

The Stuff of the Universe

Part I - Particles and Waves

The Classical Picture

1: Atoms, Ideal Gases

1.1 Introduction

How **big** is an atom? All atoms are of similar size, with a radius of about 1 Angstrom (\AA) :

$$1 \text{ Angstrom} = 1 \text{\AA} = 10^{-10} \text{ metres.} \quad (1.1)$$

What is the **mass** of an atom? There are about a hundred **atomic elements** which are ordered in the *Periodic Table of the Elements*, see Appendices F and G of Halliday, Resnick and Walker. All atomic elements have different **atomic masses**, A , depending on the number of protons - the **atomic number**, Z , - and neutrons in the atomic nucleus, respectively. Atomic masses are measured relative to the mass of the carbon-12 **isotope** defined to be exactly 12 **atomic mass units**:

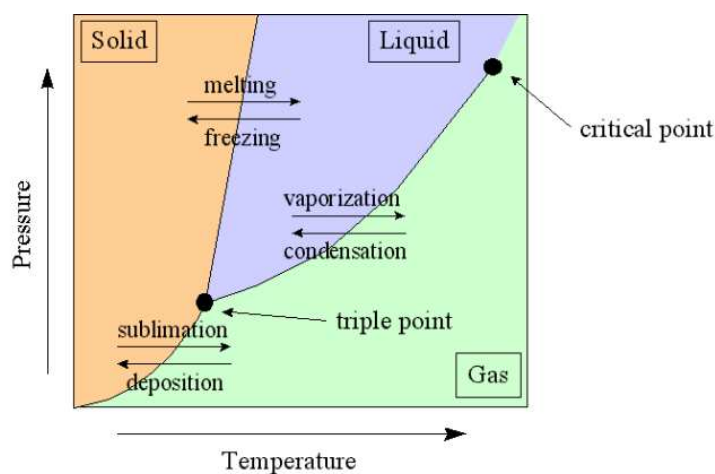
$$1 \text{ atomic mass unit} = 1 u = 1.66 \times 10^{-27} \text{ kg.} \quad (1.2)$$

One **mole** is defined as the number of atoms in a 12 g sample of the carbon-12. A mole of any element has a **molar mass** equal to its atomic mass in grams. The number of atoms in a mole is known as **Avogadro’s number**.

$$N_A = 6.02 \times 10^{23} \text{ mol}^{-1} . \quad (1.3)$$

Many common substances are formed of **molecules**, which consist of a few atoms closely bound together by an attractive force. Examples are water (H_2O), nitrogen (N_2) or oxygen gas (O_2) and carbon dioxide (CO_2). Air is a mixture of molecules (80% N_2 + 20% O_2). A mole of a molecule or a mixture has a mass equal to the sum of its constituent atomic mass in grams. Large molecules can be much bigger in size than atoms.

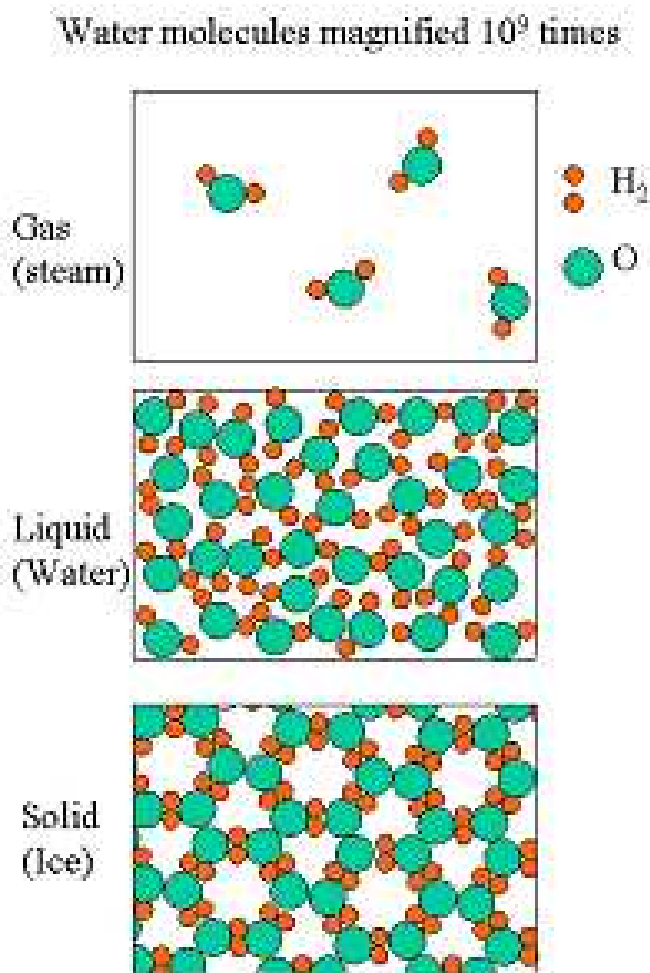
1.2 Phases of Matter



We will be looking more closely at gases in the next few sections, but before we do it is worth reminding ourselves of the phases of matter. At low Temperature and/or high Pressure the molecular separations become small. Transitions occur between solid, liquid and gaseous (vapour) phases of matter along lines of particular values of Pressure and Temperature. The figure shows how the phases of matter depend upon temperature and pressure. Notice the so called *triple point* where in principle all three phases exist at once.

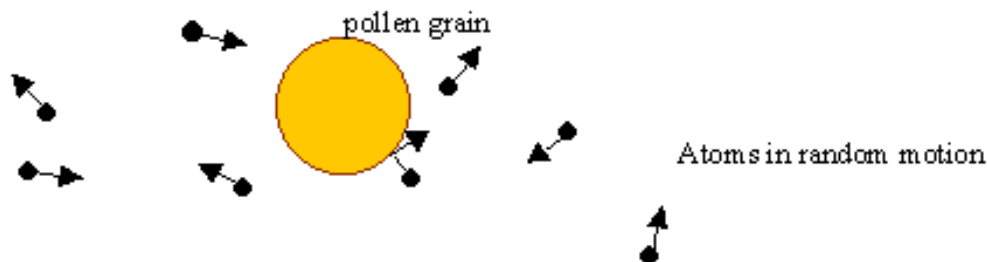
Typical densities for solids and liquids range from 1 g/cm^3 (water, ice) to about 10 g/cm^3 (lead). Most gases at STP (standard temperature and pressure) have a density of about 10^{-3} g/cm^3 . As an example, 1 m^3 of air at STP has a mass of about 1 kg.

The next figure shows a pictorial view of water molecules in the gaseous, liquid and solid phase.



1.3 Molecules are in motion all the time

We know that all molecules of a material are in constant motion. A good example of this is Brownian motion, which was discovered in 1827 by Robert Brown. Under a microscope he saw plant pollen spontaneously moving around in water. He correctly assumed that the pollen was not alive. He subsequently confirmed this motion by observing particles moving around in water that was in bubbles in quartz (inclusions). Those inclusions had been created millions of years ago.



Atoms in liquids are randomly moving around due to their kinetic energy. When they hit a suspended particle they give it a small random momentum. This continues giving the particles the appearance of random motion under the microscope. Do get it clear in your mind that it is the very large number of molecules of the liquid (which you can't see) which are moving around with random velocities which induces the random motion of the particle (which you can see).

It is this random motion of the molecules (= kinetic energy) which gives rise to what we call the *temperature* of the material. We will see shortly that there is a direct relationship between average kinetic energy and temperature.

1.4 Ideal Gases

We are going to look at gases in more detail in the next couple of sections.

A collection of atoms or molecules can be treated as an **ideal gas** if:

- The volume occupied by the gas molecules is negligible compared to the volume of the gas. The molecules can be treated as point masses.
- The time during which molecules collide with each other is negligible. All collisions can be treated as instantaneous.
- All collisions are elastic and obey Newton's laws. There are no long range forces between the molecules.
- The directions of motion of the molecules are random. There is no net motion of the gas in the absence of an external force (no "Maxwell's demon").

These approximations give surprisingly good predictions for the bulk properties of gases (temperature, pressure, ...) under normal conditions.

1.5 The Ideal Gas Law

Experimentally it is observed that:

- The **Volume**, V , of a gas is inversely proportional to its **Pressure**, P , at **constant Temperature**, T (Boyle's Law).
- The **Volume** of a gas is proportional to its **Temperature** at **constant Pressure** (Charles's Law).
- **A mole of an ideal gas at STP occupies a volume of 22.4 litres.** For Standard Temperature and Pressure (STP) $T = 0^\circ\text{C}$ (273 K) and $P = 1$ Atmosphere (1.01 bar) $= 10^5 \text{ N/m}^2$.

Be careful of units, particularly P and V . A consistent set is P in N/m^2 and V in m^3 .

These experimental results can all be combined into a single equation known as the **Ideal Gas Law**:

$$PV = nRT = NkT \quad (1.4)$$

where the temperature T is measured in Kelvins and n is the number of moles of the gas. The quantity R has the same value for all ideal gases, and is known as the **gas constant**. Alternatively the ideal gas law can be written in terms of the number of atoms or molecules in the gas, $N = nN_A$, and the **Boltzmann constant** k :

$$R = 8.31 \text{ J mol}^{-1}\text{K}^{-1} \quad k = 1.38 \times 10^{-23} \text{ J K}^{-1} \quad \text{and} \quad k = \frac{R}{N_A} \quad (1.5)$$

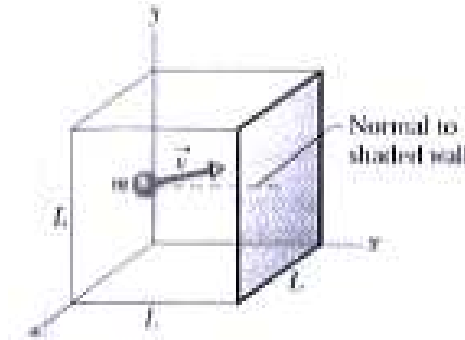
Note that both sides of the ideal gas law have units of Energy (Joules), so the pressure P has units of J/m^3 (Energy/Volume).

2: Kinetic Theory of Gases

In this section we try to understand the behaviour of an ideal gas at the microscopic level of individual atoms or molecules and to derive relations between bulk properties (temperature, pressure, ...) and atomic quantities (speed, kinetic energy) of ideal gases.

2.1 Pressure and Molecular Speed

A gas exerts pressure on the walls of a container due to the collisions of the gas molecules with the walls. We start with a single molecule or atom of mass, m , moving backwards and forwards with a velocity, $\vec{v} = (v_x, v_y, v_z)$, between two walls separated by a distance, L .



