot \frac{1}{4\pi\varepsilon_o} \sum_{i,j}^{\infty} \frac{q_i q_j}{r_{ij}} \\
\to W &= \frac{1}{2} \sum_{i,j}^{\infty} q_i \phi_i \qquad (1.18)
\end{aligned}
$$

However, in this case, we have assumed that the charge distribution is discrete; that it consists of individual charges. If the charge distribution is continuous:

$$
\begin{aligned}
W &= \frac{1}{2} \int_V \rho\phi \, dV \\
&= \frac{1}{2} \int_V \varepsilon_o (\nabla \cdot \underline{E})\phi \, dV \\
&= \frac{1}{2}\varepsilon_o \int_V (\nabla \cdot \underline{E})\phi \, dV
\end{aligned}
$$

Applying the divergence theorem to the left-hand side of the following:

$$
\begin{aligned}
\oint_{\partial V} \phi \underline{E} \cdot d\underline{S} &= \int_V \nabla \cdot (\phi \underline{E}) \, dV \\
&= \int_V \phi(\nabla \cdot \underline{E}) \, dV + \int_V \underline{E} \cdot \nabla\phi \, dV \\
\to W &= \frac{1}{2}\varepsilon_o \int_V \underline{E}^2 \, dV \qquad (1.19)
\end{aligned}
$$

Let us consider an example of a uniformly charged sphere of radius R and total charge Q. We need to take account of both the energy required to bring the charges in (from ∞) to the surface of the sphere, as well as the energy required to distribute the charge throughout the sphere.

First, use Gauss' Electric Law (1.2) to find the electric fields inside and outside the sphere:

for $r \geq R$

$$\begin{aligned} q_{enclosed} &= Q \\ \rightarrow E &= \frac{q}{4\pi\varepsilon_o} \cdot \frac{1}{r^2} \\ U_{outside} &= \frac{1}{2}\varepsilon_o \int_R^\infty \frac{Q^2}{16\pi^2{\varepsilon_o}^2} \cdot \frac{1}{r^4} \cdot 4\pi r^2 dr \\ &= \frac{Q^2}{32\pi\varepsilon_o}\left(\frac{4}{R}\right) \end{aligned}$$

for $r < R$

$$\begin{aligned} q_{enclosed} &= Q\frac{r^3}{R^3} \\ \rightarrow E &= \frac{Q}{4\pi\varepsilon_o} \cdot \frac{r^3}{R^3} \\ U_{outside} &= \frac{1}{2}\varepsilon_o \int_0^R \frac{Q^2}{16\pi^2{\varepsilon_o}^2} \cdot \frac{r^6}{R^6} \cdot 4\pi r^2 dr \\ &= \frac{Q^2}{32\pi\varepsilon_o}\left(\frac{4}{5R}\right) \end{aligned}$$

Hence, the total electrostatic potential energy of the sphere is given by:

$$U_{Total} = \frac{3Q^2}{20\pi\varepsilon_o} \cdot \frac{1}{R}$$

1.3 Conductors, Image Charges and Capacitors

This section looks at the way in which electric fields established due to static charges can have an influence on conductors and capacitors.

1.3.1 Properties of Conductors

The following properties of conductors follow as a consequence from the results derived in the previous two sections:

- $\underline{E} = 0$ inside a conductor. Electrons are free to move inside a conductor, and so will move to create an internal electric field that will oppose the external field.
- $\rho = 0$ inside a conductor. From 1.2, $\underline{E} = 0 \to \rho = 0$, as required.
- This means that any net charge must reside on the surface of a conductor.
- A conductor is an equipotential.
- The field lines of $\underline{E}$ are always perpendicular to the surface of a conductor.

1.3.2 The Image Charge Method

This is a technique used to simplify electrostatics problems. It involved replacing certain elements of the problem with static charges that reproduce the boundary conditions of the problem. The solution to this analogous problem will be the unique solution to the problem according to the Uniqueness Theorem, which states that *for a well defined reqion, there is only one possible solution for the potential* ϕ. The proof is as follows.

Suppose that the potentials ϕ_1 and ϕ_2 are solutions to Laplace's equation (1.3). This means that at the boundary $\phi_1 = \phi_2$ for them to both satisfy the boundary conditions. Now consider $\phi_3 = \phi_2 - \phi_2$.

$$\nabla^2 \phi_3 = 0$$
$$\phi_3 = 0 \, \textit{at the boundary of } \partial V$$

This means that $\phi_3 = 0$ over ∂V. and this thus implies that $\phi_1 = \phi_2$ over ∂V, and the theorem follows.

Now let us consider a grounded (hollow) conducting sphere located in the field of a point charge q_2 (as shown in Figure 1.3), and ask ourselves the question *" What is the force felt by the charge? "*.

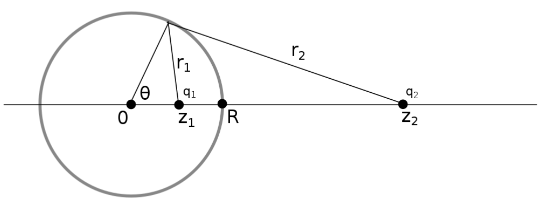

Figure 1.3: Spherical Shell in the presence of q_2

We know that the sphere must remain in equipotential, so the image charge that we have to place must also satisfy this condition.

$$
\begin{aligned}
{r_1}^2 &= R^2 + {z_1}^2 - 2Rz_1\cos\theta \\
{r_2}^2 &= R^2 + {z_2}^2 - 2Rz_2\cos\theta \\
q_1{r_2}^2 + q_2{r_1}^2 &= 0 \\
\rightarrow \frac{q_1}{{r_1}^2} + \frac{q_2}{{r_2}^2} &= 0
\end{aligned}
$$

And hence the boundary condition is satisfied by a point charge located at z_1 inside the sphere. We can thus find the force on the charge q_2 by Coulomb's Law (1.2.1).

$$
\begin{aligned}
|\underline{F}| &= \frac{1}{4\pi\varepsilon_o}\cdot\frac{|q_1q_2|}{r^2} \\
r &= z_1 - z_2 \\
&= z_1 + R\frac{q_2}{q_1}
\end{aligned}
$$

Hence:

$$
F = \frac{1}{4\pi\varepsilon_o}\cdot\frac{|q_1q_2|}{\left(z_1 + R\frac{q_2}{q_1}\right)^2}
$$

Note that in simpler image charge problems involving conducting plates and the like, the surface charge density can be found using:

$$
\sigma = -\varepsilon_o\frac{\partial\phi}{\partial z}\bigg|_{z=0} \tag{1.20}
$$

This comes from considering (1.2) and using (1.14), assuming that close to the surface the electric field is perpendicular to it.

1.3.3 Capacitors

Capacitance is the storage of energy through the separation of two oppositely poled charge accumulations. Capacitance is defined as:

$$
Capacitance = \frac{Charge\ Stored}{Voltage\ Applied} \tag{1.21}
$$

$$
C = \frac{Q}{V} \tag{1.22}
$$

Capacitance is by convention positive, and we add capacitors in parallel, and add their inverses in series. We can use Gauss' Law (1.2) to find the electric field between the two capacitor surfaces, and then use $\phi = \int \underline{E}\cdot d\underline{r}$ to find the voltage. Below are the capacitances of some common objects, derived by applying Gauss' law in different coordinate systems:

- Parallel Plate Capacitor:

$$
C = \varepsilon_o\frac{A}{d}
$$

- Cylindrical Capacitor:

$$
\frac{C}{l} = \frac{2\pi\varepsilon_o}{\log\left(\frac{b}{a}\right)}
$$

- Spherical Capacitor:

$$C = 4\pi\varepsilon_o \left(\frac{ab}{b-a}\right)$$

- Twin Cables:

$$\frac{C}{l} = \pi\varepsilon_o \log\left(\frac{a}{d}\right)$$

Note that for an isolated spherical shell, it's capacitance can be found by letting the radius of the outer shell (in this case b) tend to ∞.

The energy stored by a capacitor is given by:

$$dW = VdQ = V\,d(CV) = VCdV$$
$$W = \int_0^V VCdV$$
$$W = \frac{1}{2}CV^2 = \frac{1}{2}\frac{1}{C}Q^2 \tag{1.23}$$

when capacitor plates are moved apart, they experience a force between them that resists the motion; i.e work has to be done to pull them apart in order to conserve energy. Consider the work done:

$$dW_{mechanical} = dW_{capacitor} - dW_{battery}$$
$$F = -\frac{dW_m}{dx}$$
$$\rightarrow F = \frac{d}{dx}(W_B) - \frac{d}{dx}(W_C) \tag{1.24}$$

This gives rise to two cases:

- If the battery is not connected, then the charge on the capacitor plates must stay constant in order to conserve charge. Thus, $\frac{d}{dx}(W_B) = 0$, and we have to use the second of the two expressions in (1.23).
- If the battery is connected, the charge is able to change, but the battery will maintain the system at a constant voltage. Thus, $\frac{d}{dx}(W_B) \neq 0$, and we have to consider the work done by the battery:

$$P = VI = V\frac{dQ}{dt}$$
$$W = VQ$$
$$dW = VdQ$$
$$= Vd(CV)$$
$$= V^2dC$$

Curiously, if one were to carry out the calculation in both cases, it would become clear that they actually obtain the same result, except differing by the sign. However, do not be fooled; the force in both cases is attractive, the sign is due to the way that we consider the battery to do work on the system.

1.4 Biot-Savart Law and Magnetic Field Energy

The Biot-Savart Law allows us to find the value of the magnetic field at some location relative to a steady current, and unlike Ampere's Law (1.1.4), it holds true for all geometries, not just in the 'infinite' cases where we can ignore edge effects. It is as follows:

$$d\underline{B} = \frac{\mu_o I}{4\pi} \cdot \frac{d\underline{l} \times \underline{r}}{r^3} \tag{1.25}$$

where r is the position vector from the line element $d\underline{l}$. Assuming that appropriate expressions can be found for $d\underline{l} \times \underline{r}$ and r, the magnetic field can easily be calculated. Some common derivations follow:

1.4.1 Straight Current Carrying Wire

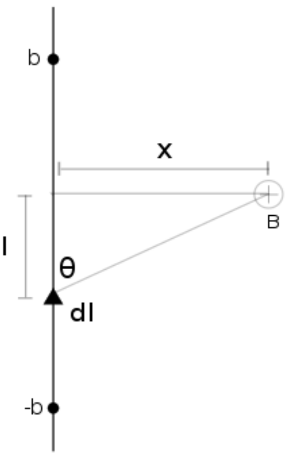

Figure 1.4: Straight Current Carrying Wire

$$\begin{aligned} |d\underline{l} \times \underline{r}| &= r \cdot dl \cdot \sin\theta \\ &= r \cdot dl \cdot \frac{x}{r} \\ &= x \cdot dl \\ B &= \frac{\mu_o I}{4\pi} \int_{-b}^{b} \frac{x \cdot dl}{(x^2 + l^2)^{\frac{3}{2}}} \\ &= \frac{\mu_o I}{4\pi x} \left[\frac{l}{(x^2 + l^2)^{\frac{1}{2}}} \right]_{-b}^{b} \end{aligned}$$

$$\to B = \frac{\mu_o I}{2\pi x} \cdot \frac{l}{(x^2 + l^2)^{\frac{1}{2}}} \tag{1.26}$$

Consider the value of the magnetic field in the limiting case where $|b| \to \infty$; we re-obtain the result that we would get if we found the magnetic field using Ampere's Law (1.9) (ignoring edge effects).

$$\lim_{|b|\to\infty} B = \frac{\mu_o I}{2\pi x}$$

1.4.2 Current Loop

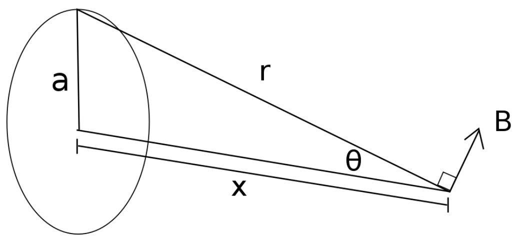

Figure 1.5: Current Loop

$$
\begin{aligned}
|\underline{dl} \times \underline{r}| &= r \cdot dl \cdot \sin\theta \\
&= r \cdot dl \cdot \frac{a}{r} \\
&= a \cdot dl
\end{aligned}
$$

Note that both a and r are independent of l, and so we can extract them from the integral.

$$
\begin{aligned}
B &= \frac{\mu_0 I}{4\pi} \int \frac{a \cdot dl}{r^3} \\
&= \frac{\mu_0 I}{4\pi} \cdot \frac{a}{r^3} \int_0^{2\pi a} dl
\end{aligned}
$$

$$
\rightarrow B = \frac{\mu_o I}{2} \cdot \frac{a^2}{(x^2 + a^2)^{\frac{1}{2}}} \tag{1.27}
$$

1.4.3 Straight Solenoid

We will first treat this scenario using the Biot-Savart Law (1.25), and then examine the derivation using Ampere's law (1.9).