Every time the molecule hits the shaded wall it bounces back, reversing the x-component of its momentum. Its momentum $\vec{p} = (p_x, p_y, p_z)$ therefore changes by:

$$\Delta p_x = -2mv_x$$

The time between two collisions with the same wall is how long it takes to traverse the box twice and get back to the shaded wall:

$$\Delta t = \frac{2L}{v_x}$$

Using Newton's second law, the force by the molecule acting on the wall is:

$$F_x = -\frac{\Delta p_x}{\Delta t} = \frac{2mv_x}{2L/v_x} = \frac{mv_x^2}{L}$$

To get the overall gas pressure P on one wall, we need to sum together the contributions due to N molecules in a cubic container of side, L , and take account of the fact that

molecules move with varying speeds by taking an average value for v_x^2 :

$$\begin{aligned} P &= \frac{\text{Force}}{\text{Area}} = \frac{\Sigma F_x}{L^2} \\ &= N \frac{m}{L^3} \langle v_x^2 \rangle \\ &= \frac{Nm}{V} \langle v_x^2 \rangle \end{aligned} \quad (2.1)$$

where we have replaced L^3 with V , the volume of the gas.

In practice we don't measure or consider the average velocity in one specific direction $\langle v_x^2 \rangle$, but only the average speed in any direction $\langle v_{rms}^2 \rangle$ ($v_{rms} = \sqrt{\langle v^2 \rangle}$) is the average or **root-mean-square (rms) speed** of the molecules). These are related in a simple way

$$v_{rms}^2 = \langle v_x^2 \rangle + \langle v_y^2 \rangle + \langle v_z^2 \rangle$$

and since there is no preferred direction

$$v_{rms}^2 = 3 \langle v_x^2 \rangle = 3 \langle v_y^2 \rangle = 3 \langle v_z^2 \rangle$$

hence finally we get

$$P = \frac{Nm}{3V} v_{rms}^2 \quad (2.2)$$

from which we see clearly that the pressure is determined directly by the random kinetic energy of the gas molecules.

If we rearrange this trivially to read:

$$PV = \frac{2N}{3} \times \frac{1}{2} m v_{rms}^2 \quad (2.3)$$

we begin to see how the ideal gas law arises - we have already got PV on the left hand side and we have something proportional to the average kinetic energy on the right hand side. We complete this in the next section.

2.2 Temperature and Kinetic Energy

If we compare the ideal gas law:

$$PV = NkT$$

to the expression we have just derived in 2.3 we see the direct relation between the temperature and the average kinetic energy $\frac{1}{2} m v_{rms}^2$ per molecules of a gas.

The average **Kinetic Energy**, $\langle E_k \rangle$, of a gas molecule is:

$$\langle E_k \rangle = \frac{1}{2} m v_{rms}^2 = \frac{3}{2} kT \quad (2.4)$$

This is a very important result that is quite easy to remember.

We observe that temperature is simply a measure of this average kinetic energy per molecule, and the Boltzmann constant, k , is the constant which relates them (along with a numerical factor of $\frac{3}{2}$). The faster the molecules move the higher is their temperature. The average rms speeds for oxygen molecules at room temperature (300 K) is $v_{\text{rms}} = 483 \text{ m/s}$. This is faster than a bullet! The lowest possible temperature is $T = -273.15^\circ\text{C} = 0 \text{ K}$. What happens at $T = 0 \text{ K}$?

2.3 The Distribution of Molecular Speeds

We have seen that the properties of a gas are mainly determined by v_{rms} . It is also important to know the distribution of the speeds of individual molecules.

This is a statistical problem in which random collisions between molecules lead to changes in their speeds. It was first solved by the Scottish physicist James Clerk Maxwell in 1852. After some time *any initial distribution* of speeds, will reach the same *equilibrium distribution* of speeds. This can be demonstrated experimentally, e.g. by using a set of pucks on an ice rink.

The Maxwell distribution of speeds has a universal shape for all ideal gases at the same Temperature.

The exact shape of Maxwell’s speed distribution, $P(v)$, is the following:

$$P(v) = 4\pi \left(\frac{m}{2\pi kT} \right)^{3/2} v^2 e^{-mv^2/2kT} \quad (2.5)$$

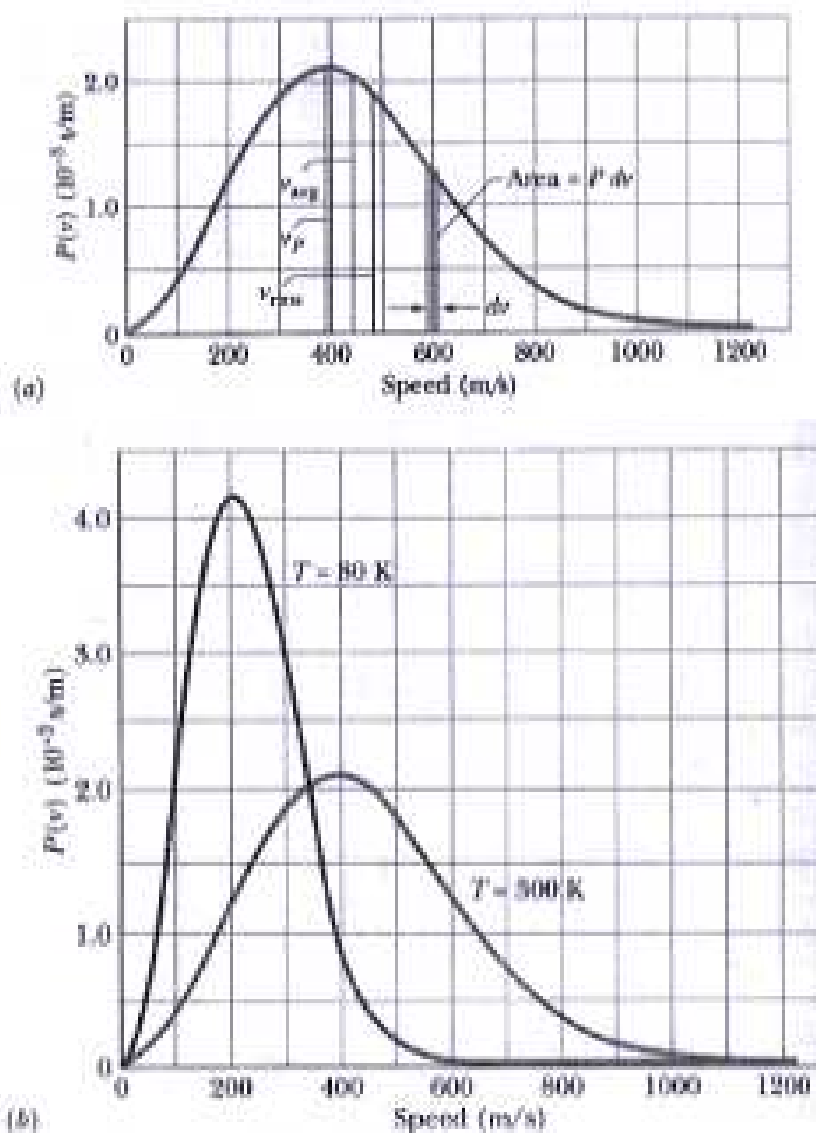
where here P means probability, not pressure. This looks quite complicated, but that is just because of all of the constants. The core of the expression is

$$P(v) \sim v^2 e^{-\frac{E_k}{kT}} \quad (2.6)$$

where E_k = kinetic energy. The second part of this will become very familiar - it is the distribution for the probability for something having an energy E (in this case $E = E_k$) and you will see the expression E/kT arise many many times in your course. The first part (v^2) is simply related to the number of ways a molecule can have a particular speed (it’s all to do with spherical shells at radius v but we will not cover this further in this course).

The Maxwell distribution is shown in the figures below for different temperatures. There is a peak for the most probable speed, which increases with Temperature, and there are tails extending down to zero, and up to very large velocities. Note that all velocities are possible!

The Maxwell distribution of velocities for oxygen molecules: at $T = 300$ K:



There are three different values of speed marked on the graph. In order of increasing speed these are:

- The most probable speed (peak of the distribution)
- The average speed ($\langle v \rangle$)

- The rms speed ($v_{rms} = \sqrt{\langle v^2 \rangle}$)

“A small number of molecules, lying at the extreme right-hand tail of the distribution curve, can have speeds that are several times the average speed. This simple fact makes possible both rain and sunshine.”

(Halliday, Resnick & Walker: 19-7, P.518)

2.4 Example: Nitrogen Molecules

As an example we calculate the rms velocity, v_{rms} of nitrogen molecules at room temperature ($T = 300\text{ K}$), and atmospheric pressure ($p = 1.01\text{ bar}$).

The average Kinetic Energy of a nitrogen molecule is:

$$\langle E_k \rangle = \frac{3}{2}kT = 6.21 \times 10^{-21}\text{ J} = 0.039\text{ eV}$$

Note it is often easier to express atomic or molecular energies in units of electron-Volts (eV):

$$1\text{ eV} = 1.60 \times 10^{-19}\text{ J} \quad (2.7)$$

The molar mass of nitrogen molecules is $M = 28\text{ g}$, and $M = N_A m$

Using $\langle E_k \rangle = \frac{1}{2}mv_{rms}^2$:

$$v_{rms}^2 = \frac{2 \langle E_k \rangle N_A}{M} = \frac{2 \times 6.21 \times 10^{-21}\text{ J} \times 6.02 \times 10^{23}\text{ mol}^{-1}}{28 \times 10^{-3}\text{ kg mol}^{-1}} = 2.67 \times 10^5\text{ (m/s)}^2$$

$$v_{rms} = \sqrt{v_{rms}^2} = 517\text{ m/s}$$

3: Thermodynamics

In this section we first consider different ways in which the state of a gas can be changed, and in particular the $P - V$ diagram. We then discuss the zeroth and first law of thermodynamics which includes the study of thermal equilibrium, internal energy, heat transfer and heat capacities, and the equipartition of energy.

3.1 Changes of state of an Ideal Gas

An ideal gas has an amount of **Internal Energy**, U . As discussed so far this is all tied up in kinetic energy. We will later see that there are other energy “stores” but for now consider it to all be *translational* kinetic energy. Therefore U depends upon T . If we don't change T , then U cannot change.

The internal energy U of an ideal gas can be changed by two types of energy transfer:

- **Work**, W , can be done by an external force to compress the gas or by the gas if it expands. The work done by an ideal gas **expanding** from V_1 to V_2 at constant pressure P is given by:

$$W = \int_{V_1}^{V_2} P dV \quad (3.1)$$

This equation (see lectures) comes simply from Force x Distance. The gas loses this amount of energy. If instead the gas were compressed (the reverse process) then the gas would be given energy.

- Energy in the form of **Heat**, Q , can be transferred into the gas from a hotter body or out of the gas to a cooler body.

Some possible changes to ideal gases:

- An **Isothermal** change keeps the temperature constant. The pressure varies inversely with volume. Work done on the gas is balanced by heat transfer out of the gas.
- An **Isometric** change keeps the volume constant. The pressure is proportional to the temperature. No work is done, $W = \int P dV = 0$, so the change is just due to heat transfer.
- An **Isobaric** change keeps the pressure constant. The volume is proportional to temperature. Work is done to change the volume, and heat is transferred.
- An **Adiabatic** change occurs when there is no heat transfer. This is either because the gas is thermally isolated, or because the change is too rapid for heat transfer to be a factor. The volume, pressure and temperature all change due to work done.

3.2 The $P - V$ diagram

We can visualise different changes using the $P - V$ diagrams (pressure -vs- volume) shown in figures 3.1 to figure 3.3.

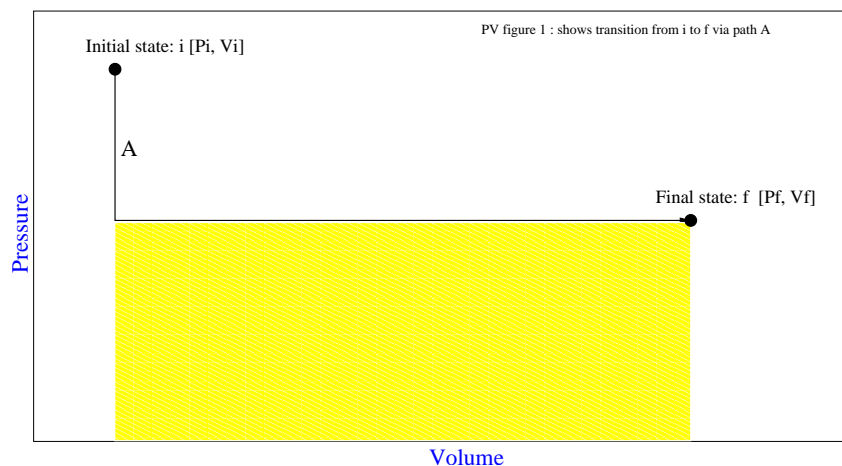


Figure 3.1: P-V diagram showing change of state from $i \rightarrow f$ by path A. The shaded area represents the work done, W , by the gas

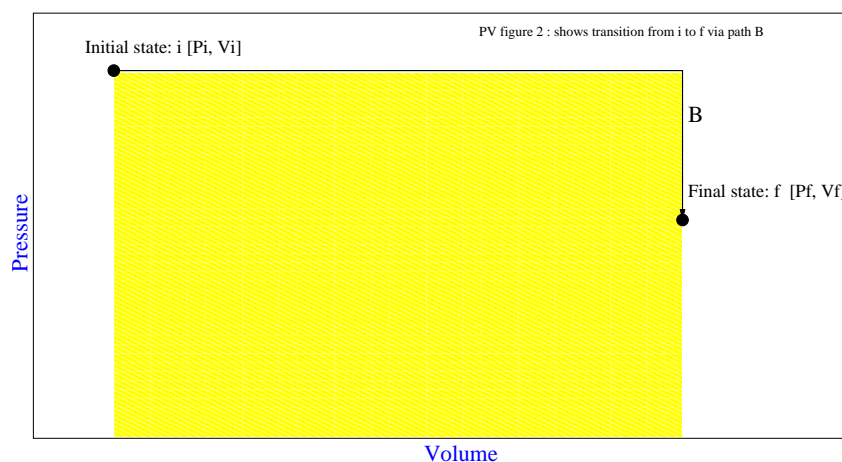


Figure 3.2: P-V diagram showing change of state from $i \rightarrow f$ by path B. The shaded area represents the work done, W , by the gas

These depicts the change of state of a gas from initial state i to final state f by different paths. These are shown as two points in the diagram.

The first and most important thing to realise is that since both initial and final state correspond to particular values of pressure and volume, P_i, V_i and P_f, V_f respectively then the temperatures of both states are also uniquely determined by virtue of the gas

law, i.e.

$$T_i = \frac{P_i V_i}{kN}$$

$$T_f = \frac{P_f V_f}{kN}$$

This is true regardless of how the gas gets from i to f . Since the temperatures are independent of the path taken between the two states, then the internal (kinetic) energy, U , of each the two states is uniquely determined and hence so is the change in internal energy ΔU . This is a very important statement.

The change in internal energy, ΔU , in going from initial state i to final state f is independent of the path taken between the states.

Now we need to visualise what happens when different paths are taken between i and f . We are going to consider how the heat transferred, ΔQ and the work done by the gas ΔW are different in each case, whilst keeping $\Delta U = \Delta Q - \Delta W$ fixed. Students are strongly recommended to use the course text book to get this clear, which includes *gedanken* experiments by which different paths can be realised.

The first path (A) shows first the pressure being reduced (stage 1) and then the volume being increased (stage 2). To achieve stage 1 requires cooling the gas for there is no other way to reduce the pressure at constant volume. Cooling the gas \equiv removing heat, i.e. ΔQ_1 is negative but no work is done so $\Delta W_1 = 0$. To achieve stage 2 the gas must be allowed to expand at constant pressure, which means the gas does work, i.e. ΔW_2 is positive. To cause this to happen the gas must be heated, i.e. ΔQ_2 is also positive. To see why this must be so, think about a container with a piston holding the gas in a volume. To increase the volume means pulling the piston out, but this would naturally reduce the pressure. The only way to maintain constant pressure is therefore to heat the gas at the same time.

Drawing all this together

$$\Delta U_A = \Delta Q_A - \Delta W_A \quad (3.2)$$

where $\Delta Q_A = \Delta Q_1 + \Delta Q_2$ and $\Delta W_A = \Delta W_1 + \Delta W_2$

Before moving on to path B we want to bring out a very important feature of the $P - V$ diagram, namely that the work done by the gas = the area under the $P - V$ curve as shown in the diagram. To see this recall that

$$W = \int_{V_i}^{V_f} P dV \quad (3.3)$$

which is evidently the definition of the area under the $P - V$ curve.

Now consider path (B). Using similar arguments we see that in stage 1 the gas expands at constant pressure (higher than for A). At the same time heat must be put in. In stage 2 heat is removed to reduce the volume and no work is done. It should be clear that the work done by the gas $\Delta W_B > \Delta W_A$ because it expanded at a higher pressure. This can also be seen as the area under the $P - V$ curve is greater for path (B) than path (A). Therefore since

$$\Delta U_B = \Delta Q_B - \Delta W_B \quad (3.4)$$

and by definition

$$\Delta U_B = \Delta U_A \quad (3.5)$$

then it is clear that $\Delta Q_B > \Delta Q_A$.

To summarise

- ΔU is independent of the path between a particular i and f
- Both ΔQ and ΔW depend upon the path
- ΔW always equals the area under the curve.

It is left to the student to analyse path (C) and conclude that the work done is in between that of the other paths, i.e. $\Delta W_B > \Delta W_C > \Delta W_A$.

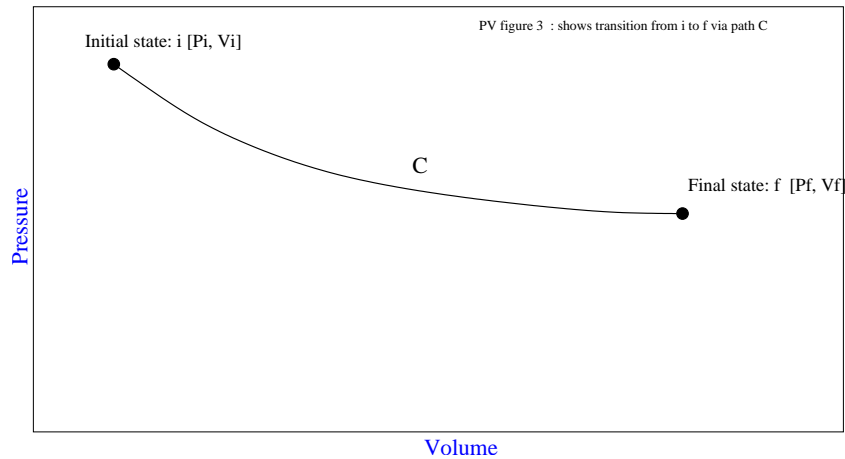


Figure 3.3: P-V diagram showing change of state from $i \rightarrow f$ by path C.

3.3 The P-V diagram for a change at constant temperature

We can easily draw the line on the $P - V$ diagram corresponding to all possible initial and final states which have the same temperature. Re-writing the gas law as

$$P = (NkT) \frac{1}{V} \quad (3.6)$$

we see that (at constant temperature) everything in the brackets is constant and hence

$$P \propto \frac{1}{V} \quad (3.7)$$

This gives the lines shown in figure 3.4. Each line is an iso-therm, corresponding to one fixed temperature.

Students should prove (see tutorial question) that the work done by the gas for such a path is

$$\Delta W = NkT \ln \left(\frac{V_f}{V_i} \right) \quad (3.8)$$

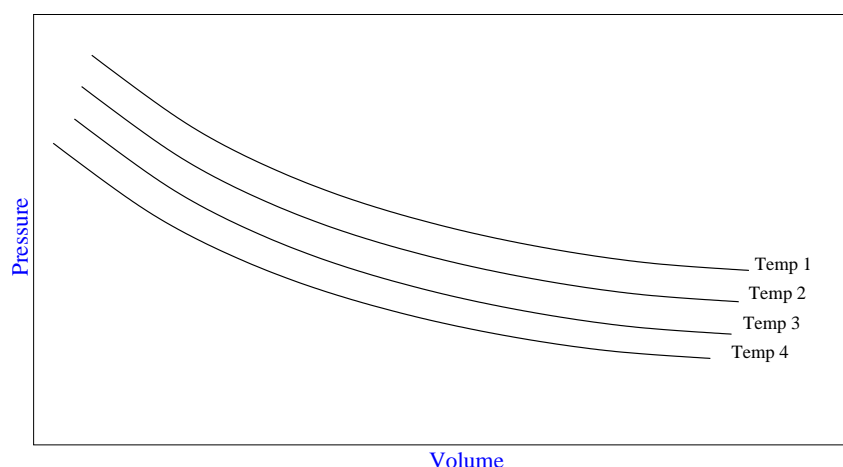


Figure 3.4: P-V diagram showing isotherms. These are changes of state of a gas at **constant temperature**. In this case the lines all obey a $P \propto 1/V$ relationship.

3.4 Zeroth and First Laws of Thermodynamics

We are ready to formally state some laws of thermodynamics which should appear “obvious” given all we have covered concerning the kinetic theory of gases, the internal energy and its relation to temperature.

If two bodies of matter, A and B, have the **same Temperature** they are said to be in thermal equilibrium.

Zeroth Law of Thermodynamics:

A third body of matter, C, which is in thermal equilibrium with body A, must also be in thermal equilibrium with body B.