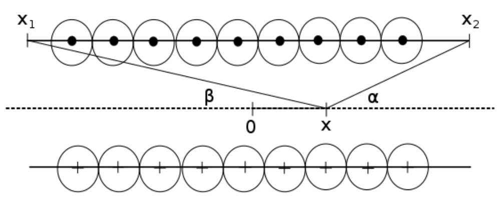

Figure 1.6: Straight Solenoid with N turns

$$
\begin{aligned}
dI &= \frac{NI}{L} \\
dB &= \frac{\mu_0 NI}{2L} \cdot \frac{a^2}{((x-x')^{\frac{1}{2}} + a^2)^{\frac{3}{2}}} \\
B &= \frac{\mu_0 NI}{2L} \left[\frac{x_2 - x}{\sqrt{((x-x_2)^{\frac{1}{2}} + a^2}} - \frac{x_1 - x}{\sqrt{((x-x_1)^{\frac{1}{2}} + a^2}} \right] \\
B &= \frac{\mu_o NI}{2L} \left[\cos\alpha + \cos\beta \right]
\end{aligned}
$$

Letting $\alpha = 0$ and $\beta = 0$, we obtain the 'infinite' result of:

$$\rightarrow B = \frac{\mu_o NI}{L} \tag{1.28}$$

Now, let us look at how we can obtain this result using Ampere's Law (1.9). Consider the Amperian loop *abcd* shown in Figure 1.7.

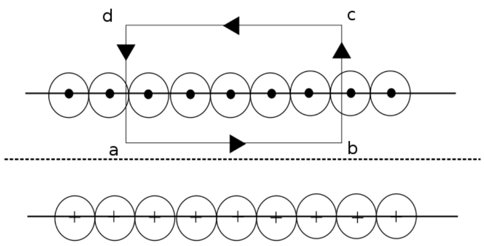

Figure 1.7: Infinite Straight Solenoid with N turns

Using (1.5) with a cylindrical Gaussian surface inside the solenoid, it becomes clear that:

$$
\begin{aligned}
B_r &= 0 \\
B_\theta &= 0 \\
B_x &\neq 0
\end{aligned}
$$

These are also quite obvious by the symmetry of the problem.

Applying Ampere's Law (1.9):

$$\oint \underline{B} \cdot d\underline{l} = \int_{a\rightarrow b} \underline{B} \cdot d\underline{l} + \int_{b\rightarrow c} \underline{B} \cdot d\underline{l} + \int_{c\rightarrow d} \underline{B} \cdot d\underline{l} + \int_{d\rightarrow a} \underline{B} \cdot d\underline{l}$$

The paths $b \rightarrow c$ and $a \rightarrow d$ cancel one another as they are opposite and equal, and the path *cd* makes no contribution as the field is zero outside the solenoid (or rather, we can arbitrarily place the side *cd* at infinity). Thus, we just have to consider $a \rightarrow b$.

$$
\begin{aligned}
B(d-a) &= \mu_o NI \\
B &= \frac{\mu_o NI}{L}
\end{aligned}
$$

As $d - a = L$. At no point have we said where the side *ab* lies inside the solenoid. This means that the field inside the solenoid is uniform as it is independent of the lengths of *ad* and *bc*.

1.4.4 Magnetic Field Energy

The results obtained here are analogous to those for the electric field, and can be derived using similar methods.

- Within a circuit:

$$\begin{aligned} U &= \int VI \cdot dt \\ &= \int L\frac{dI}{dt} \cdot dt \\ &= \int LI \cdot dI \\ \rightarrow U &= \frac{1}{2}LI^2 \end{aligned}$$

- For continuous distributions:

$$\rightarrow U = \frac{1}{2\mu_o}\int \underline{B}^2 dV$$

1.5 Motion under Magnetic and Electric Fields

For a charged particle moving through space, it will be influenced by the presence of electric and magnetic fields. This motion is governed by what is known as the Lorentz Force Equation:

$$\underline{F} = q\left(\underline{E} + \underline{v} \times \underline{B}\right) \tag{1.29}$$

This can be simplified to a set of separable equations of the form:

$$\begin{pmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{z} \end{pmatrix} = \frac{q}{m}\begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} + \frac{q}{m}\begin{pmatrix} \dot{y}B_z - \dot{z}B_y \\ \dot{z}B_z - \dot{x}B_z \\ \dot{x}B_y - \dot{y}B_x \end{pmatrix} \tag{1.30}$$

This looks quite complicated (and it is if you try to solve it in this form). However, usually the $\underline{E}$ and $\underline{B}$ fields are usually mutually orthogonal and orthogonal to the velocity, which makes the calculation much easier. Note that when solving these types of problems, we usually neglect gravity as it's magnitude is vastly smaller that the magnitudes of the forces due to the $\underline{E}$ and $\underline{B}$ fields

Trivially, it becomes clear that if we require a particle to be un-deflected, then the velocity, $\underline{E}$ field and $\underline{B}$ field have to form a right-handed basis set, and the velocity must have magnitude:

$$\begin{aligned} 0 &= q\left(\underline{E} + \underline{v} \times \underline{B}\right) \\ \underline{E} &= -\underline{v} \times \underline{B} \\ \rightarrow |\underline{v}| &= \frac{|\underline{E}|}{|\underline{B}|} \end{aligned}$$

1.5.1 Motion in an isolated Magnetic Field

Let us derive the force a charge experiences due to magnetic field. Consider the current due a moving charge or a set of charges.

$$\begin{aligned} I &= \frac{dq}{dt} \\ &= \frac{dq}{dl} \cdot \frac{dl}{dt} \\ &= \frac{dq}{dl} \cdot v \end{aligned}$$

$$\begin{aligned} d\underline{F} &= I\left(d\underline{l} \times \underline{B}\right) \\ &= dq\left(\underline{v} \times \underline{B}\right) \end{aligned}$$

Hence, we obtain the previous result of:

$$\underline{F} = q\left(\underline{v} \times \underline{B}\right) \tag{1.31}$$

If a particle initially moving in a plane perpendicular to the magnetic field enters the field, then it will perform circular motion in this plane for as long as it remains inside the field. Suppose that $\underline{v} = \left(\dot{x}, \dot{y}, 0\right)$ and $\underline{B} = \left(0, 0, B\right)$. Then:

$$\begin{pmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{z} \end{pmatrix} = \frac{qB}{m}\begin{pmatrix} \dot{y} \\ -\dot{x} \\ 0 \end{pmatrix}$$

This system has general solution:

$$x = R\cos(\Omega t)$$
$$y = R\sin(\Omega t)$$

for $\Omega = \frac{qB}{m}$ (this is known as the "Cyclotron Frequency"). It thus performs circular motion with radius:

$$\ddot{r} = \Omega v$$
$$\frac{v^2}{R} = \Omega v$$
$$\rightarrow R = \frac{mv}{Bq} \tag{1.32}$$

In general, the motion in an isolated B field can be described by the following equation:

$$\underline{r} = \underline{p} + (\underline{\widehat{B}} \cdot \underline{v}_o)\underline{\widehat{B}}\, t \; + \; \frac{1}{\Omega}\, \underline{v}_{perp} \times \underline{\widehat{B}} \tag{1.33}$$

where $\underline{v}_{perp} = \underline{v}_o \times \underline{\widehat{B}} \cos(\omega t) + \underline{v}_o \sin(\omega t)$. This comes from the fact that it will move along the $\underline{B}$ field with the component of the initial velocity ($\underline{v}_o$) that is in that direction.

1.6 Electromagnetic Induction

Electromagnetic induction is the action of creating a voltage/potential difference (denoted by ε) in a conductor due a changing magnetic flux (denoted by Φ).

$$\Phi = \oint_S \underline{B} \cdot d\underline{S} \tag{1.34}$$

There are two mains laws that govern electromagnetic induction; Faraday's Law and Lens' Law. Note that the form of Faraday's Law quoted here is not the same as that outlined in Section (1.1.3), as we have added an extra term that we only have to take account of in certain specific scenarios (see Section (1.6.1)).

Faraday's Law states that:

$$\varepsilon = -\frac{\partial}{\partial t}\Phi + \oint_{\partial S} (\underline{v} \times \underline{B}) \cdot d\underline{l} \tag{1.35}$$

Lens' Law states that *when a current flows in a conductor as a result of EM induction, it will flow such that the magnetic field it creates acts to oppose the external change in magnetic flux*; essentially, it is the minus sign out the front of (1.35). This can be thought of as a consequence of the conservation of energy, as otherwise energy would be continually created in the system as a result of the induced current.

Below are a couple of examples of electromagnetic induction problems. In general, the only difficult part is finding an appropriate expression for Φ, and the rest will follow relatively trivially.

1.6.1 Faraday Disk

The Faraday Disk is simply a conducting disk spinning in a magnetic field that is parallel with it's principle axis. Note that in this case $\Phi = 0$ as the surface area in the magnetic field does not change. This is where the second term in (1.35) comes into play.

$$\begin{aligned}
\varepsilon &= \oint_{\partial S} (\underline{v} \times \underline{B}) \cdot d\underline{l} \\
&= \oint_{\partial S} ((\underline{\omega} \times \underline{r}) \times \underline{B}) \cdot d\underline{l} \\
&= \oint_{\partial S} ((\underline{B} \cdot \underline{r})\omega - (\underline{B} \cdot \underline{\omega})) \cdot d\underline{l}
\end{aligned}$$

and because $\underline{B} \cdot \underline{r} = 0$ we have:

$$\begin{aligned}
\varepsilon &= -\oint_{\partial S} B\omega r \cdot dr \\
\rightarrow \varepsilon &= \frac{B\omega r^2}{2}
\end{aligned}$$

Assuming that the disk is connected to an external resistance, it will slow down when it is no longer driven at ω. The average power will be equal to the rate of change of kinetic energy. Assume that the disk has moment of inertia I.

$$P_{average} = \frac{\partial}{\partial t}\left(\frac{1}{2}I\omega^2\right) \tag{1.36}$$

Consider the average power dissipated in the resistance R.

$$\begin{aligned} P &= \frac{V^2}{R} \\ &= \frac{B^2\omega^2R^4}{4R} \\ \frac{B^2\omega^2R^4}{4R} &= \frac{1}{2}\cdot I\cdot\frac{\partial}{\partial t}(\omega^2) \end{aligned}$$

This is a separable differential-equation that can be solved to find ω at some time t given some initial conditions.

1.6.2 The Induction Pendulum

Imagine that we have a solid conducting bar attached to an horizontal support by two light, insulated strings. It swings in a magnetic field. *What is the induced voltage between the ends of the bar?* The trick with this question is to recognise that the magnetic flux that passes through the bar per half cycle is the dotted region in Figure (1.8).

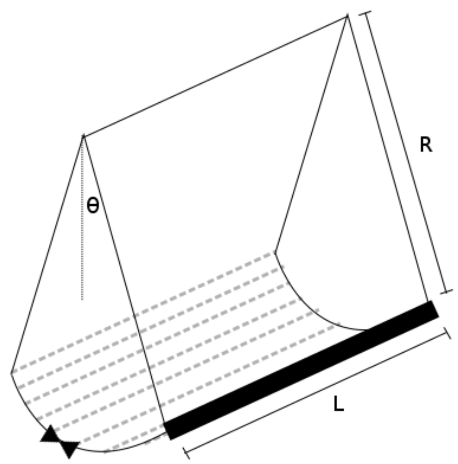

Figure 1.8: The Induction Pendulum

- For a $\underline{B}$ that is oriented along $\theta = 0$:

$$\begin{aligned} \Phi &= \oint_S \underline{B}\cdot d\underline{S} \\ &= \int B\cos\theta\cdot RL\cdot d\theta \\ \frac{\partial\Phi}{\partial t} &= -BRL\cos\theta\cdot\frac{\partial\theta}{\partial t} \\ &= -\omega\cdot\theta_o\cdot BRL\cos\theta\cdot\sin(\omega t) \end{aligned}$$

  Hence, for small angles in θ:

$$\implies V = \omega\cdot\theta_o\cdot BRL\sin(\omega t)$$

- For a $\underline{B}$ that is oriented along $\theta = 90°$:

$$\begin{aligned}
\Phi &= \oint_S \underline{B} \cdot d\underline{S} \\
&= -\int B\cos(\frac{\pi}{2} - \theta) \cdot RL \cdot d\theta \\
\frac{\partial \Phi}{\partial t} &= -BRL\sin\theta \cdot \frac{\partial \theta}{\partial t}
\end{aligned}$$

Hence, for small angles in θ:

$$\begin{aligned}
V &= \omega \cdot {\theta_o}^2 \cdot BRL\cos(\omega t)\sin(\omega t) \\
\rightarrow V &= \frac{\omega \cdot {\theta_o}^2 \cdot BRL}{2}\sin(2\omega t)
\end{aligned}$$

1.6.3 Self Inductance

Self inductance is defined as the induction of a voltage in a current-carrying wire when the current in the wire itself is changing. It is defined as:

$$L = \frac{d\Phi}{dt} \cdot \frac{dt}{dI} = \frac{d\Phi}{dI} \equiv \frac{\varepsilon}{I} \tag{1.37}$$

The chain rule expression has been included here as sometimes it becomes useful when trying to work from a mutual inductance result to a magnetic flux result, for example. Below are some common self inductances:

- Solenoid:

$$\begin{aligned}
B &= \frac{\mu_o NI}{l} \\
L &= \frac{\partial}{\partial I}\left(\frac{\mu_0 NI}{l} \cdot (N\pi a^2)\right) \\
\rightarrow L &= \frac{\mu_o N^2 \pi a^2}{l}
\end{aligned}$$

- Square Toroid

$$\begin{aligned}
B &= \frac{\mu_o NI}{2\pi r} \\
\Phi &= N\int \underline{B} \cdot d\underline{S} \\
&= N\int_{R+a}^{R-a} \frac{\mu_o NI}{2\pi r} \cdot dr \\
&= \frac{\mu_o N^2 I}{2\pi} ln\left(\frac{R+a}{R-a}\right) \\
\rightarrow L &= \frac{\mu_o N^2}{2\pi} ln\left(\frac{R+a}{R-a}\right)
\end{aligned}$$

- Coaxial Cable:

$$B = \frac{\mu_o I}{2\pi} \cdot \frac{1}{r}$$
$$\phi = \frac{\mu_o I l}{2\pi} \int_a^b \frac{1}{r} \cdot dr$$
$$\rightarrow \frac{L}{l} = \frac{\mu_0}{2\pi} ln\left(\frac{b}{a}\right)$$

1.6.4 Mutual Inductance

Mutual inductance is a measure of the extent to which the magnetic field created due a changing current in one component induces a current in the second component. It is given by:

$$M = M_{12} = M_{21} = \sqrt{L_1 L_2} \tag{1.38}$$

where:

$$M_{12} = \frac{\Phi_2}{I_1}$$
$$M_{21} = \frac{\Phi_1}{I_2}$$

To prove this result, recall (1.37).

$$\begin{aligned} L_1 L_2 &= \frac{\Phi_1}{I_1} \cdot \frac{\Phi_2}{I_2} \\ &= \frac{\Phi_1}{I_2} \cdot \frac{\Phi_2}{I_1} \\ &= M_{12} M_{21} \\ &= M^2 \end{aligned}$$

and so the result follows.

Consider the two coupled circuits as shown in Figure (1.9). Using Kirchoff's Voltage Law

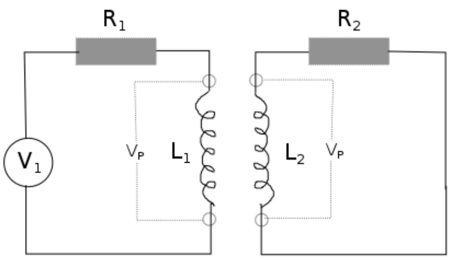

Figure 1.9: Coupled Circuits, with the first sinusoidally voltage driven

(2.5), we can write down the coupled circuit equations:

$$V_1 = I_1 R_1 + L_1 \frac{dI_1}{dt} - M \frac{dI_2}{dt}$$
$$0 = I_2 R_2 + L_2 \frac{dI_2}{dt} - M \frac{dI_1}{dt}$$

We can then solve this coupled system of equations to find the currents in each circuit. Supposing that there are no parasitic voltage losses, consider the voltage drops V_p and V_s.

$$V_p - N_p\frac{d\Phi_p}{dt} = 0$$
$$V_s - N_s\frac{d\Phi_s}{dt} = 0$$

If we assume that there is perfect flux linkage between the coils, then $d\Phi_p = d\Phi_s$. Thus, re-arranging the above equations, we arrive at:

$$\frac{V_s}{V_p} = \frac{N_s}{N_p} \tag{1.39}$$

This is the 'transformer equation' that described how voltage is 'stepped-up' or 'stepped-down' by a transformer (which is essentially just two coupled coils).

1.7 Electromagnetic Waves

Electromagnetic Waves (that is to say, light) arise when we consider Maxwell's equations (see Section 1.1) in a vacuum (i.e with $\rho = 0$ and $\underline{J} = 0$):

$$\nabla \cdot \underline{E} = 0 \tag{1.40}$$
$$\nabla \cdot \underline{B} = 0 \tag{1.41}$$
$$\nabla \times \underline{E} = -\frac{\partial \underline{B}}{\partial t} \tag{1.42}$$
$$\nabla \times \underline{B} = \mu_o \varepsilon_o \frac{\partial \underline{E}}{\partial t} \tag{1.43}$$

Consider $\nabla \times (\nabla \times \underline{E})$.

$$\begin{aligned}\nabla \times (\nabla \times \underline{E}) &= \nabla(\nabla \cdot \underline{E}) - \nabla^2 \underline{E} \\ &= -\nabla^2 \underline{E}\end{aligned}$$

From (1.40). Then, using (1.42) and (1.43):

$$\begin{aligned}\nabla \times (\nabla \times \underline{E}) &= -\frac{\partial}{\partial t}(\nabla \times \underline{B}) \\ -\nabla^2 \underline{E} &= -\frac{\partial}{\partial t}\left(\mu_o \varepsilon_o \frac{\partial \underline{E}}{\partial t}\right)\end{aligned}$$

Hence:

$$\nabla^2 \underline{E} = \mu_o \varepsilon_o \frac{\partial^2 \underline{E}}{\partial t^2} \tag{1.44}$$

Following the analogous process for the $\underline{B}$ field, we obtain the equation:

$$\nabla^2 \underline{B} = \mu_o \varepsilon_o \frac{\partial^2 \underline{B}}{\partial t^2} \tag{1.45}$$

These two equations have solutions of the form:

$$\underline{E} = E_o e^{i(\omega t - \underline{k} \cdot \underline{r})}$$
$$\underline{B} = B_o e^{i(\omega t - \underline{k} \cdot \underline{r})}$$

In general, ω is the angular frequency and k the wave-number of the wave.

1.7.1 Properties of Electromagnetic Waves

Electromagnetic waves have a series of fundamental properties, as follows:

- If we substitute the above solutions into the corresponding wave equations, then we find that:

$$k^2 = \mu_o \varepsilon_o \omega^2$$
$$\rightarrow \text{Phase Velocity} = \frac{w}{k} = \frac{1}{\sqrt{\mu_o \varepsilon_o}}$$

  Interestingly, we find that the product of the two constants μ_o and ε_o is closely related to the speed of light.

- If we substitute the above solutions into Maxwell's equations in vacuo, we obtain the following sets of results:

$$\underline{k} \times \underline{E} = \omega \underline{B}$$
$$\underline{k} \times \underline{B} = -\omega \underline{E}$$
$$\underline{B} = \frac{\underline{k} \times \underline{E}}{\omega}$$
$$\underline{E} \cdot \underline{B} = \underline{E} \cdot \left(\frac{\underline{k} \times \underline{E}}{\omega} \right)$$
$$\rightarrow \underline{E} \cdot \underline{B} = 0$$

  This shows that the electric and magnetic fields are perpendicular to one another. Furthermore:

$$\underline{k} \cdot \underline{B} = 0$$
$$\underline{k} \cdot \underline{E} = 0$$

  Thus, the $\underline{E}$ and $\underline{B}$ fields, and the direction of propagation ($\underline{k}$) form an orthogonal right-handed basis set.

- Given that we know the electric field, we can consequently find the magnetic field (or vise-versa) using:

$$\underline{B} = \frac{1}{\omega} \underline{k} \times \underline{E} \longleftrightarrow \underline{E} = -\frac{c^2}{\omega} \underline{k} \times \underline{E}$$

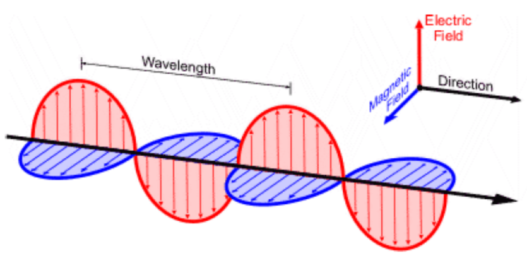

Figure 1.10: A simple Electromagnetic Wave

1.7.2 Polarisation

Polarisation is a property of waves that can oscillate with more than one orientation, and as we have seen, Electromagnetic Waves exhibit this property. There are three main types of polarisation:

- Linear Polarisation - Where the $\underline{E}$ and $\underline{B}$ field components are of equal amplitude and are in phase.
- Circular Polarisation - Where the $\underline{E}$ and $\underline{B}$ field components are of equal amplitude, but are $\frac{\pi}{2}$ out of phase with one another.
- Ellipitcal Polarisation - Where the $\underline{E}$ and $\underline{B}$ field components are of unequal amplitude, and are also $\frac{\pi}{2}$ out of phase with one another.

In this course, we do not have to perform computations involving non-linearly polarised waves (thank goodness!) but it is still worth being able to recognise these properties.

2. *Circuit Theory*

This chapter aims to cover the basics of Circuit theory including:

- Resistance, Capacitance and Inductance
- Conventions and Kirchoff's Laws
- DC Circuits
- Complex Impedance and AC circuits
- Ideal Operational Amplifiers

Students who have only covered generic circuit analysis using Ohm's Law ($V = IR$ in it's basic form) may find some of the methods outlined in this section initially difficult to grasp as instead they focus on analysis via Kirchoff's Laws (2.2.2). However, the mathematics involved is quite simple (only requiring knowledge of basic complex number arithmetic and integral calculus) and so any problems encountered should be only passing. Note that for complex number analysis j is used in the place of i so as not to confuse it with current. For clarity: 'DC' refers to 'direct current' circuits where the current is always in one direction; 'AC' refers to 'alternating current circuits' where the direction of the current reverses every half-cycle.

2.1 Resistance, Capacitance and Inductance

Most circuits will consists of a combination of sources (current and/or voltage sources), resistors, capacitors and inductors. The only other component that we have to worry about is the Ideal OP-Amp, but this is covered later (see Section (2.5)).

2.1.1 Resistors

Resistors are the only components that dissipate power within a circuit. The voltage drop across a resistor is given simply by Ohm's Law:

$$V = I \cdot R \tag{2.1}$$

These are the simplest components to deal with as the current running through them is always in phase with the source current, which minimises complications (more on this later in Section (2.4)). We add resistances in series, and add their inverses in parallel.

2.1.2 Capacitors

These store energy via the separation of two oppositely poled charge accumulations. When fully charged, it behaves like an open circuit as no current is able to flow. If uncharged, it will initially behave like a short circuit. *It is impossible to change the voltage on a capacitor instantly.*

$$C = \frac{Q}{V} \longleftrightarrow V_c = \frac{1}{C}\int I dt \tag{2.2}$$

2.1.3 Inductors

Inductors store energy in a magnetic field created due to a changing current through the inductor. For steady currents, there is no voltage drop across the inductor, and it will behave like a short-circuit. *It is impossible to change the current though and inductor instantly.*

$$V = L\frac{dI}{dt} \tag{2.3}$$

2.2 Conventions and Kirchoff's Laws

This section outlines some of the basic theorems that are used throughout this circuit theory course, as well as the all-important Kirchoff's Laws. This section involved quite a lot of listing to cover all of the basic concepts; they are not particularly difficult, there are just quite a few of them.

2.2.1 Passive Sign Convention

For passive components (that is, components that do not drive the circuit, such as sources), we get to choose either the direction of the current or voltage, but not both. In general, it is a good idea to choose the voltage based on the current direction.

Figure 2.1: Passive Sign Convention for an impedance Z

In this case, the positive terminal of the component is labelled as the one that the current flows into. This means that the sign of the voltage that we write down is the first sign encountered when doing Kirchoff Loops.

2.2.2 Kirchoff's Laws

- Kirchoff's Current Law (KCL) - This is a simple statement of the conservation of charge; *the sum of currents into a junction is equal to the sum of currents out of a junction.* This means that we can find mathematical relationships between currents that are linked at junctions. For example:

  $$\text{Current In} = \text{Current Out}$$

  $$\sum_n^\infty I_n = 0 \tag{2.4}$$

- Kirchoff's Voltage Law (KVL) - This is a statement of the conversation of energy. It states that *around any closed loop in a circuit, the net change in voltage is zero.*

  $$\sum_n^\infty V_n = 0 \tag{2.5}$$

  We can use this to perform analysis on circuits by writing down all the closed loop equations in the circuit, and then solving the simultaneous system using the relationships between the various currents derived using KCL.

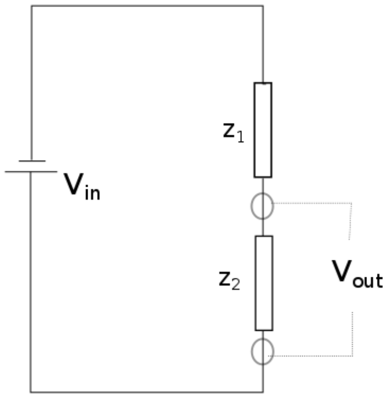

Figure 2.2: Voltage Divider for a simple circuit

2.2.3 Divider Theorems

- Voltage Divider - Can be used to find the voltage at a particular point or 'node' in a circuit based on the voltage drop across a previous component. By considering the ratios of the voltages and resistances in Figure (2.2), it can be shown that:

$$V_{out} = V_{in}\left(\frac{z_2}{z_1 + z_2}\right) \tag{2.6}$$

- Current Divider - Evidently, the current that flows down a branch in a parallel circuit is inversely proportional to the resistance of that particular branch. We can thus write:

$$I_n = I_{total}\,\frac{R_{total}}{R_n} \tag{2.7}$$

  We can thus use this to find the current flowing down a particular branch in DC circuit analysis.

2.2.4 Equivalence Theorems

These two theorems allow us to simplify circuits greatly by replacing complex elements of the circuits by 'equivalent' components that replicate the same conditions, in a similar fashion to The Image Charge Method (1.3.2).