All this is really saying is that the temperatures will equalise. This should be no surprise as temperature simply means kinetic energy. If the two gases, A and B, are not in thermal equilibrium, then when they are in contact with each other, the collisions between the molecules of the two gases will transfer energy until thermal equilibrium is reached. **When two gases are in thermal equilibrium their molecules have the same Kinetic Energy.**

The energy transfers due to Heat, Q , and Work, $W = \int PdV$, change the **internal energy** U of the molecules.

First Law of Thermodynamics:

The internal energy U of a system tends to increase if energy is added as heat Q , and tends to decrease if energy is lost through work W done by the system.

$$\Delta U = \Delta Q - \Delta W \quad (3.9)$$

3.5 Specific Heat Capacities

The **heat capacity** of a body of matter is defined by the amount of energy, ΔQ , which raises the temperature of the gas by ΔT :

$$C = \frac{\Delta Q}{\Delta T} \quad (\text{units } \text{J/K}) \quad (3.10)$$

C is therefore the amount of heat transfer required to change the Temperature of the body by 1 degree.

The **specific heat capacity** is defined per unit mass or per mole:

$$\begin{aligned} c &= \frac{\Delta Q}{m\Delta T} & (\text{units } \text{J/kg K}) \\ c &= \frac{\Delta Q}{n\Delta T} & (\text{units } \text{J/mole K}) \end{aligned} \quad (3.11)$$

The latter is sometimes quoted as **molar heat capacity**. The specific heat capacity is a property of a particular material. Table 19-3 on P.435 of HRW lists c for some solids and liquids. Note the similarity of the molar specific heats for the solids.

3.6 Latent Heat

A **change of phase** of a body of matter between gas and liquid or between liquid and solid requires the transfer of heat known as **Latent Heat**. This transfer occurs at a particular Pressure and Temperature.

The Latent Heat is defined per unit mass:

$$L = \frac{\Delta Q}{m} \quad (\text{units } \text{J/kg}) \quad (3.12)$$

There is no change in the Temperature of the body during the transfer of Latent Heat. It is a property of a material making a particular phase transition.

The Latent Heat of Fusion of melting ice to water or solidifying water to ice is 333 kJ/kg. The Latent Heat of Vaporisation of water to steam or vice versa is 2256 kJ/kg.

3.7 Specific Heat Capacities of Gases at constant Volume

If a gas is heated at **constant volume**, then no work is done on it or by it. In this case all the heat transferred goes into increasing the temperature. Therefore

$$\Delta Q = \Delta U \quad (3.13)$$

Recall that the average kinetic energy of the gas molecules is

$$\langle K.E \rangle = \frac{3}{2}NkT = \frac{3}{2}nRT$$

(this is strictly true only for monatomic gases - see later for more complicated cases). Therefore if the temperature is raised by ΔT it means the K.E has been raised by ΔQ :

$$\Delta Q = \Delta U = \frac{3}{2}nR\Delta T$$

The *molar specific heat capacity* is defined by ΔQ per mole divided by the temperature rise:

$$C_V = \frac{\Delta Q}{n\Delta T} = \frac{\Delta U}{n\Delta T} = \frac{\frac{3}{2}nR\Delta T}{n\Delta T} \quad (3.14)$$

after cancellation we get the important result:

$$C_V = \frac{3}{2}R = 12.5 \text{ J/mole K} \quad (3.15)$$

so we see that the specific heat capacity of gases is directly related to the gas constant.

Table 20-3 on P.471 of HRW lists values for the molar specific heats of gases. For **monatomic gases** (He, Ar,...) the measurements agree with the predicted value of $\frac{3}{2}R$. For **diatomic gases** (H_2 , N_2 , O_2), the value increases to

$$C_V = \frac{5}{2}R = 20.8 \text{ J/mole K.} \quad (3.16)$$

What is the explanation for this increase?

3.8 Specific Heat Capacities of Gases at constant Pressure

At **constant pressure**, the volume changes and work is done by the gas. Therefore

$$\Delta Q = \Delta U + \Delta W \quad (3.17)$$

where because the expansion is at constant pressure

$$\Delta Q = \Delta U + P\Delta V \quad (3.18)$$

The specific heat capacity at constant pressure is thus:

$$C_P = \frac{\Delta Q}{n\Delta T} = \frac{\Delta U}{n\Delta T} + \frac{P\Delta V}{n\Delta T}$$

The expression for ΔU is the same as in the constant volume case. Using the Ideal Gas Law at constant pressure $P\Delta V = nR\Delta T$, we obtain:

$$C_P = C_V + R \quad (3.19)$$

3.9 Equipartition of Energy

Equipartition of Energy:

Every kind of molecule has a certain number of *degrees of freedom*, which are independent ways of storing energy. Each degree of freedom has associated with it - on average - an energy of $\frac{1}{2}kT$ per molecule.

(Halliday, Resnick & Walker: 19.9, P.523)

The total internal energy stored in an ideal gas of N molecules or atoms is:

$$U = \frac{f}{2}NkT = \frac{f}{2}nRT \quad (3.20)$$

where f is the number of degrees of freedom.

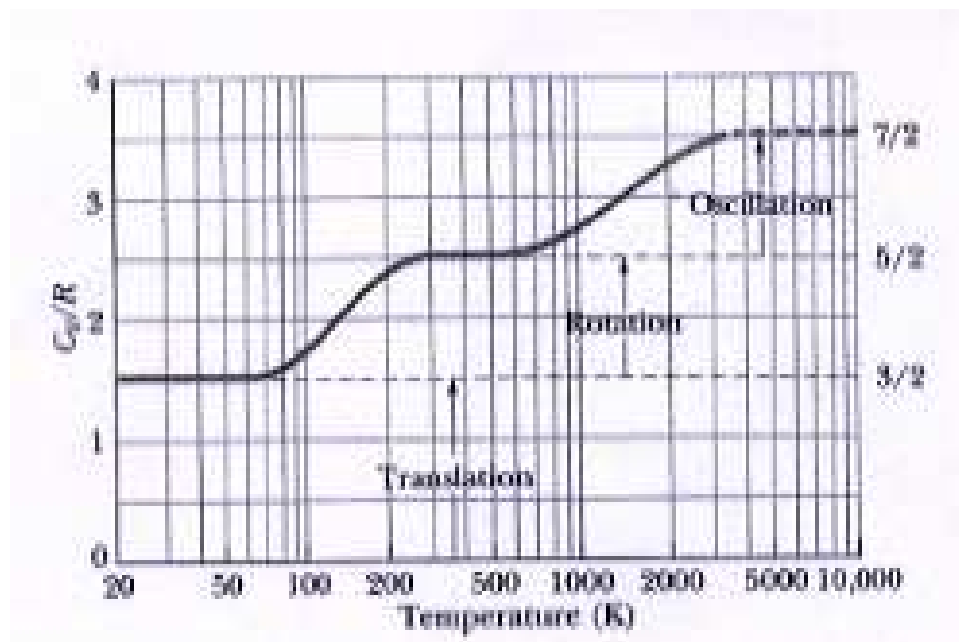
We have already met the simple case where $f = 3$. These correspond to the **three degrees of freedom** which are associated the translational motion in three dimensions. This is the **translational Kinetic Energy** of the molecules. All gases have this form of energy.

Monatomic gases have only three available degrees of freedom. This leads to $C_V = \frac{3}{2}R$.

Diatomic gases have two atoms per molecule. There are two additional degrees of freedom associated with **rotations** about the centre of gravity. In this case $f = 5$. This leads to $C_V = \frac{5}{2}R$

More complex molecules have additional degrees of freedom associated with other rotational modes. There can also be **vibrational energy** (oscillations) stored between the atoms of a molecule. This will increase C_V even further.

You might think that a particular gas either has $C_V = \frac{3}{2}R$ or $C_V = \frac{5}{2}R$ or something else, but that whatever it is it is fixed. This is wrong! The figure below shows that C_V for this diatomic gas example changes with temperature. This must mean that **the number of degrees of freedom depends on Temperature**. Some modes “freeze out” below particular temperature. The explanation for this is all tied up with quantum mechanics.



C_V/R versus temperature for hydrogen gas.

The Classical Wave Picture

4: Properties of Waves

In this section we introduce the concept of classical waves. Their properties are described and illustrated by examples of mechanical waves.

4.1 Introduction

Waves arise in many branches of physics such as *mechanics*, *thermodynamics*, *electromagnetism* and *quantum physics*. Here we group them in three main types:

- Mechanical waves,
- Electromagnetic waves,
- Matter waves.

Mechanical and electromagnetic waves are the classical waves which are the subject of this part of the course, “*The Classical Wave Picture*”. The concept of matter or particle waves will be introduced in the following part: “*The Quantum World*”.

4.2 Mechanical Waves

A **mechanical wave** is a **local disturbance** of a **material (medium)**. The disturbance **varies periodically in space and time**, which causes **variations of the energy density** stored in the material. These periodic energy variations can travel over **large distances**.

The disturbance of a medium is sustained by local interactions governed by Newton’s laws. Unlike classical particles, a wave is not a pointlike object. A wave is a broad distribution of energy filling a particular region of space.

There are many familiar examples of mechanical waves:

- Waves travelling along a rope, string or spring
- Water waves
- Sound waves.
- Seismic waves associated with earthquakes.

As an example, a mechanical **travelling wave** can be generated by tapping periodically a string under tension. This causes a transverse displacement of the string material varying in space and time. We can produce a variety of **waveforms** according to how we disturb the string and what happens at the end of the string.

4.3 Mathematical Description of Waves

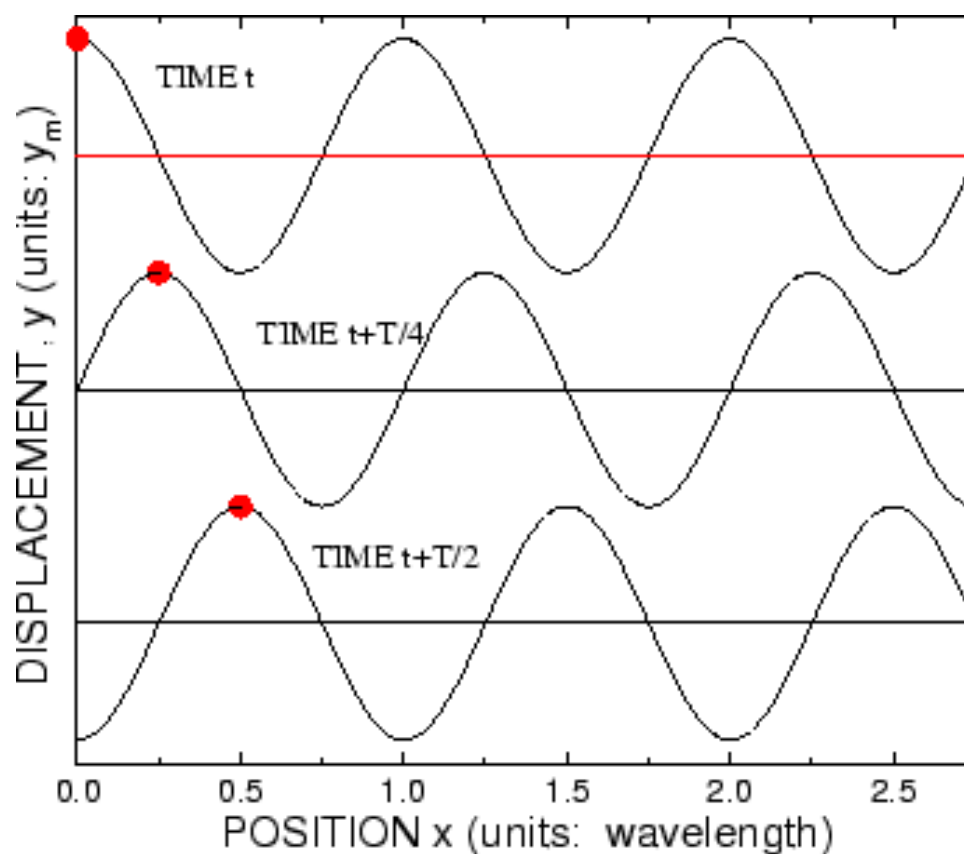
For mechanical waves the disturbance is often a **displacement**, y , of an element of a medium at position x and time t . For a concrete example think of the vertical displacement of a string, y , at a distance x along the string.