- Thevenin's Theorem - Any linear network of voltage/current sources and resistors can be written as an equivalent circuit containing only the equivalent voltage (V_{eq}) and the equivalent resistance (R_{eq}). We can find V_{eq} by considering the voltage drop across the open circuit. The R_{eq} is the resistance "seen" (what we mean by this is that we can ignore how current flows through the particular components) across the interval with all voltage sources replaced by short circuits, and all current sources replaced by open circuits. This may seem confusing, so have a look at the example below.

- Norton's Theorem - This is very similar to Thevenin's Theorem, except in this case we obtain a current source and resistor in parallel. We can convert between the two very easily, as R_{eq} will not change, and we can simply use Ohm's Law for $V_{eq} \to I_{eq}$.

As an example to illustrate these concepts, consider the circuit shown below:
Considering Thevenin's Theorem, we replace the voltage source with a short circuit, and the current source with an open circuit. This means that for the equivalent resistance, we

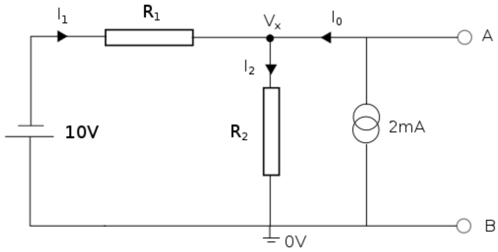

Figure 2.3: Simple Circuit with Voltage and Current Source

obtain the following circuit:

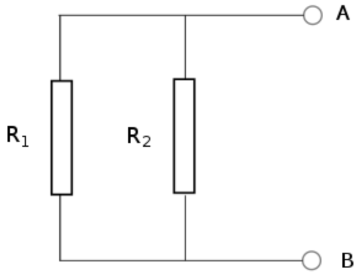

Figure 2.4: The Thevenin equivalent resistance

Consequently:

$$\begin{aligned} R_{eq} &= \left(\frac{1}{R_1} + \frac{1}{R_2}\right)^{-1} \\ &= \frac{R_1 R_2}{R_1 + R_2} \\ &= 0.5\, k\Omega \end{aligned}$$

Gasp, they have used numbers! We can then find the equivalent voltage using node voltage analysis. This has been done in the next Section (2.2.5). It turns out that $V_{eq} = 6V$.

2.2.5 Methods of Analysis

Other than using Kirchoff's Voltage Law (2.5), there are two other main techniques for analysing circuits. However, one tends to use them less than KVL/KCL as they can only be applied to DC circuits.

- Node Voltage Analysis:
    1. Choose a ground node for reference, preferably the one with the most currents running into it, as this simplifies the calculations
    2. Label the voltages on all nodes
    3. Apply KCL (2.4) and Ohm's law to find expressions for the currents running in and out of each node

Consider again the circuit shown in Figure (2.3). Applying node voltage analysis:

$$I_1 + I_o - I_2 = 0$$
$$I_1 = \frac{10 - V_x}{R_1}$$
$$I_2 = \frac{V_x}{R_2}$$
$$I_o + \frac{10 - V_x}{R_1} - \frac{V_x}{R_2} = 0$$

Rearranging, and recalling that $I_o = \ 2\,mA$, we obtain:

$$V_x = \frac{I_1 R_1 R_2 + 10R_2}{R_1 + R_2}$$
$$\rightarrow V_x = 6V$$

We have thus shown the result quoted above.

- Mesh Current Analysis:
  1. Label all the interior current loops
  2. Apply Passive Sign Convention (2.2.1) according to the mesh currents; the sign of the smaller currents can be sorted out later
  3. Apply both KCL and KVL, relating the mesh currents to one another
  4. Use Ohm's Law to solve the resultant equations

Consider the following circuit.

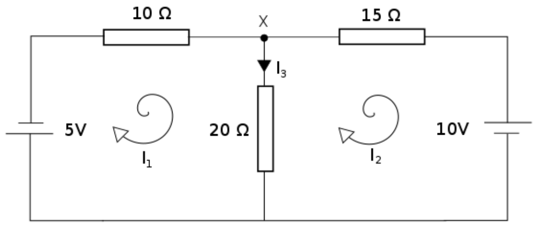

Figure 2.5: Mesh Current Circuit

First, apply KCL at node X:

$$I_3 - I_1 - I_2 = 0$$
$$I_3 = I_1 + I_2$$

Then, apply KVL around each loop containing a mesh current, and apply the above equation to reduce the equations to a system in two variables where appropriate:

$$-5 + 10I_1 + 20I_3 = 0$$
$$\rightarrow -5 + 30I_1 + 20I_2 = 0$$

and

$$
\begin{aligned}
-10 + 15I_2 + 20I_3 &= 0 \\
\rightarrow -10 + 35I_2 + 20I_1 &= 0
\end{aligned}
$$

We can then solve these two equations simultaneously to find the currents. It turns out that $I_1 = -\frac{1}{26}A$, $I_2 = \frac{8}{26}A$, $I_3 = \frac{7}{26}A$. These can be checked by the reader.

2.3 DC Circuits

Generally, these are the easier types of questions, as the hardest part is finding the appropriate differential equation for the circuit; the rest is just a bit of mathematics. In general, we look at the conditions of the circuit at $t = 0 + \delta t$, $t = 0 - \delta t$ and as $t \to \infty$ to give us the initial conditions of the circuit, and the rest is applying Kirchoff's Laws (2.2.2) to the available loops in the circuit. Below are some simple examples.

2.3.1 RC Circuit

Consider the circuit shown in Figure (2.6). The capacitor is initially uncharged. At $t = 0$ the switch is moved from position A to position B. *What is the subsequent response of the circuit?*

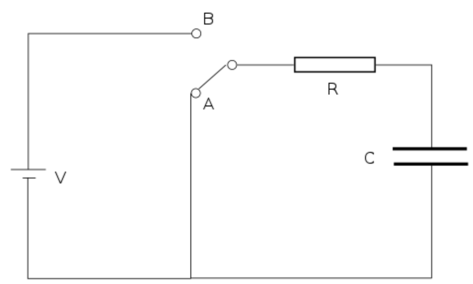

Figure 2.6: A typical RC circuit

By KVL:

$$\begin{aligned} V_o - V_R - V_C &= 0 \\ V_o &= V_R + V_C \\ &= IR + \frac{Q}{C} \end{aligned}$$

Differentiating with respect to time to obtain the differential equation:

$$\begin{aligned} 0 &= R\frac{dI}{dt} + \frac{1}{C}\frac{dQ}{dt} \\ &= R\frac{dI}{dt} + \frac{I}{C} \end{aligned}$$

Solving this separable equation and imposing the initial condition that $I_o = \frac{V_o}{R}$, we obtain the final solution of:

$$I(t) = \frac{V_o}{R} \cdot e^{-\frac{1}{RC}t}$$

Note the behaviour of the circuit at large times; as $t \to \infty$, $I \to 0$. This makes sense, as the capacitor becomes fully charged and current can no longer flow.

2.3.2 RL Circuit

Consider the circuit shown in Figure (2.7). At $t = 0$ the switch is moved from position A to position B. *What is the subsequent response of the circuit?*

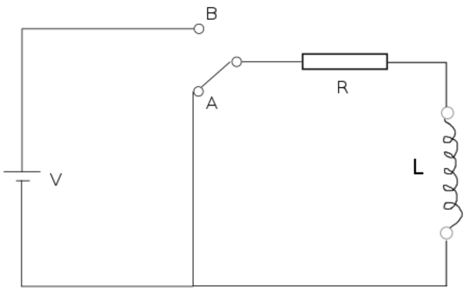

Figure 2.7: A typical RL circuit

By KVL:

$$V_o - V_R - V_C = 0$$
$$V_o - IR - L\frac{dI}{dt} = 0$$
$$V_o - IR = L\frac{dI}{dt}$$

Solving this separable equation and imposing the initial condition that $I_o = 0$ as you cannot change the current through an inductor instantly (and it was initially zero!), we obtain the final solution of:

$$I(t) = \frac{V_o}{R} \cdot \left(1 - e^{-\frac{R}{L}t}\right)$$

Again, note the behaviour of the circuit at large times; as $t \to \infty$, $I \to \frac{V_o}{R}$. This makes sense, at large times the inductor behaves like a short circuit.

2.3.3 RCL Circuit

Consider the circuit shown in Figure (2.8). The capacitor is initially charged with Q_o. At $t = 0$ the switch closed. *What is the subsequent response of the circuit?*

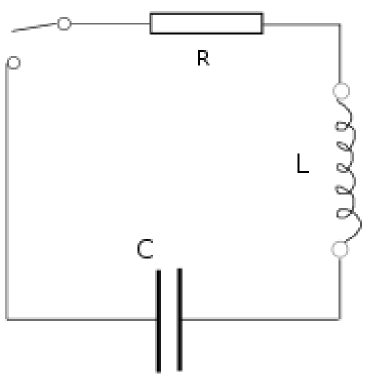

Figure 2.8: A typical RCL circuit

By KVL:

$$V_R + V_C + V_L = 0$$
$$IR + L\frac{dI}{dt} + \frac{Q}{C} = 0$$

Differentiating with respect to time:

$$R\frac{dI}{dt} + L\frac{d^2I}{dt^2} + \frac{1}{C}\frac{dQ}{dt} = 0$$
$$L\frac{d^2I}{dt^2} + R\frac{dI}{dt} + \frac{I}{C} = 0$$

This is a linear, second order differential equation. Using a text solution of $I = Ae^{nt}$:

$$Ln^2 + Rn + \frac{1}{C} = 0$$
$$n = \frac{1}{2L}\left(-R \pm \sqrt{R^2 - \frac{4L}{C}}\right)$$

This gives rise to a number of cases depending on the relative magnitudes of R, C and L:

- For ${w_1}^2 = R^2 - \frac{4L}{C} > 0$, the response is over-damped and of the form:

$$I(t) = e^{-\frac{R}{2L}t} \cdot \left(c_1 e^{w_1 t} + c_2 e^{w_2 t}\right)$$

- For ${w_2}^2 = \frac{4L}{C} - R^2 > 0$, the response is oscillatory and of the form:

$$I(t) = e^{-\frac{R}{2L}t} \cdot \left(c_3 \cos w_2 t + c_4 \sin w_2 t\right)$$

- For ${w_3}^2 = \frac{4L}{C} - R^2 = 0$, the response is critically-damped and of the form:

$$I(t) = e^{-\frac{R}{2L}t} \cdot (c_5 + c_6 t)$$

where c_1, c_2, c_3, c_4, c_5 and c_6 are all constants that can be determined from initial conditions (this has been left an exercise to the reader).

2.4 Complex Impedance and AC circuits

'Complex impedance' a measure of the opposition of a particular component to the flow of current in a circuit. It is used extensively in AC circuit analysis, as it takes account both of the *magnitude* and *phase* of the response with respect to the source current. For a component, complex impedance is defined as:

$$z \equiv \frac{V}{I} \tag{2.8}$$

where $V = V_o e^{j\omega t}$. Impedance behaves like resistance in the sense that we add impedances in series, and add their inverses in parallel.

- For a resistor:

$$\rightarrow z_R = R$$

- For a capacitor:

$$\begin{aligned} z_C &= \frac{V_C}{I} \\ &= \frac{\frac{1}{C}\int I dt}{I} \\ \rightarrow z_C &= -\frac{j}{\omega C} \end{aligned}$$

- For an inductor:

$$\begin{aligned} V_L &= L\frac{dI}{dt} \\ &= j\omega L I_o e^{j\omega t} \\ \rightarrow z_L &= j\omega L \end{aligned}$$

We can also use Kirchoff Loops with complex impedance to analyse AC circuits.

2.4.1 Bridge Circuits

This is a common type of problem that we can use AC circuit analysis circuit to analyse. A bridge circuit (see Figure (2.9)) is typically used to determine the value of the capacitance/resistance/inductance of particular components as it will only be 'balanced' for particular values of the predetermined components. This is a very accurate method for determining component values as it is independent of the source voltage.

For the bridge to be balanced, we required $V_A = V_B$. By (2.6):

$$\begin{aligned} V_o\left(\frac{z_2}{z_1 + z_2}\right) &= V_o\left(\frac{z_4}{z_3 + z_4}\right) \\ z_2(z_3 + z_4) &= z_4(z_1 + z_2) \end{aligned}$$

Hence, we obtain the 'bridge equation':

$$z_2 z_3 = z_1 z_4 \tag{2.9}$$

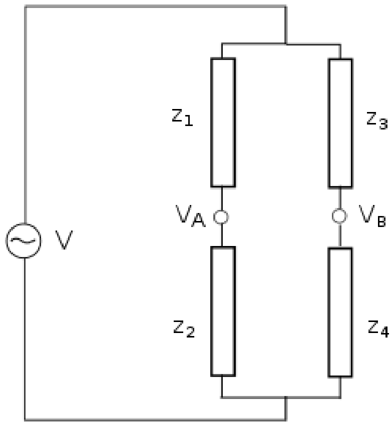

Figure 2.9: A general Bridge Circuit

2.5 Ideal Operational Amplifiers

Operational Amplifiers, otherwise known as OP-Amps, are powered components used as filters in circuits. They are particularly useful as they amplify the input signal instead of resulting in attenuation or loss.

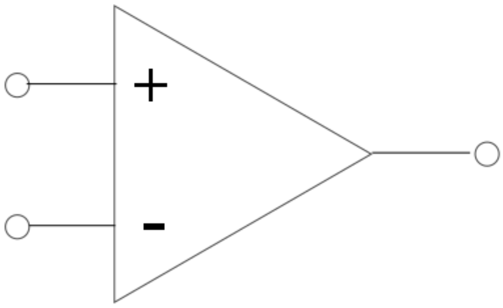

Figure 2.10: An OP-Amp

In general, we say that *under ideal OP-Amp conditions, the voltage at both terminals is equal.* This allows us to put constraints on the circuit that it is connected to to find the output voltage via node voltage analysis (2.2.5).

The input resistance of an OP-Amp is the resistance 'seen' by the OP-Amp from the source voltage. In general, the output resistance is zero.

Note that we will adopt the shorthand notation that $V_{in} \equiv V_i$, and that $V_{out} \equiv V_o$ for the following sections.

2.5.1 Non-Inverting OP-Amp

From node voltage analysis:

$$\frac{V_1}{R_1} = \frac{V_o - Vi}{R_2}$$

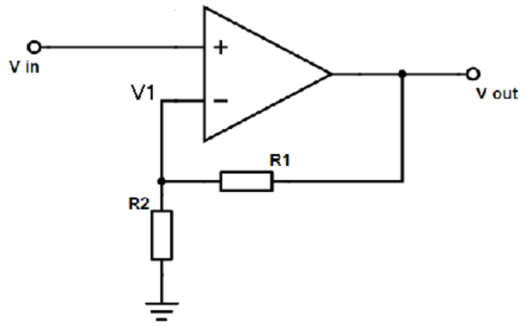

Figure 2.11: A non-inverting OP-Amp

Under ideal conditions, $V_{in} = V_1$. Thus:

$$\rightarrow \frac{V_o}{V_i} = 1 + \frac{R_2}{R_1} \tag{2.10}$$

2.5.2 Inverting OP-Amp

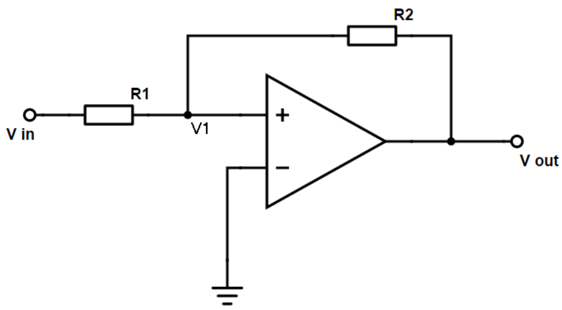

Figure 2.12: An inverting OP-Amp

From node voltage analysis:

$$\frac{V_1 - V_i}{R_1} = -\frac{V_1 - V_o}{R_2}$$

Under ideal conditions, $V_1 = 0$. Thus:

$$\rightarrow \frac{V_o}{V_i} = -\frac{R_2}{R_1} \tag{2.11}$$

2.5.3 Difference Amplifier

From node voltage analysis:

$$\frac{V_1 - V_{ref}}{R_1} = -\frac{V_o - V_1}{R_2}$$
$$\frac{V_2 - V_i}{R_1} = -\frac{V_2}{R_2}$$

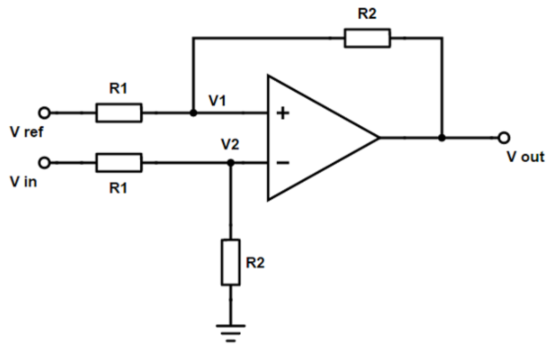

Figure 2.13: A difference amplifier

Under ideal conditions, $V_1 = V_2$. Thus:

$$\rightarrow V_o = \frac{R_2}{R_1} \cdot (V_i - V_{ref}) \tag{2.12}$$

2.5.4 Differentiating OP-Amp

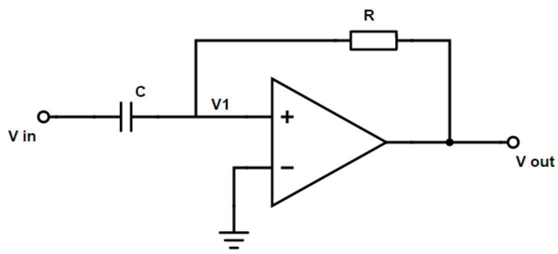

Figure 2.14: A differentiating OP-Amp

Recall that for a capacitor, $I = C\frac{dV}{dt}$. By KVL (2.5) and KCL (2.4):

$$V_o = -IR$$
$$\frac{V_o}{R} = -C\frac{dV}{dt}$$

Thus:

$$\rightarrow V_o = -RC\frac{dV_i}{dt} \tag{2.13}$$

2.5.5 Integrating OP-Amp

Recall that for a capacitor, $I = C\frac{dV}{dt}$. By KVL (2.5) and KCL (2.4):

$$\begin{aligned} V_o &= -V_c \\ &= -\frac{1}{C} \cdot \int I dt \end{aligned}$$

Thus:

$$\rightarrow V_o = -\frac{1}{RC} \cdot \int V_i \cdot dt \tag{2.14}$$

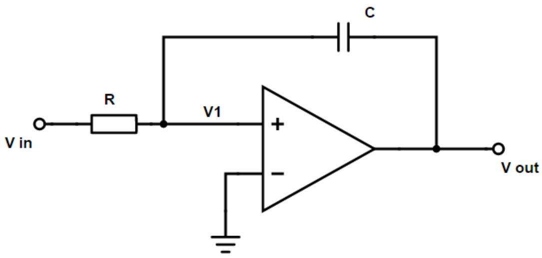

Figure 2.15: An integrating OP-Amp

3. *Geometric Optics*

This chapter aims to cover the basics of Geometric Optics, including:

- Approximations and General Principles
- Reflection and Refraction
- Standard Derivations
- Magnification, Real and Virtual Images
- Aberrations

Geometric optics is exactly what the name suggests; the solving of optics problems by considering the geometric path length of the light rays through a system. It is quite a basic, and relatively cumbersome, method but should not prove too difficult for those familiar with basic geometry. Note that we will denote the Optical Path Length by 's' throughout this chapter.

3.1 Approximations and General Principles

A lot of the work done in this topic is based off a few key approximations and principles that we will outline before tackling the main body of the material.

3.1.1 Approximations

There are three main approximations that we make use of, as otherwise the computation would become too complicated to be covered by the scope of this course. Always be sure to state what assumptions you are using when answering problems on this topic, as otherwise your working will not be valid! These are as follows:

- Paraxial Approximation - This assumes that all of the rays passing through the optical system lie very close to the optical axis, as otherwise aberrations (see Section (3.5)) start to come into large effect.
- Small Angle Approximation - As a consequence of the paraxial approximation, we can use the small angle formulae that $\sin(\theta) \approx \tan(\theta) \approx \theta$ and $\cos(\theta) \approx 1$. This allows us to simplify calculations greatly.
- Thin Lens Approximation - This assumes that the time that light takes to travel through a lens or similar optical system is negligible in comparison to the remainder of the travel time. In essence, we can thus assume that the lens is two-dimensional (a flat sheet).

3.1.2 General Principles

Essentially, all of the work that is done in this course on geometric optics will use either Fermat's Principle, or Huygen's Principle.

- Fermat's Principle - Simply put, it states that *the path adopted by a light wave is the one down which the time is a stationary value of some functional.* Mathematically, this means that we need to minimise the optical path length of the system.
- Huygen's Principle - This proposes that light is made up of a series of pulsations in the aether (a zero-viscosity, all permeating medium through which light travels) that causes a set of circular 'wavelets' to spread out from these points. These constructively and destructively interfere with one another to create a wave-front. Each point on the wave-front can then we regarded as a source of secondary wavelets. When encountering a boundary, such as a medium boundary, a spherical wave will originate from each point of the wave-front that hits the boundary. As we know from Chapter (1), this formulation of light is in fact completely erroneous, but it works for the purposes of certain geometric and, in fact, wave optics problems.

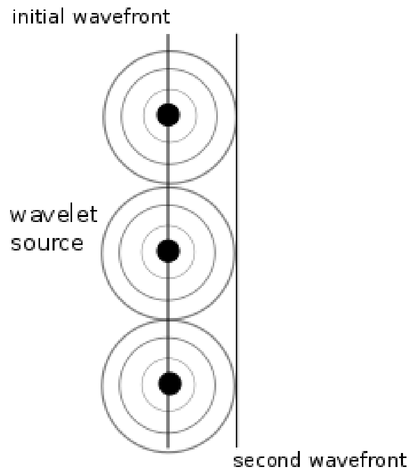

Figure 3.1: Huygen's Principle over two wave-fronts

3.2 Reflection and Refraction

Geometric optics mainly deals with reflection and refraction; diffraction and other interference concepts are dealt with more easily using Wave Optics. For this, see Chapter (4).

3.2.1 Reflection

The Law of Reflection is probably the concept that most people think of when it comes to optics. It simply states that *the angle of incidence of light equals the angle of reflection, when measured from the surface normal.*

$$\rightarrow \theta_i = \theta_r \tag{3.1}$$

This can be proven quite easily using Fermat's Theorem. Consider Figure (3.2) below.

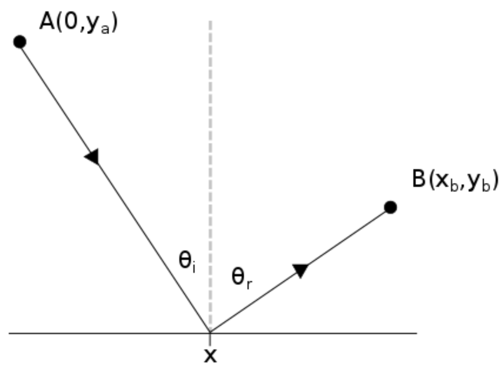

Figure 3.2: Law of Reflection by Fermat's

We shall assume that the refractive index stays constant over $0 < y < y_a$. This is valid, as we can arbitrarily choose point A, and so the derivation remains valid in the limiting case where A is essentially on the reflection surface. We can write the optical path length (s)

as:

$$
\begin{aligned}
s &= n\sqrt{x^2 + {y_a}^2} + n\sqrt{{y_b}^2 + (x_b - x)^2} \\
\frac{\partial s}{\partial x} &= n\left(\frac{x}{\sqrt{x^2 + {y_a}^2}} - \frac{x_b - x}{\sqrt{{y_b}^2 + (x_b - x)^2}}\right) \\
\frac{\partial s}{\partial x} &= 0 \\
\rightarrow \frac{x}{\sqrt{x^2 + {y_a}^2}} &= \frac{x_b - x}{\sqrt{{y_b}^2 + (x_b - x)^2}}
\end{aligned}
$$

By inspection, it is clear that $\sin\theta_i = \sin\theta_r$, and the result above follows.

3.2.2 Refraction

The refractive index (n) of a medium is the measure of the speed at which light travels in said medium; the higher the refractive index, the slower the speed.

$$n = \frac{c}{v} \tag{3.2}$$

When light encounters a medium boundary, it will bend towards or away from the normal depending on the relative refractive indices of the two media. In general, when light moves from a medium with a higher refractive index to that of a lower refractive index, it bends away from the normal and vise-versa. This can be thought of as *light bends away from the normal when it speeds up*.

The key to solving refraction problems is being familiar with *Snell's Law*, which is stated mathematically as:

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \tag{3.3}$$

This can be proven using either Fermat's Principle or Huygen's Principle, both of which are outlined below:

- Derivation by Fermat's:

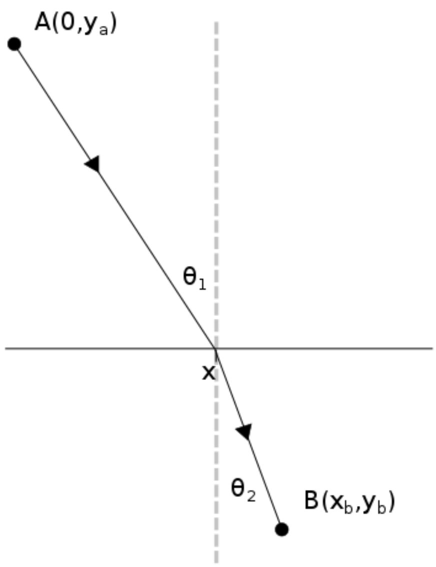

Figure 3.3: Law of Refraction by Fermat's

$$s = n_1\sqrt{x^2+{y_a}^2} + n_2\sqrt{{y_b}^2+(x_b-x)^2}$$
$$\frac{\partial s}{\partial x} = n_1\frac{x}{\sqrt{x^2+{y_a}^2}} - n_2\frac{x_b-x}{\sqrt{{y_b}^2+(x_b-x)^2}}$$
$$\frac{\partial s}{\partial x} = 0$$
$$n_1\frac{x}{\sqrt{x^2+{y_a}^2}} = n_2\frac{x_b-x}{\sqrt{{y_b}^2+(x_b-x)^2}}$$

However, we remark that:

$$\sin\theta_1 = \frac{x}{\sqrt{x^2+{y_a}^2}}$$
$$\sin\theta_2 = \frac{x_b-x}{\sqrt{{y_b}^2+(x_b-x)^2}}$$

And hence the result of (3.3) follows.