For a **sinusoidal travelling wave** the displacement $y(x, t)$ can be written as:

$$y(x, t) = A \sin(kx \pm \omega t + \phi) \quad (4.1)$$

where the -ve (+ve) sign describes a wave travelling to the right (left), respectively.

It is worth plotting this on graph paper for a few different times. This is depicted in the figure below for $(kx - \omega t)$ which shows the wave for three successive times. The wavefront clearly moves to the right.



The terms appearing in the expression are:

- The **amplitude** of the oscillation, A , is the maximum displacement of the medium from its equilibrium position.
- A snapshot of a travelling wave shows the **displacement in space** at fixed time t :

$$y(x) = A \sin(kx + \pi/2) \quad (4.2)$$

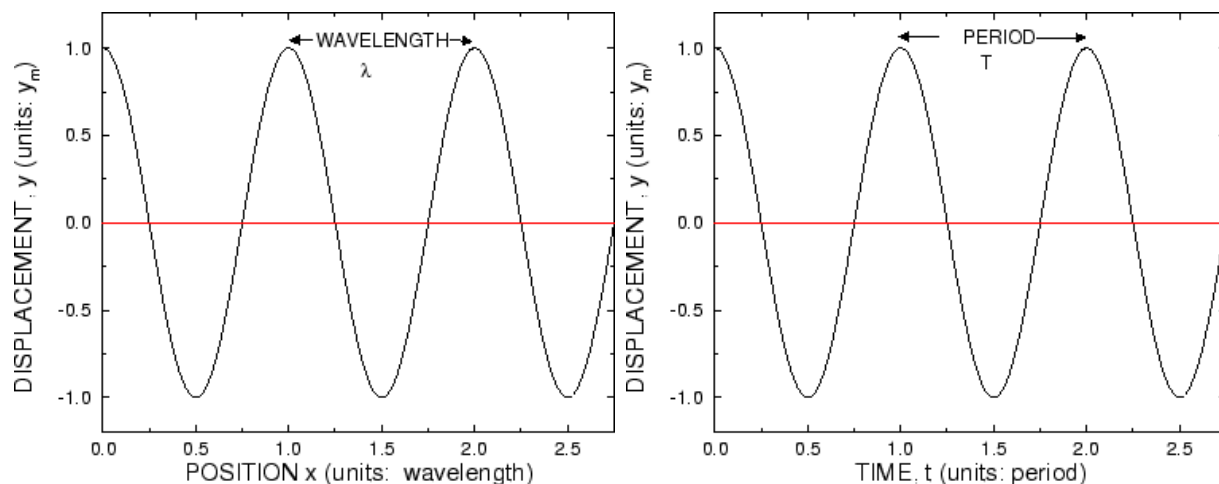


Figure 4.5: Snapshots at fixed time (left) and fixed position (right)

This is shown in the left hand figure.

- The **wavelength**, λ , is the distance between successive maxima of the energy fluctuations. Note that λ is not the distance between a minimum and a maximum of the waveform.
- The **wavenumber**, k , is a useful quantity to describe waves with sine or cosine functions. It is indirectly proportional to the wavelength.

$$k = \frac{2\pi}{\lambda} \quad (4.3)$$

- The **phase shift**, ϕ , gives the distance of the first minimum from the origin $x = t = 0$.
- The **displacement in time** shows the time evolution of a wave element at a fixed position x . For a sinusoidal wave this is a **Simple Harmonic Motion**:

$$y(t) = A \sin(-\omega t + \pi/2) \quad (4.4)$$

This is shown in the right hand figure.

- The **period**, T , of oscillation is the time between successive maxima of the oscillation.
- The **frequency**, f , is the rate of oscillation of a wave element and is the inverse of the period, $f = 1/T$. The frequency is measured in units of $1 \text{ s}^{-1} = 1 \text{ Hertz (Hz)}$.
- The **angular frequency** ω , which is defined as $\omega = 2\pi f$, is also convenient to use in describing oscillations and waves. A very useful relation is:

$$\omega = \frac{2\pi}{T} \quad (4.5)$$

4.4 Alternative form

It may be useful to recognise an alternative form for the travelling wave expression. The sine function for wave traveling to the right was written earlier as:

$$y(x, t) = A \sin(kx - \omega t + \phi) \quad (4.6)$$

This can be written (using $\sin(p + q) = \sin(p) \cos(q) + \cos(p) \sin(q)$) as:

$$\begin{aligned} y(x, t) &= A[\cos \phi \sin(kx - \omega t) + \sin \phi \cos(kx - \omega t)] \\ &= C \sin(kx - \omega t) + D \cos(kx - \omega t) \end{aligned} \quad (4.7)$$

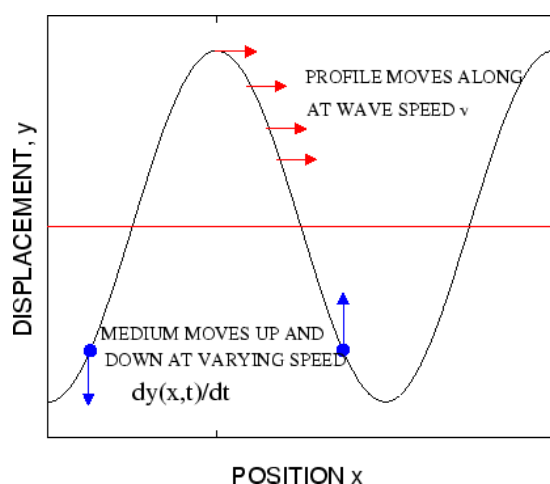
where $C = A \cos \phi$ and $D = A \sin \phi$. This represents the superposition of a sine wave and a cosine wave with different amplitudes depending upon the phase shift.

The forms are equivalent.

4.5 Wave Speed

The **wave speed** v describes the rate at which the wave front advances in space. It is important to understand that this does not mean that any piece of the medium itself is travelling anywhere at this speed.

The displacement of a **wave node**, x_n , must be zero at all times. This yields the relation:



$$kx_n - \omega t = n\pi$$

Wave nodes advance at the wave speed v . We define $v = \frac{dx_n}{dt}$.

The **wave speed** v of a travelling wave is related to the angular frequency, ω , and the wavenumber, k :

$$v = \frac{\omega}{k} = f\lambda \quad (4.8)$$

The **wave speed** v is determined by the **properties of the medium**. It can be written in the generic form:

$$v = \sqrt{\frac{\text{stiffness}}{\text{inertia}}} \quad (4.9)$$

where *stiffness* is a restoring effect of the medium, like a force or a tension, and *inertia* is a mass or moment of inertia resisting changes to the motion of the medium, respectively.

For waves on a string the wave speed is determined by the the string tension, T , and the mass per unit length, μ :

$$v = \sqrt{\frac{T}{\mu}} \quad (4.10)$$

4.6 Medium speed

We also define a **medium speed** which describes the rate at which matter elements of the medium move as the waveform travels. An example would be the speed at which a piece of a string (say you paint a short section yellow) moves up and down. The medium speed is by definition the time derivative of the displacement. For a wave travelling to the right:

$$\begin{aligned} v_m &= \frac{dy(x, t)}{dt} \\ &= \frac{d}{dt} A \sin(kx - \omega t) \\ &= -\omega A \cos(kx - \omega t) \end{aligned} \quad (4.11)$$

The medium speed depends on the frequency ω and the amplitude A of the wave. When the displacement s maximum the medium speed is zero and when the displacement is zero the medium speed is maximum (because when $\sin = 1$ then $\cos = 0$ and vice versa).

4.7 Doppler Effect

If you hear a source of sound, such as a siren moving first towards you and then away from you, you will notice that the pitch decreases when the siren passes by you. These are examples of the **Doppler effect**.

The frequency of an emitted sound increases (decreases) if either the source of a sound wave or the sound detector are moving towards (away from) each other. The emitted frequency f and detected frequency f' are related by:

$$f' = f \frac{v \pm v_d}{v \pm v_s} \quad (4.12)$$

where v is the speed of sound in air, v_d is the detector's speed relative to air, and v_s is the source's speed relative to air. Getting the signs right can seem confusing, and of course it depends upon which direction you chose to be positive v . However it is very simple to see what the correct signs are. If the source and detector are moving towards each other then the frequency must increase.

To understand this result first set v_d to zero and consider a source, S, moving at speed v_s towards the detector, D. It is emitting a sine wave of frequency f , period T and wavelength λ . The velocity of the wave in the medium is v .

For the purposes of this derivation we shall only consider the positions of the peaks. The diagram below shows a “peak” being emitted at $t = 0$. The next peak is emitted at $t = T$. If the source had been stationary, then the distance between the peaks would have simply been $vT = \lambda$. However the source has been moving and gained a distance $v_s T$ so that the two successive peaks are nearer together and hence the detector will observe a shorter wavelength. This shorter wavelength is easy to calculate, denoting it as λ' .

$$\lambda' = vT - v_s T = T(v - v_s) \quad (4.13)$$

Thus the wave is *compressed* such that to the detector it has a shorter wavelength. In terms of the original wavelength use $vT = \lambda$ to get:

$$\lambda' = \lambda \frac{(v - v_s)}{v} \quad (4.14)$$

From this we can determine the observed frequency f' using $f'\lambda' = v$ (we are using the fact that even though compressed, the velocity of the wave in the medium cannot change and for all waves $v = f\lambda$).

$$f' = \frac{v}{\lambda'} = \frac{v}{T(v - v_s)} = f \frac{v}{(v - v_s)} \quad (4.15)$$

This shows how the denominator of 4.12 arises.