- Derivation by Huygen's:

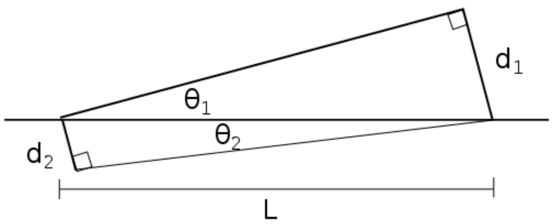

Figure 3.4: Law of Refraction by Huygen's

We require that the time taken for each end of the wave-front to travel d_1 and d_2 must be equal in order for the wave-front to remain cohesive.

$$t_1 = t_2$$
$$\frac{d_1}{v_1} = \frac{d_2}{v_2}$$
$$n_1\frac{l\sin\theta_1}{c} = n_2\frac{l\sin\theta_2}{c}$$
$$\rightarrow n_1\sin\theta_1 = n_2\sin\theta_2$$

And once again we obtain the desired result.

This refractive property of light leads to an underwater object having a reduced apparent depth when viewed from above. Consider Figure (3.5). Let the object be located at O and it's apparent image located at I.

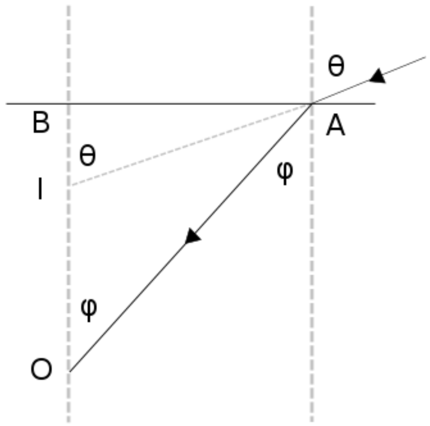

Figure 3.5: Apparent Reduced Depth

Let the first medium be air with $n = 1$. Using the small angle approximations (3.1.1):

$$\begin{aligned}\sin\theta &= n\sin\varphi \\ n &= \frac{\sin(\theta)}{\tan\varphi} \\ &\approx \frac{\tan\theta}{\tan\varphi} \\ &= \frac{AB}{IB}\cdot\frac{AB}{OB} \\ \rightarrow IB &= \frac{OB}{n}\end{aligned}$$

That is, the image depth is the object depth reduced by a factor of n. This is why water always looks shallower than it actually is.

3.2.3 Optical Fibres

Optical fibres are used to transport signals via *total internal reflection* (TIR). This is a process whereby light enters a medium with a lower refractive index at an angle greater than what is known as the "critical angle", and is reflected instead of passing through. The condition for total internal reflection is thus:

$$\theta_i > \theta_c \geqslant \arcsin\left(\frac{n_2}{n_1}\right) \tag{3.4}$$

The cable consists of an optically conductive medium surrounded by a material of a higher refractive index known as the cladding. This causes TIR to occur, and so assuming that the light enters the fibre at less than a certain angle, it should be totally internally reflected along the length of the fibre. This is more effective than simply using mirrors as in TIR, all of the signal is reflected, instead of only $\approx 98\%$ in the case of mirrors.

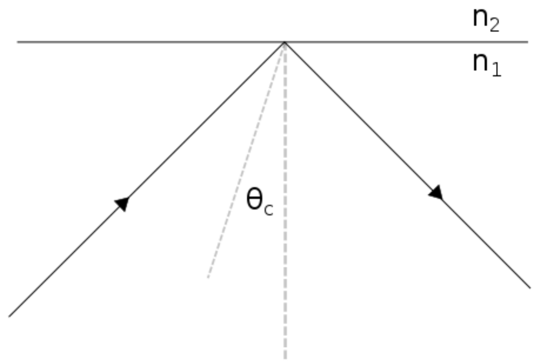

Figure 3.6: Total Internal Reflection

As previously stated, there is a maximum angle from the axis of the fibre at which light can enter and still be totally internally reflected. Consider Figure (3.7).

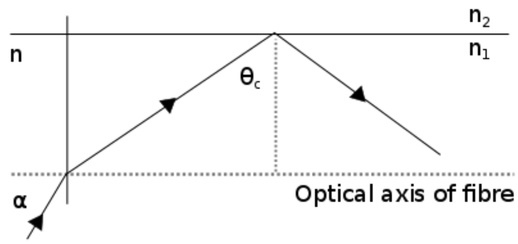

Figure 3.7: Light travelling through an optical fibre

By Snell's Law (3.3):

$$\begin{aligned} n\sin\alpha &= n_1\sin(90^\circ - \theta_c) \\ &= n_1\cos\theta_c \\ &= n_1\sqrt{1-\sin^2\theta_c} \\ &= n_1\sqrt{1-\left(\frac{n_2}{n_1}\right)^2} \\ \rightarrow n\sin\alpha &< \sqrt{n_1^2 - n_2^2} \end{aligned}$$

This is the limiting condition on the incident angle.

3.3 Standard Derivations

This sections simply covers some of the standard derivations associated with this geometric optics course.

3.3.1 The Lens Equation

This is probably the most fundamental result in all of geometric optics, except possibly that light travels in straight lines. If u is the distance to the object from the lens, v is the distance to the image and f is the focal length of the lens, then:

$$\frac{1}{u}+\frac{1}{v}=\frac{1}{f} \tag{3.5}$$

This can be derived by considering the image formed by a thin convex lens, as shown in Figure (3.8). All the parallel light from ∞ passes through the focus.

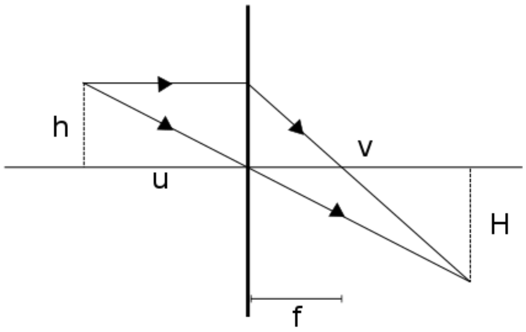

Figure 3.8: Deriving the Lens Equation

By similar triangles:

$$\begin{aligned}
\frac{H+h}{v} &= \frac{h}{f} \\
\frac{H}{v}+\frac{h}{v} &= \frac{h}{f} \\
\frac{h}{u} &= \frac{H}{v} \\
\rightarrow h\left(\frac{1}{u}+\frac{1}{v}\right) &= \frac{h}{f}
\end{aligned}$$

And so the result follows.

3.3.2 Lens-Maker's Formula

For this derivation, we are assuming the Paraxial Approximation, and consequently the small angle approximation (3.1.1). Let the radius of curvature of the mirror be R.

$$\begin{aligned}
\theta_1 &= \alpha_1 + \varphi \\
\theta_2 &= \varphi - \alpha_2
\end{aligned}$$

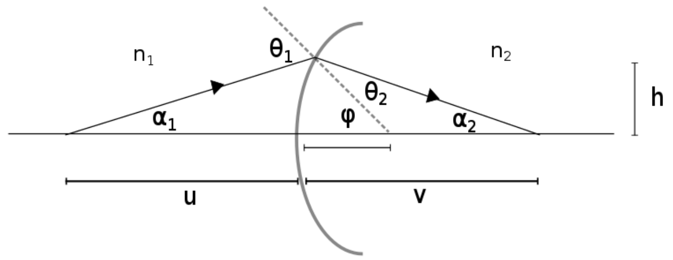

Figure 3.9: Deriving the Surface Power Equation

By Snell's Law (3.3):

$$\begin{aligned} n_1 \sin\theta_1 &= n_2 \sin\theta_2 \\ n_1 \sin(\varphi + \alpha_1) &= n_2 \sin(\varphi - \alpha_2) \\ n_1 (\sin\varphi\cos\alpha_1 + \cos\varphi\sin\alpha_1) &= n_2 (\sin\varphi\cos\alpha_2 - \cos\varphi\sin\alpha_2) \\ n_1 (\sin\varphi + \sin\alpha_1) &= n_2 (\sin\varphi - \sin\alpha_2) \\ n_1 \left(\frac{h}{R} + \frac{h}{u}\right) &= n_2 \left(\frac{h}{R} - \frac{n}{v}\right) \end{aligned}$$

Rearranging, we obtain the surface power equation:

$$\frac{n_1}{u} + \frac{n_1}{v} = (n_2 - n_1) \cdot \frac{1}{R} \tag{3.6}$$

Then, if we consider two lenses of different radii and let the distance between them tend to zero, then we obtain the Lens-Maker's Formula:

$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v} = \left(\frac{n_{lens}}{n_{medium}} - 1\right) \cdot \left(\frac{1}{R_1} - \frac{1}{R_2}\right) \tag{3.7}$$

3.3.3 Minimum distance between Object and Image

Not that this is intuitively obvious, but for any object/lens combination, there is a minimum distance between the object and the image. Let D be the distance between the object and the image:

$$\begin{aligned} D &= u + v \\ \frac{\partial D}{\partial u} &= 1 + \frac{\partial v}{\partial u} \\ &= 1 + \frac{\partial}{\partial u}\left(\left(\frac{1}{f} - \frac{1}{u}\right)^{-1}\right) \\ &= 1 - \frac{1}{u^2}\left(\left(\frac{1}{f} - \frac{1}{u}\right)^{-2}\right) \\ &= 1 - \frac{v^2}{u^2} \\ \frac{\partial D}{\partial u} &= 0 \\ \rightarrow v &= u \end{aligned}$$

Plugging this result into the Lens Equation (3.5), we find that:

$$D = 4f \tag{3.8}$$

3.3.4 The Mirror Formula

This is similar to the Lens Equation (3.5), except that it is applied to mirrors. Consider Figure (3.10) below. Let the image be located at I, the object at O and the centre of curvature of the mirror at C.

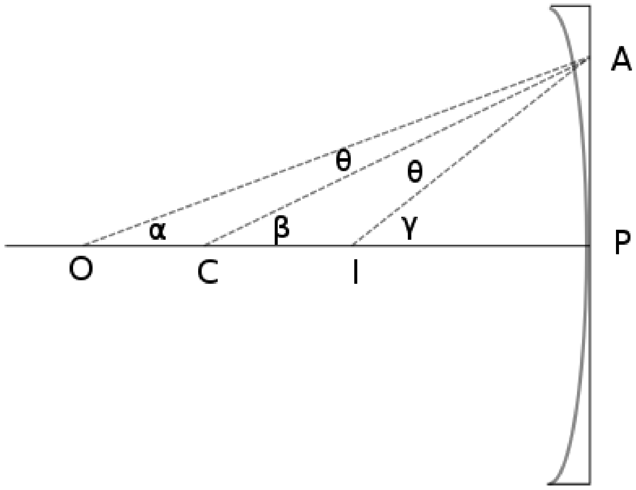

Figure 3.10: Deriving the Mirror Formula

$$\begin{aligned} \beta &= \alpha + \theta \\ \gamma &= \alpha + 2\theta \\ \alpha + \gamma &= 2\beta \end{aligned}$$

Using the paraxial and small angle approximations:

$$\begin{aligned} \tan\gamma \approx \gamma \approx \frac{AP}{IP} \\ \tan\beta \approx \beta \approx \frac{AP}{CP} \\ \tan\alpha \approx \alpha \approx \frac{AP}{OP} \end{aligned}$$

Substitute these results into $\alpha + \gamma = 2\beta$:

$$\frac{1}{u} + \frac{1}{v} = \frac{2}{R} \longrightarrow f = \frac{R}{2} \tag{3.9}$$

Thus, the focal length of a spherical mirror surface is equal to half of it's radius of curvature.

3.4 Magnification, Real and Virtual Images

This section deals with some more of the underlying concepts that are required to answer optics questions.

3.4.1 Magnification

Magnification is a measure of the relative size of the object and the image. Conventionally, if the image is inverted, then the sign of the magnification is negative, though this has no bearing on the final size of the image. There are three main types of magnification:

- Transverse Magnification - This is defined as the ratio of the apparent size of the object to it's real size. Using similar triangles, it is easy to show that:

$$M_T = \frac{v}{u} = \frac{H}{h} \tag{3.10}$$

- Angular Magnification - This is defined as the ratio of the angular size (the angle subtended at the eye) of the image to the angular size of the object.

$$M_\alpha = \frac{\theta_{image}}{\theta_{object}} \tag{3.11}$$

- Longitudinal Magnification - This is the rate of change of image distance with respect to object distance.

$$M_L = -{M_T}^2 \tag{3.12}$$

  This means that it is difficult to focus at images at a large transverse magnification as small adjustments have to be made for it to remain in focus.

3.4.2 Real and Virtual Images

There are two types of images, as follows:

- Real Images - A real image is formed at a point in space where light rays converge. A virtual image will be formed for $\infty > u > f$ in the case of convex lenses and mirrors.
- Virtual Images - A virtual image is formed when outgoing rays from a point always diverge, and the image appears to be located at the point of apparent divergence. A virtual image will be formed for $u < f$ in the case of convex lenses and mirrors. Concave surfaces always form a virtual image.