Now consider the case of $v_s = 0$ but the detector is moving towards the source with a speed v_d . Visualise successive peaks arriving at the detector. After one peak arrives the detector moves towards the next peak such that they meet nearer to each other than if the detector were stationary. This is shown in the diagram below. If the detector had not been moving it would have seen the next peak at a time T later, but because it is moving it sees the next peak earlier at time T' . Since $T' < T$ the frequency will be higher. You can see that the the distances add up like this:

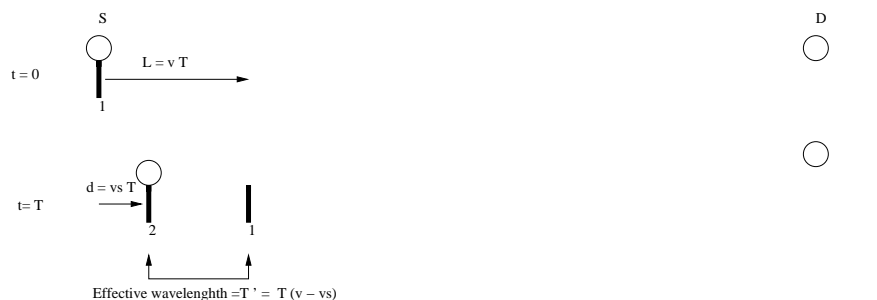
$$\lambda = T'v + T'v_d = T'(v + v_d) \quad (4.16)$$

We can rearrange this using $f\lambda = v$ and $f' = \frac{1}{T'}$ to give

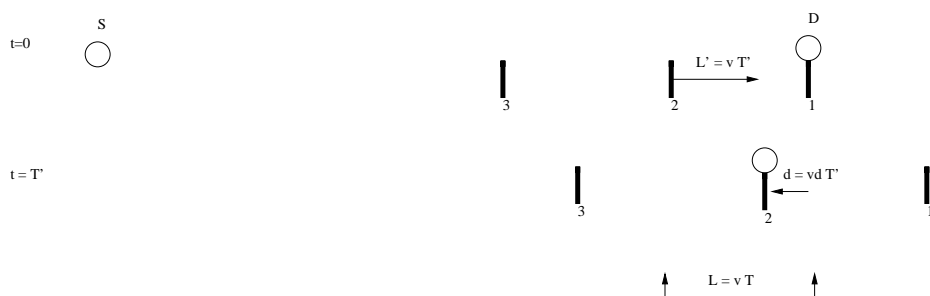
$$f' = f \frac{(v + v_d)}{v} \quad (4.17)$$

This shows how the numerator of 4.12 arises.

If you follow the same arguments but allowing both v_s and v_d to be non zero you will obtain the full result of 4.12 where we identify the $+$ on the top and $-$ on the bottom as both are moving towards each other and so both must cause $f' > f$.



The sketch above shows that after the first peak is emitted (1) it moves distance $L = vT$ (L =wavelength) after one period T . The source meanwhile moves $d = T v_s$ before emitting the next peak and hence the shortened wavelength is $T(v - v_s)$.



This sketch shows successive peaks 1,2,3 passing the detector. After 1 passes the detector at $t=0$, it meets the next one at $T' < T$ because it has moved towards it. The lengths add up so that $vT = vT' + v_d T'$

4.8 The wave equation

Wave equations are very important in physics. A wave equation is a **differential equation** which represents the dynamics of the physical situation. The solutions of the wave equation are the allowed solutions of the problem. Very often one can derive the wave equation from the physical system.

We are not going to derive the wave equation for our waves in these lectures, but the interested reader can find this in the course text book for a string. If you are keen you are recommended to do the derivation in your own time. The beauty is that there are only about three wave equations you ever come across in physics. Sure, they have different constants and symbols which make them look different, but at their heart they are mostly of the form:

$$\begin{aligned}
 \frac{dy(x,t)}{dt} &\propto \frac{d^2y(x,t)}{dx^2} \\
 \frac{d^2y(x,t)}{dt^2} &\propto \frac{d^2y(x,t)}{dx^2} \\
 \frac{dy(x,t)}{dt} &\propto \frac{dy(x,t)}{dx}
 \end{aligned} \tag{4.18}$$

The relevant wave equation in our situation is the middle one. All of our waves are solu-

tions of the wave equation:

$$\frac{d^2 y(x, t)}{dt^2} = v^2 \frac{d^2 y(x, t)}{dx^2}$$

or using $v = \omega/k$

$$\frac{1}{\omega^2} \frac{d^2 y(x, t)}{dt^2} = \frac{1}{k^2} \frac{d^2 y(x, t)}{dx^2}$$

To demonstrate that our expression for a travelling wave is a solution substitute it in to each side. First the LHS

$$\frac{d^2 y(x, t)}{dt^2} = \frac{d^2 A \sin(kx - \omega t)}{dt^2} = -\omega^2 A \sin(kx - \omega t)$$

Next the RHS

$$v^2 \frac{d^2 y(x, t)}{dx^2} = v^2 \frac{d^2 A \sin(kx - \omega t)}{dx^2} = -v^2 k^2 A \sin(kx - \omega t)$$

and since $\omega^2 = v^2 k^2$ then the LHS=RHS.

In fact it is trivial to show that any function of the form:

$$f(x, t) = f(kx \pm \omega t)$$

is a solution of this wave equation. The solutions do not have to be sinusoidal!

4.9 Examples of Travelling Waves

Examples of mechanical waves are where the disturbance $y(x, t)$ can be the displacement of a string, the density variation caused by sound waves, the height of water waves or seismic waves.

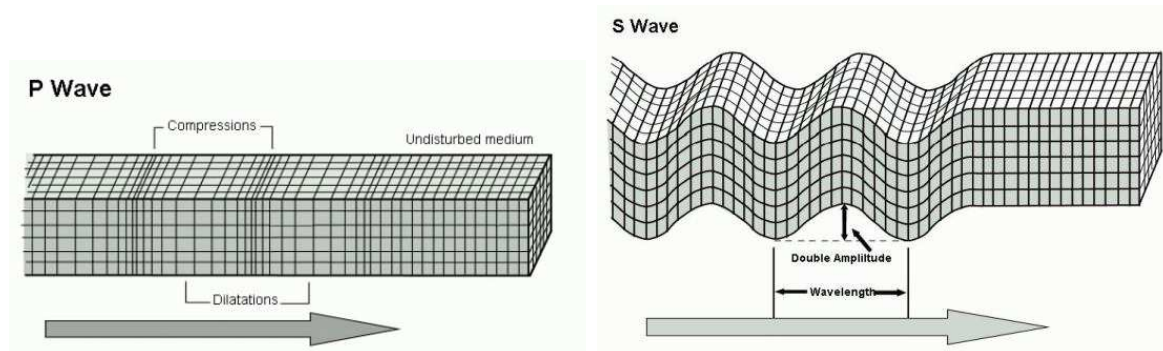
Longitudinal waves have a displacement amplitude parallel to the direction the wave is travelling in. Sound waves are longitudinal.

Transverse waves have a displacement in the $y - z$ plane perpendicular to the direction x the waveform is travelling in.

The amplitude of a transverse wave can be **polarised** in any direction in the $y - z$ plane. A plane-polarised wave has a fixed $y - z$ direction. A circularly-polarised wave has a direction that rotates in the $y - z$ plane. Note that longitudinal waves are always

unpolarised.

In an earthquake a sudden rupture of the earth’s structure releases a vast quantity of energy which travels across the earth’s surface in the form of **seismic waves**. Longitudinal and transverse seismic waves are called P-waves and S-waves (P stands for primary, S for secondary), respectively. P-waves travel faster and arrive first at the seismic monitor stations.



4.10 Worked Examples

[1: from HRW Ch6:P1]

A sinusoidal wave moves along a string. The time for a particular point to move from maximum displacement to zero is $0.17s$. What are (a) the period and (b) the frequency? The wavelength is $1.40m$ what is the wave speed?

Imagine this wave travelling on a stretched string. Imagine a point on the string marked with paint. The first part of the question tells you that this mark moves up and down (transverse to wave motion) such that it takes $0.17s$ to go from a min to a max. This is $1/4$ of a cycle therefore it takes $4 \times 0.17s$ to do a complete cycle. Hence:

$$T = 0.68s \quad (4.19)$$

$$f = \frac{1}{T} = 1.47s^{-1} \quad (4.20)$$

The next part tells you the wavelength from which we can write

$$\lambda = 1.4m \quad (4.21)$$

The wave speed is given by the relation in the notes

$$v = f\lambda = 1.47 \times 1.4 = 2.06ms^{-1} \quad (4.22)$$

[2: from HRW Ch6:P7]

A transverse sinusoidal wave is moving along a string in the positive x direction with a speed of 80ms^{-1} . At $T = 0$ the string particle at $x = 0$ has a transverse displacement of 4.0cm from its equilibrium position and is not moving. The maximum transverse speed of the string particle is 16ms^{-1}

1. *What is the frequency of the wave?*
2. *What is the wavelength of the wave?*

If the wave is described by

$$y(x, t) = y_m \sin(kx \pm \omega t + \phi) \quad (4.23)$$

what are

1. y_m
2. k
3. ω
4. ϕ
5. *The correct choice of sign?*

This is a good example as it makes us go through almost everything connected with a travelling wave.

The first part involves the medium speed. Adopting the equation suggested, then it tells us that

$$y(x, t)_{\max} = 0.04\text{m} \quad (4.24)$$

$$\frac{dy(x, t)}{dt}_{\max} = 16\text{ms}^{-1} \quad (4.25)$$

Look at either the course notes or just the equation given to see that

$$y(x, t)_{\max} = y_m \quad (4.26)$$

when the sin term = 1, and

$$\frac{dy(x, t)}{dt} = \pm \omega y_m \cos(kx \pm \omega t + \phi) \quad (4.27)$$

is max when the cos = 1 giving

$$\frac{dy(x, t)}{dt}_{\max} = \pm \omega y_m \quad (4.28)$$

Ignoring the sign for now we can find ω and f

$$\omega = 2\pi f = \frac{16}{0.04} = 400 \text{ } r s^{-1} \quad (4.29)$$

$$f = \mathbf{64 \text{ } s^{-1}} \quad (4.30)$$

The next part asks for the wavelength. The only other piece of information you are given so far is the wave speed (call this v_w) which is just what we need for:

$$v_w = f\lambda \quad (4.31)$$

$$\lambda = \frac{v_w}{f} = \frac{\mathbf{80}}{\mathbf{64}} = \mathbf{1.25 \text{ m}} \quad (4.32)$$

Turning to the second part of the question

- We have already identified y_m as the maximum transverse displacement (0.04m)
- k is the wave number given by $2\pi/\lambda$
- We have already found ω which is the angular frequency
- ϕ is the phase offset by which the wave is shifted to the **negative** x direction at $t = 0$
- The correct choice of sign is negative as we know $(kx - \omega t)$ means the wave travels in the positive x direction - if in any doubt you must plot this out for yourselves.