In general, to find the location, type and size of an image, it is best to draw a ray diagram. Always remember than any ray passing through the mirror at the optical axis is undeviating, and any rays initially parallel to the optical axis will be refracted through the focus (or will appear to be coming from the focus) and vice-versa. It is recommended that the reader practises drawing some ray diagrams with construction rays to become familiar with this. If stuck, have a look at Figure (3.8) for guidance.

A quick note at the end of this section on a convention that has not been mentioned thus far. It is known as the *Real is Positive Sign convention*; this essentially says that the sign of distances associated with real images is positive, while the sign of the distances associated with virtual images is negative. Always ask yourself as to whether the image is real or virtual before writing down equations.

3.5 Aberrations

These arise due to a break-down in the assumptions/approximations used in geometric optics, as well as the fact that not all surfaces that are involved are perfectly smooth.

- Spherical Aberration - The focal length of a spherical lens is reduced at the edges due to curvature, meaning that there is no clearly defined focus. This can be 'fixed' to a certain extent with a-spheric lenses.

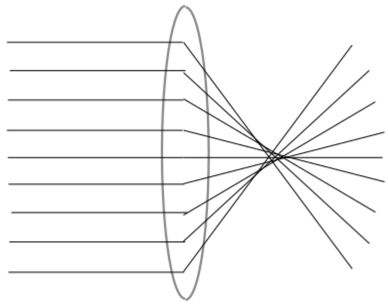

Figure 3.11: Spherical Aberration

- Coma - This arises due to off axis sources, causing the image to lose clarity around the edges. Again, it can be 'fixed' with a-spheric lenses.

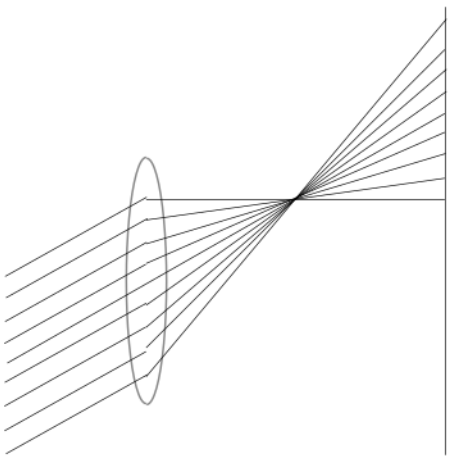

Figure 3.12: Coma

- Chromatic Aberration - The variation in the refractive index for a medium for different wavelengths (or frequencies) of light means that it is difficult to achieve a single focal point for light consisting of several wavelengths. The *achromatic doublet* uses errors in two lenses to cancel each other out, and give red and green the same focus; an achromatic triplet gives red, green *and* blue the same focus.

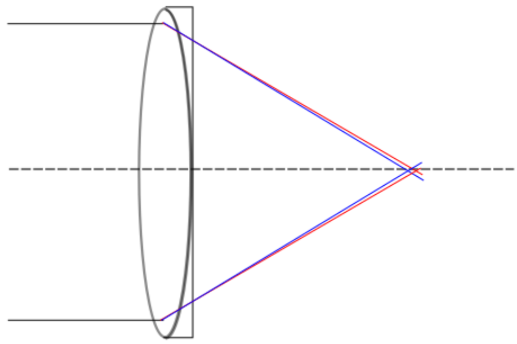

Figure 3.13: An achromatic doublet for Chromatic Aberration

4. *Wave Optics*

This chapter aims to cover the basics of Wave Optics, including:

- General Principles
- Single Slit Diffraction
- Young's Slits
- Diffraction Grating

The material in this section focusses on problems of diffraction and interference, which are solved more easily by considering light as a series of waves, instead of as a 'linear' light pulse in the case of geometric optics. This approach is more mathematical than in the case of geometric optics, though only requires knowledge of basic integration and complex numbers. Note that the principle of superposition is very important here, but it can only be applied to the amplitude of the wave and *not* the intensity. Drawing diagrams is always recommended.

4.1 General Principles

The entirety of the material covered in this chapter is based on the formulation of light as a wave; that is to say that it can be represented as a solution ψ in the following equation:

$$\nabla^2\psi = \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2}$$

This wave equation has "analytic" solutions of the form:

$$\psi(r,t) = \psi_o \cdot e^{i(\omega t - kr)} \tag{4.1}$$

$$\psi(r,t) = \frac{\psi_o}{r} \cdot e^{i(\omega t - kr)} \tag{4.2}$$

The second is in the case where we are considering Huygen's spherical waves when light encounters a slit boundary, as the slit will behave like a point source of these spherical waves.

For most questions, we want to use these analytic solutions to find the amplitude of the light hitting the plane of observation. Then, the observed intensity distribution can be found by:

$$I \propto |\psi(r,t)|^2 \tag{4.3}$$

Note that the time dependant parts of ψ will disappear when taking it's modulus, meaning that the intensity is not time dependant (thankfully!).

4.1.1 The Kirchoff Diffraction Integral

The Kirchoff Diffraction Integral is simply a way of referring to the way in which we deal with not point-like apertures; it involves summing up all of the individual 'point sources' in the slit by integrating over it's surface.

$$\psi = \psi_o \cdot e^{-i\omega t}\int_S \frac{e^{ikr}}{r} \cdot dS \tag{4.4}$$

In writing this integral, we have made the assumptions that:

- Huygen's Principle holds
- Edge effects at the extremities of the aperture are negligible
- The incident light is coherent
- The aperture is of a greater order of magnitude that a wavelength

4.1.2 The Rayleigh Criterion

The Rayleigh Criterion simply states that *two objects are considered distinguishable if and only if the limit of the maxima of one falls on the minima of the other.* This implies that the resolution of a system is limited by diffraction.

Most of the light from diffraction falls within a particular angle. The Rayleigh Criterion leads to the statement that the minimum resolvable angle for a single aperture of width D is:

$$\theta_{min} = \frac{\lambda}{D} \tag{4.5}$$

For example, if viewing the headlights of a car from a large distance away, each light can only be considered distinct when the distance between the two lights subtends an angle at the eye greater than θ_{min}.

4.2 Single Slit Diffraction

The intensity distribution due to a single slit of length a is given by:

$$I(\theta) = I_o \cdot \text{sinc}^2\left(\frac{1}{2}ka\sin\theta\right) \tag{4.6}$$

When graphed, this appears as:

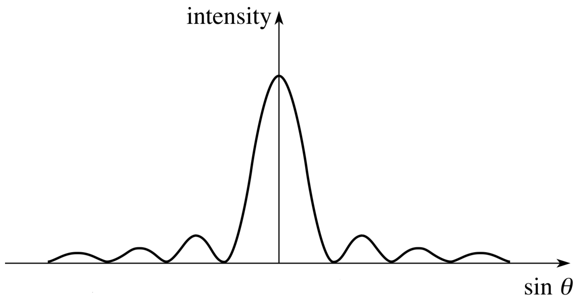

Figure 4.1: Single Slit diffraction intensity

4.2.1 Derivation

For this derivation, we will assume that the diffraction is 'far-field'; that the plane of observation is sufficiently far from the aperture such that we can consider rays leaving the aperture to be parallel (though of course that is not the case, as otherwise no intensity pattern would be observed). We need to consider the phase difference between a ray at the centre of the slit, and one at a distance x from the centre, as show in Figure (4.2).

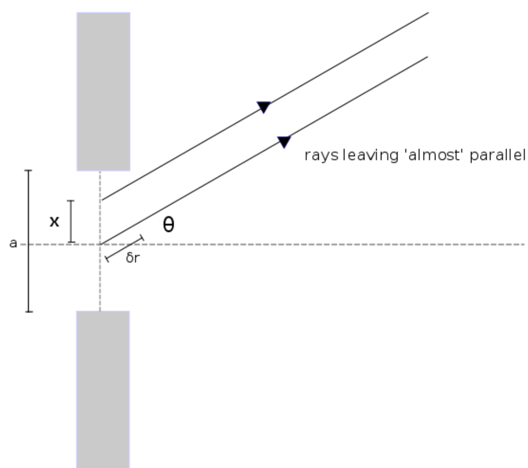

Figure 4.2: Phase difference for a single slit

From the diagram, it becomes clear that

$$\delta r = x\sin\theta \longrightarrow \phi(x) = -kx\sin\theta \tag{4.7}$$

From the Kirchoff Diffraction Integral (4.4):

$$\begin{aligned}
\psi &= \psi_o \int_S e^{i(\omega t - kr)} \cdot e^{i\phi(x)} \cdot dS \\
&\propto \psi_o \int_{\frac{a}{2}}^{\frac{a}{2}} e^{-kx\sin\theta} \cdot dx \\
&= \frac{\psi_o}{k\sin(\theta)}(-i)\left(e^{ik\frac{a}{2}\sin\theta} - e^{-ik\frac{a}{2}\sin\theta}\right) \\
&= \frac{\psi_o}{\frac{1}{2}k\sin\theta}\left(\sin(\frac{1}{2}ka\sin\theta)\right)
\end{aligned}$$

Now consider the intensity at the centre of the slit.

$$\begin{aligned}
I_o &= \left|\psi_o \int_S dS\right|^2 \\
&= {\psi_o}^2 \cdot a^2 \\
\rightarrow {\psi_o}^2 &= \frac{I_o}{a^2}
\end{aligned}$$

Finding the intensity using (4.3):

$$I(\theta) = \frac{{\psi_o}^2}{(\frac{1}{2}k\sin\theta)^2}\left(\sin(\frac{1}{2}ka\sin\theta)\right)^2$$

Then, substituting ${\psi_o}^2 = \frac{I_o}{a^2}$ we arrive at the required result:

$$\rightarrow I(\theta) = I_o \cdot \mathrm{sinc}^2\left(\frac{1}{2}ka\sin\theta\right)$$

4.2.2 Minima

The minima of the distribution will be given by the zeros of the *sinc* function. For some integer $n \neq 0$:

$$\begin{aligned}
\frac{1}{2}ka\sin\theta &= n\pi \\
\frac{1}{2}\left(\frac{2\pi}{\lambda}\right)a\sin\theta &= n\pi \\
a\sin\theta &= \lambda n \qquad (4.8)
\end{aligned}$$

We can also use the fact that θ is small to allow us to write:

$$\begin{aligned}
a\left(\frac{y}{D}\right) &= \lambda n \\
\rightarrow y &= n\frac{\lambda D}{a}
\end{aligned}$$

where y is the perpendicular distance from the optical axis in the plane of observation, and D is the distance of the plane of observation from the slit.

4.2.3 Intensity at Half Maximum

Let $\beta = \frac{1}{2}ka\sin\theta$. Then, for intensity at half maximum, we can use a taylor expansion and neglect terms of $O(\beta^6)$ and above:

$$\begin{aligned}\frac{1}{2} &= \frac{\sin^2\beta}{\beta^2} \\ \frac{1}{2}\beta^2 &\approx \beta^2 - \frac{\beta^4}{3} + \frac{2\beta^6}{45} - \dots \\ &= \beta^2 - \frac{\beta^4}{3} \\ \beta^4 &= \frac{3}{2}\beta^2 \\ \rightarrow \beta &= \sqrt{\frac{3}{2}}\end{aligned}$$

4.2.4 Laser Pulses

The energy delivered by a light wave is given by the Poynting Vector:

$$\underline{P} = \underline{E} \times \underline{B} = \frac{1}{2}c\left(\mu_o\varepsilon_o\underline{E}^2 + \underline{B}^2\right)\widehat{\underline{k}} \tag{4.9}$$

The intensity of a laser pulse is given by the time average of (4.9):

$$\begin{aligned}P_{instantaneous} &= \frac{1}{2}c\left(\mu_o\varepsilon_o\underline{E}^2 + \frac{\underline{B}^2}{c^2}\right)\widehat{\underline{k}} \\ &= \frac{E^2}{\mu_o c}\cdot\widehat{\underline{k}}\end{aligned}$$

Taking the time average:

$$I = \frac{1}{2}\cdot\frac{E^2}{\mu_o c} \tag{4.10}$$

$$= \frac{Peak\ Power}{Area} \tag{4.11}$$

The peak power is given by:

$$Peak\ \ Power\ = \frac{Energy}{Pulse\ \ Duration}$$

Now, consider the fact that at the focus of a lens or optical device, the surface area is a single point, and so has zero size. This means, theoretically, that the power delivered by light at the focus is infinite. However, as we have seen with the Single Slit diffraction, systems are diffraction limited, meaning that there is a minimum diameter over which the energy can be delivered:

$$d_{min} = \frac{\lambda}{D}\cdot f \tag{4.12}$$

This means that there is a maximum amount of power that can be delivered by light of a particular $\underline{E}$ field strength.