5: Standing Waves, Sound Waves

In this section we introduce the superposition principle for waves and discuss the properties of standing waves and of sound waves.

5.1 Superposition of Waves

Superposition Principle:

If two or more travelling waves pass simultaneously through the same region in space, the **resultant wave** disturbance or displacement $y(x, t)$ is the **sum** of the disturbances $y_i(x, t)$ ($i = 1, 2, \dots$) of the **individual waves**.

$$y(x, t) = y_1(x, t) + y_2(x, t) + \dots \quad (5.1)$$

When **adding** together the **displacements** of the waves, it is necessary to take account of their **phases** as well as of their **amplitudes**.

(This is strictly true for **linear** media, but we will not concern ourselves with non-linear media in this course)

The superposition principle is **valid for all waveforms**, the disturbances $y_i(x, t)$ need not necessarily be sinusoidal travelling waves.

As an example, we add together two **identical waves** with a **phase difference**, $\Delta\phi$:

$$\begin{aligned} y_1(x, t) &= A \sin(kx - \omega t) & y_2(x, t) &= A \sin(kx - \omega t + \Delta\phi) \\ y(x, t) &= y_1(x, t) + y_2(x, t) = 2A \cos \frac{\Delta\phi}{2} \sin(kx - \omega t + \frac{\Delta\phi}{2}) \end{aligned} \quad (5.2)$$

where we have used the trigonometric identity

$$\sin \alpha + \sin \beta = 2 \sin \frac{\alpha + \beta}{2} \cos \frac{\alpha - \beta}{2}$$

The result is a sine wave of the same frequency but (i) diminished in amplitude and (ii) with a different phase. Note that if the phase difference were π then the two waves would exactly cancel out.

5.2 Standing Waves

Having established the principle of superposition we will now use it to understand several phenomena. The first is how two travelling waves can make a standing wave.

If two identical sinusoidal waves travel in opposite directions we obtain a **standing wave**.

$$y_1(x, t) = A \sin(kx - \omega t) \quad y_2(x, t) = A \sin(kx + \omega t)$$

$$y(x, t) = y_1(x, t) + y_2(x, t) = 2A \sin kx \cos \omega t \quad (5.3)$$

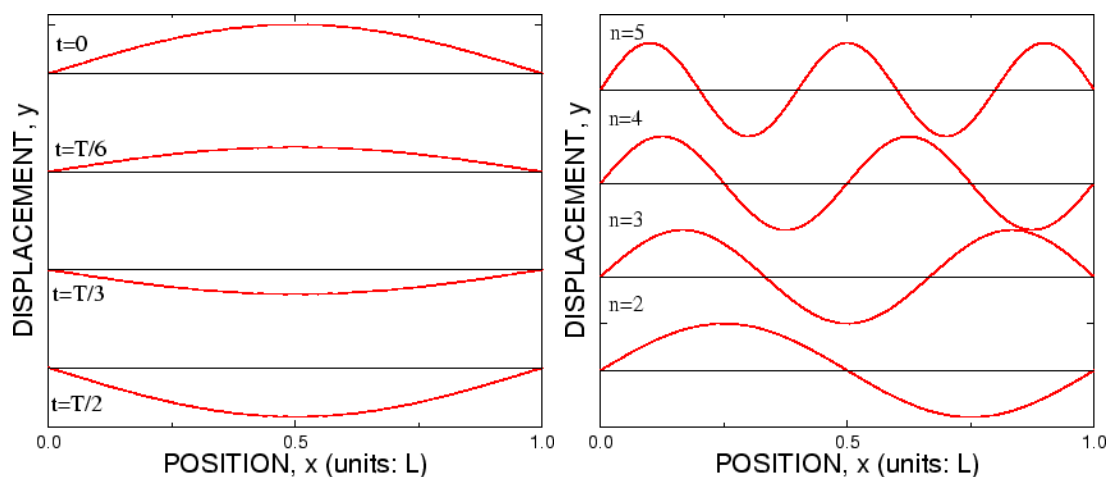
(this follows from the same trig identity used on the previous page)

In a standing wave there are oscillations as a function of both x and t , but they are decoupled. The nodes of the wave stand still and the total wave does not propagate in any direction. It is vitally important to get this clear and if you are in any doubt you should get some graph paper:

- Plot $\sin(kx - \omega t)$ as a function of x for two different times t_1 and t_2 . You will clearly see the wavefront move to the right.
- Plot $\sin(kx) \cos(\omega t)$ as a function of x for two different times t_1 and t_2 . Nothing moves to left or right, the vibration just goes up and down and the nodes (zeros) always stay zero.

The lecture demonstrations will illustrate the difference.

A primary example of standing waves is provided by a string under tension and fixed at both ends. This example has many applications from the physics of musical instruments to the fundamentals of quantum mechanics. Consider a string with length L , fixed



at its ends. A sinusoidal wave initiated at one end will be reflected at the other end, giving a second sinusoidal wave travelling in the opposite direction. The resulting standing wave must fulfil the boundary conditions at all times.

$$y(0, t) = y(L, t) = 0 \quad \implies \quad \sin kL = 0 \quad \text{or} \quad kL = n\pi \quad \text{where} \quad n = 1, 2, 3, \dots$$

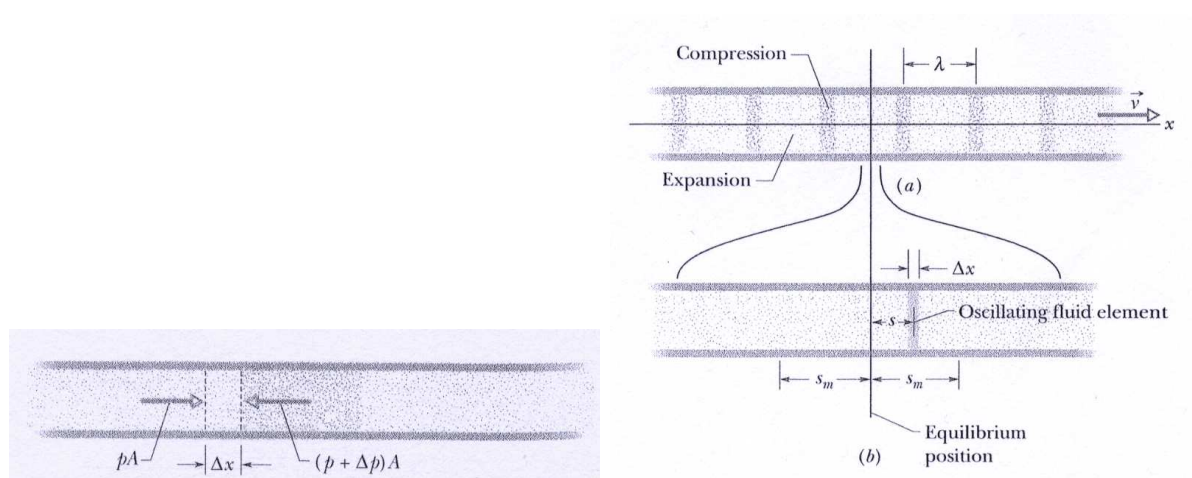
We obtain that a string of length L , fixed at each end, can only support standing waves of wavelength λ satisfying

$$L = \frac{n\lambda}{2} \quad (5.4)$$

The different solutions are called **modes** and $n = 1$ is the **fundamental mode** which has the longest allowed wavelength $\lambda = 2L$.

5.3 Example of a standing wave in a gas - leads to a good demonstration

In a gas a sound wave occurs through changes in **density** and **pressure**.



A region of gas is displaced causing changes in the density on either side of the region. This leads to changes in pressure described by the ideal gas law. The changes in pressure then create an accelerating force and cause the region of gas to move back. Sound waves are longitudinal waves since the density variations travel along the direction of the wave.

5.4 Waves of Different Frequencies

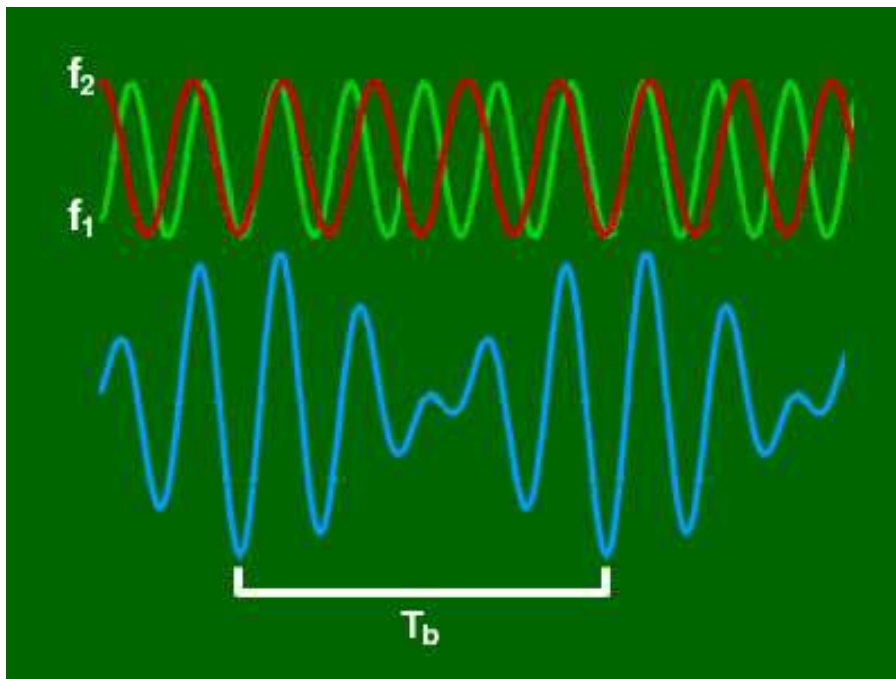
Two waves with the same amplitude, wavelength and phase, but different frequencies can be superimposed:

$$\begin{aligned} A_1(x, t) &= A_0 \sin(kx - \omega_1 t) & A_2(x, t) &= A_0 \sin(kx - \omega_2 t) \\ A(x, t) &= 2A_0 \cos(\omega' t) \sin(kx - \omega t) \end{aligned} \quad (5.5)$$

The resultant **beat wave** oscillates with both the average angular frequency $\omega = \frac{1}{2}(\omega_1 + \omega_2)$ and half the angular frequency difference $\omega' = \frac{1}{2}(\omega_1 - \omega_2)$.

If the two frequencies are very similar then $\omega \gg \omega'$. We hear a sound of frequency $f = \omega/2\pi$ with an amplitude that is modulated by the beat frequency $f_{\text{beat}} = 2\omega'/2\pi$.