4.3 Young's Slits

This is diffraction caused by two 'point-source' like slits. The intensity distribution in this case for two slits separated by a distance a is given by:

$$I(\theta) = I_o \cdot \cos^2\left(\frac{1}{2}ka\sin\theta\right) \tag{4.13}$$

When graphed, this appears as:

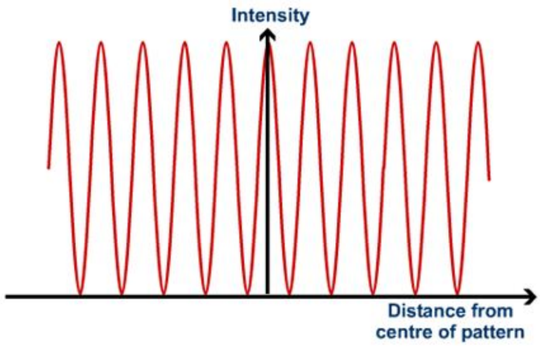

Figure 4.3: Young's Slits diffraction intensity

4.3.1 Derivation

In this case, the total amplitude arriving at the plane of observation is the sum of the amplitudes arriving from each slit (by the principle of superposition). Consider the phase difference between light arriving from each slit as shown in Figure (4.4).

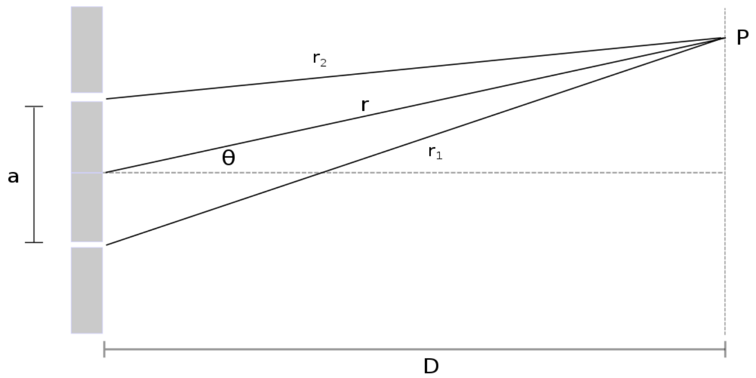

Figure 4.4: Young's Slits phase difference

By the cosine rule:

$$\begin{aligned}
{r_1}^2 &= r^2 + \left(\frac{d}{2}\right)^2 + dr\sin\theta \\
{r_2}^2 &= r^2 + \left(\frac{d}{2}\right)^2 - dr\sin\theta \\
{r_2}^2 - {r_1}^2 &= (r_2 - r_1)\cdot(r_2 + r_1)
\end{aligned}$$

For $r_1, r_2 \gg a$:

$$r_1 + r_2 \approx 2r \longrightarrow r_2 - r_1 = a\sin\theta \tag{4.14}$$

The amplitude of the waves is given by:

$$\begin{aligned} \psi &= \psi_o \left(e^{i\omega t}\right)\left(\frac{e^{ikr_1}}{r_1} + \frac{e^{ikr_2}}{r_2}\right) \\ &\approx \psi_o \cdot \frac{e^{i\omega t}}{r}\left(1 + e^{ik(r_2 - r_1)}\right) \\ &= \psi_o \cdot \frac{e^{i\omega t}}{r}\left(1 + e^{ika\sin\theta}\right) \end{aligned}$$

using (4.14). Thus, the intensity distribution is given by:

$$\begin{aligned} I(\theta) &= {\psi_o}^2 \cdot \left(\frac{e^{i\omega t}}{r}\right)^2 \cdot \left(1 + e^{ika\sin\theta}\right) \cdot \left(1 + e^{-ika\sin\theta}\right) \\ &\propto {\psi_o}^2 \cdot \left(1 + e^{ika\sin\theta}\right) \cdot \left(1 + e^{-ika\sin\theta}\right) \\ &= {\psi_o}^2 \left(2 + 2\cos(ka\sin\theta)\right) \\ \to I(\theta) &= I_o \cdot \cos^2\left(\frac{1}{2}ka\sin\theta\right) \end{aligned}$$

4.3.2 Maxima and Minima

The maxima and minima of the intensity distrubtiion are given by the maxima and minima of (4.13). For an integer $n \neq 0$:

- Maxima:

$$\begin{aligned} \frac{1}{2}ka\sin\theta &= n\pi \\ \frac{1}{2}\left(\frac{2\pi}{\lambda}\right)\sin\theta &= n\pi \end{aligned}$$

$$a\sin\theta = \lambda n \tag{4.15}$$

- Minima:

$$\begin{aligned} \frac{1}{2}ka\sin\theta &= n\pi + \frac{\pi}{2} \\ \frac{1}{2}\left(\frac{2\pi}{\lambda}\right)\sin\theta &= n\pi + \frac{\pi}{2} \end{aligned}$$

$$a\sin\theta = \lambda\left(n + \frac{1}{2}\right) \tag{4.16}$$

However, this has assumed that the incident light is monochromatic. If the slits are illuminated by two wavelengths, the number of fringes before the visibility is seriously effected will occur where the maxima of one coincides with the minima of the other (as per the Rayleigh Criterion (4.1.2)). For wavelengths λ_1 and λ_2:

$$n \cdot \lambda_2 = \left(n + \frac{1}{2}\right) \cdot \lambda_1 \tag{4.17}$$

4.3.3 Source Slit Width

We require that the light arriving from the source slit is coherent (parallel rays) for these results to hold. Let D_s be the distance from the source to the slits, and a_s be the width of the source slit. Let ϕ be the angle that is subtended by half the source at the slits. For coherent light:

$$a\sin(\phi) < \lambda$$
$$\phi < \frac{\lambda}{a}$$

But,

$$\phi = \frac{a_s}{2D_s}$$
$$\rightarrow a_s < \frac{2\lambda D_s}{a} \tag{4.18}$$

This means that there is a lower limit on the size of the source aperture.

4.3.4 Extra Path Difference

The path difference between the optical path lengths from the two slits can be modified, which will cause a shift in the intensity pattern. One of the common ways to do this is to place a piece of glass in-front of one of the slits. As light travels more slowly in the glass, this will introduce an extra path difference and thus phase difference. Suppose that the glass has thickness t. Then the extra phase difference is:

$$\delta\phi = k(n_{glass} - n_{medium})t$$

Thus, the intensity can be rewritten as:

$$I(\theta) = I_o \cdot \cos^2\left(\frac{1}{2}ka\left(\sin\theta + (n_{glass} - n_{medium})\frac{t}{d}\right)\right)$$

The central maximum *was* located at $\theta = 0$, but will now be located at:

$$y = (n_{glass} - n_{medium}) \cdot t \cdot \frac{t}{d} \tag{4.19}$$

4.4 Diffraction Grating

The intensity distribution for a diffraction grating with slits of width a and separation d is given by:

$$I(\theta) = I_o \cdot \text{sinc}^2\left(\frac{1}{2}ka\sin\theta\right) \cdot \frac{\sin^2\left(\frac{1}{2}Ndk\sin\theta\right)}{\sin^2\left(\frac{1}{2}dk\sin\theta\right)} \tag{4.20}$$

When graphed, this appears as: The intensity distribution pattern is given by the superpo-

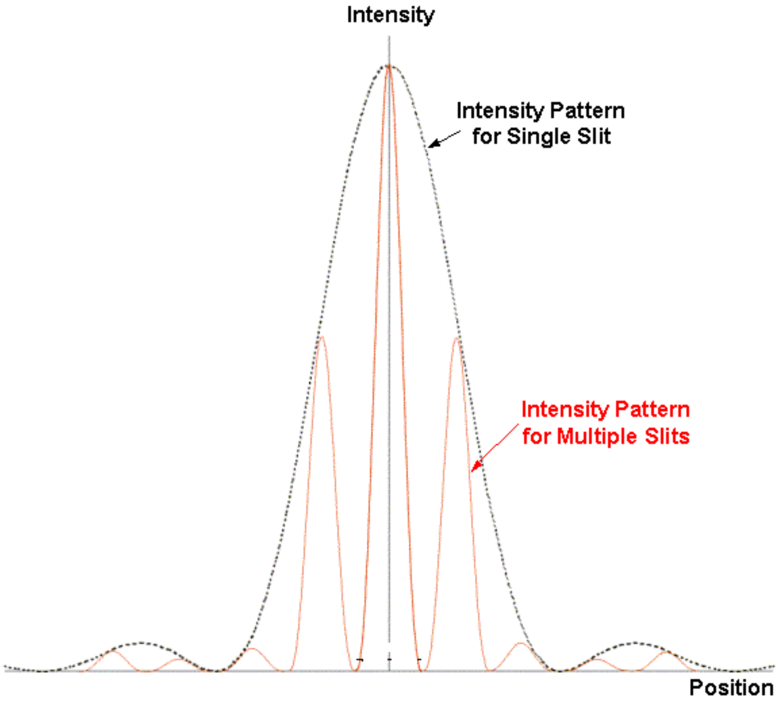

Figure 4.5: Diffraction Grating intensity

sition of that from multiple slits with that of the single slit, which provides the 'envelope' seen on the graph.

4.4.1 Derivation

The phase difference between two rays leaving equivalent points on adjacent slits is $\phi = -kd\sin\theta$. Let the light arriving from the bottom slit be ψ_o. The total amplitude arriving from all of the slits (N) is given by:

$$\psi(\theta) = \psi_o\left(1 + e^{-ikd\sin\theta} + \cdots + e^{-(N-1)kd\sin\theta}\right)$$

The $(N-1)$ appears in the last term as there are only N-1 rays out of phase with the initial ray. Summing this geometric series:

$$\psi(\theta) = \psi_o\left(\frac{1 - e^{-iNkd\sin\theta}}{1 - e^{-ikd\sin\theta}}\right)$$

The intensity distribution then follows by applying (4.3).

4.4.2 Maxima and Minima

- Maxima - The maxima are given by the minimum value of the denominator, which is when it is zero. However, (4.20) is still defined as the numberator is also zero (removable singularity). For an integer m:

$$\frac{1}{2}kd\sin\theta = m\pi$$
$$\frac{1}{2}\left(\frac{2\pi}{\lambda}\right)d\sin\theta = m\pi$$
$$a\sin\theta = \lambda m \tag{4.21}$$

- Minima - The minima are given by the minimum values of the numerator. For some integer $p = Nm$:

$$\frac{1}{2}kd\sin\theta = p\pi$$
$$\frac{1}{2}\left(\frac{2\pi}{\lambda}\right)d\sin\theta = p\pi$$
$$a\sin\theta = \lambda p \tag{4.22}$$

i.e there are N-1 minima for each large maxima.

4.4.3 Resolving Power

The resolving power of a grating is a measure of the extent to which we can distinguish between two wavelengths given the order of diffraction that is being worked in. We want to find the condition for the first m^{th} minimum of λ to coincide with the m^{th} order maximum of $\lambda + \delta\lambda$. From (4.21) and (4.22):

$$m(\lambda + \delta\lambda) = d\sin\theta$$
$$p\lambda = Nd\sin\theta$$

But we want $p = Nm + 1$ for the condition to be satisfied.

$$d\sin\theta = \frac{Nm+1}{N}\cdot\lambda$$
$$m(\lambda + \delta\lambda) = \frac{\lambda}{N}(Nm+1)$$
$$\rightarrow R = \frac{\lambda}{\delta\lambda} = Nm \tag{4.23}$$

This is the resolving power of the grating.

4.4.4 Angular Dispersion

The angular dispersion of a diffraction grating is defined as:

$$D = \frac{\partial\theta}{\partial\lambda} \tag{4.24}$$

From (4.21),

$$\cos\theta\cdot\frac{\partial D}{\partial\lambda} = \frac{m}{d}$$
$$\rightarrow D = \frac{m}{d\cos\theta}$$

4.4.5 Grating Spectrometers

Grating spectrometers separate light into it's component wavelenghts. A typical grating spectrometer set-up is shown below in Figure (4.6).

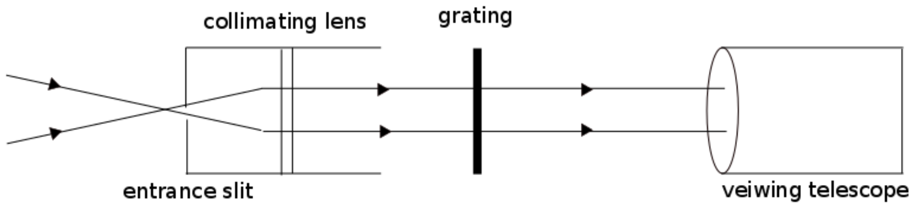

Figure 4.6: A typical Grating Spectrometer

The entrance slit must lie at the focus of the collimating lens in order to make the rays completely parallel or *collimated.* The light is then diffracted off the grating and separated into it's component frequencies. This can then be recorded with a camera or photographic film. There are three main types:

- Spectrograph - Obtains a picture of the separated spectrum
- Spectrometer - Measures the intensity of the individual spectral lines
- Monochromator - This is similar to the spectrometer, except that it deals with a small number of wavelengths.

Factors that limit the highest order we can use include both the intensity of the incoming light (intensity is greatly reduced at higher orders, so the spectrum might not be able to be seen) and the range of wavelengths used. For orders to not overlap, we want the angular displacement of the highest wavelength (λ_2) in a given order to be less than the angular displacement of the lowest wavelength (λ_1) in the next order. From (4.21):

$$m\lambda_2 = d\sin\theta$$
$$(m+1)\lambda_1 = d\sin\theta$$

$$\rightarrow \frac{\lambda_2}{\lambda_1} < 1 + \frac{1}{m} \tag{4.25}$$

Thus for higher orders, the wavelengths need to be closer together to prevent overlapping.