The phenomenon of beating is depicted in the figure, which shows the combined waves, with their overall amplitude modulated by the beat *envelope*. The period of the beat envelope is shown as T_b .



5.5 Standing wave worked example

From HRW-7 Chapter 16 Problem 50

A standing wave pattern on a string is described by

$$y(x, t) = 0.04 \sin(5\pi x) \cos(40\pi t) \quad (5.6)$$

where x and y are in metres and t is in seconds.

- *For $x \geq 0$ what is the location of the first three nodes?*
- *What is the period of the oscillation?*
- *What are the speed and amplitude of the two travelling waves that interfere to produce this standing wave?*
- *For $t \geq 0$ what are the first three times that all points on the string have zero transverse velocity?*

This question makes you reverse the standing wave expression to obtain the two travelling waves.

It probably helps you to compare it to the standing wave generic expression before doing anything else:

$$y(x, t) = 2A \sin(kx) \cos(\omega t) \quad (5.7)$$

where the factor of 2 is explicit in the amplitude so that A becomes the amplitude of each travelling wave. You might also want to write down the equivalence:

$$A \sin(kx - \omega t) + A \sin(kx + \omega t) = 2A \sin(kx) \cos(\omega t) \quad (5.8)$$

From this we can identify:

$$2A = 0.04 \rightarrow A = 0.02m \quad (5.9)$$

$$k = \frac{2\pi}{\lambda} = 5\pi \rightarrow \lambda = 0.4m \quad (5.10)$$

$$\omega = 2\pi f = 40\pi \rightarrow f = 20s^{-1} \quad (5.11)$$

The first part of the question asks for the first three nodes. This means the positions x on the string which are always at zero displacement. Inspecting the $\sin(5\pi x)$ term this happens when $5\pi x = 0, \pi, 2\pi, 3\pi$ i.e. $x = 0.0, 0.2, 0.4, 0.6$ or in other words $x = 0, \frac{1}{2}\lambda, \lambda, \frac{3}{2}\lambda$ which is exactly what you expect : a standing wave has nodes at $x = n\frac{1}{2}\lambda$

The next part of the question asks for the period of oscillation. This is trivially $T = 1/f = 0.05s$.

The next part asks about the travelling waves. The amplitude of each is $A = 0.02m$ because of the way we chose to write the equivalences above. To obtain the wave velocity use

$$v_w = f\lambda = \frac{\omega}{k} = 8ms^{-1} \quad (5.12)$$

Finally you are asked for the first three times that all points on the string have zero transverse velocity. This bit might fool you if you don't read it carefully. It is not asking for the first three times when the displacements are zero (i.e. the string is completely flat). These would be when $\cos(\omega t) = 0$. Instead this asks for when the transverse velocity is zero. This is when the displacement is maximum, i.e the string is at maximum deformation. This was done in the lectures and you see it by taking dy/dt and finding when it is zero. The answer is when $\cos(\omega t) = \pm 1$. Convince yourself of this. This is when $\omega t = 0, \pi, 2\pi$...etc. or $t = 0, 0.025, 0.05s$...etc. This is what you expect, it is every half period.

6: Electromagnetic Waves

6.1 Light

Light requires no medium for its travel. Light travels with a **constant velocity**, c , in a **vacuum**:

$$c = 2.99792458 \times 10^8 \text{ m/s} \quad (6.1)$$

This fundamental constant of Nature is the **fastest allowed speed** in Einstein’s theory of **Special Relativity**. **The speed of light c is the same for observers in all inertial frames of reference.** A consequence of Special Relativity is that when observing objects which are travelling at large velocities comparable to the speed of light clocks tick slower and distances become contracted.

In a famous set of experiments Newton showed that **white light** from the sun could be broken up into a **spectrum of colours** with the help of a glass prism. Different colours of light correspond to different wavelengths or frequencies. The **visible light spectrum** extends from $\lambda = 400 \text{ nm}$ (violet) to $\lambda = 700 \text{ nm}$ (red), with the eye being most sensitive at 560 nm (yellow), the peak of the Sun’s emission spectrum. The sun also shines at invisible wavelengths longer than $\lambda = 700 \text{ nm}$ (**infrared**) and shorter than $\lambda = 400 \text{ nm}$ (**ultraviolet**), respectively.

In the 19th century James Clerk Maxwell developed a **theory of electromagnetism** which combined the known properties of **electric** and **magnetic fields**.

Light is an Electromagnetic Wave (Maxwell):

Light propagates as a traveling wave of oscillating electric and magnetic fields with a velocity c through a vacuum.

The amplitudes of the electric and magnetic fields are related to the intensity or power of the light wave. The energy stored in the electric and magnetic fields fluctuates in a similar way to the energy stored in matter by sound waves. Light waves are **transverse** and can be **polarised**. The frequency of visible light waves ($f = c/\lambda$) is about $6 \times 10^{14} \text{ Hz}$.

6.2 The Electromagnetic Spectrum

The **electromagnetic spectrum** has since been extended by Hertz, Röntgen and others to include **radio waves**, **X-rays** and **gamma rays**. Hertz discovered radio waves and first demonstrated that they move through space at the speed of light. Electromagnetic waves are the basis of many modern devices. The figure shows the electromagnetic spectrum.

Radio waves - AM and FM radios, TV, Telescopes

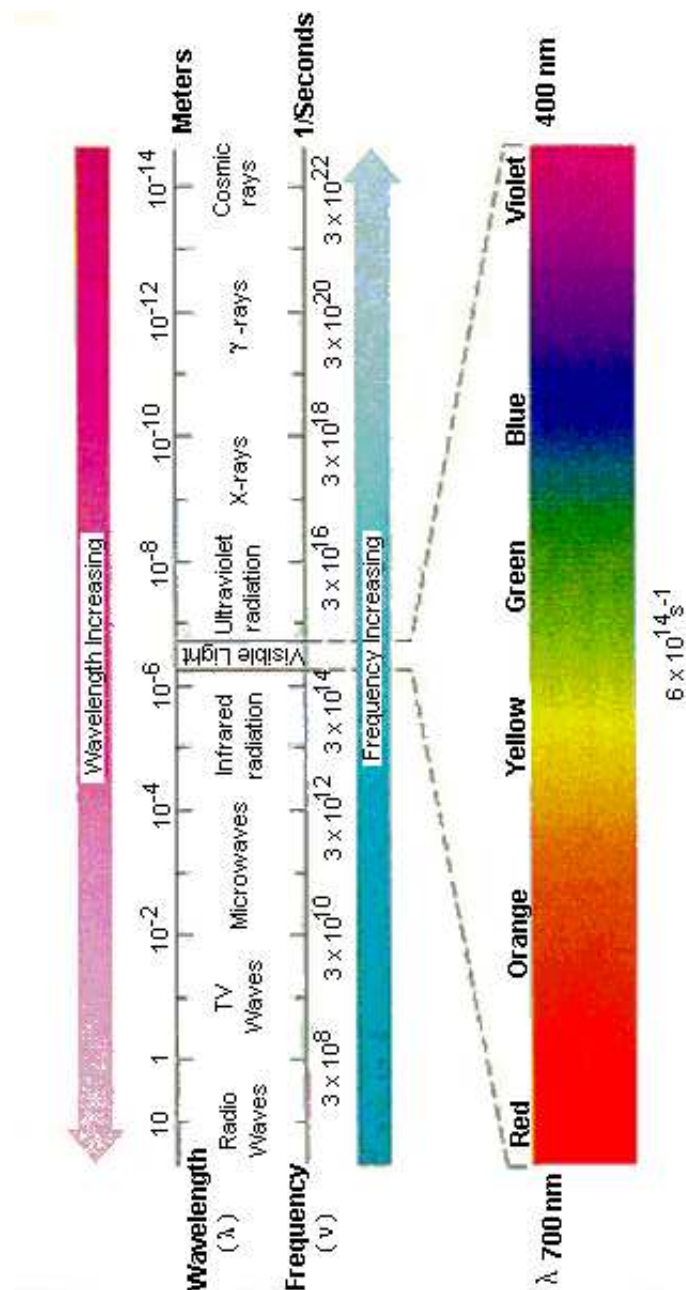
Micro waves - radar, microwave ovens, mobile phones

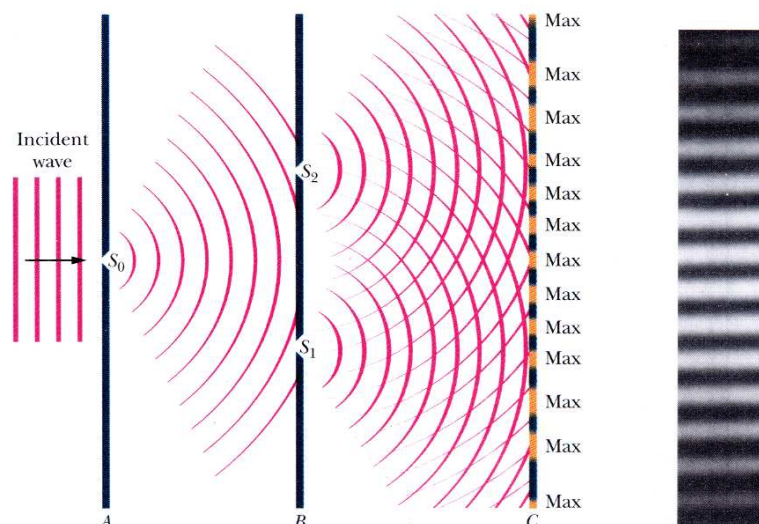
Infrared - night vision

Visible - cameras, lasers (also in infrared and UV)

Ultraviolet - food preservation

X-rays and gamma rays - medical imaging, x-ray tubes, MRI, CAT and PET scan





6.3 Light is a Wave

Young's double slit experiment is the best known example of an interference pattern due to a path length difference. In 1801, T. Young demonstrated this for the first time and proved that light is a wave.

A parallel beam of monochromatic light is incident on a mask with one slit perpendicular to the beam direction. The emerging light then spreads (Huygens' principle) and illuminates a second mask which is opaque except for two slits separated by a distance d . The slits are assumed to be narrow compared to d . The intensity profile observed at a large distance L (which is D on the diagram) on the screen is the result of **interference** between waves passing through the two slits. This is shown in the figure. We will analyse this in the next section.

6.4 Reflection and Refraction

At a **boundary** between two **different media** light can be both reflected and refracted.

The figure shows the simpler case when an incident ray travels in a medium with a lower refractive index, n_1 , and the refracted ray is in a medium with a higher index, n_2 .

The **reflected ray** is of course reflected with an angle to the normal equal to the angle of the incident ray, $\theta_1 = \theta'_1$.

The **refracted ray** is transmitted into the new medium, with an angle which is related to the incident ray by **Snell's Law**:

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \quad (6.2)$$

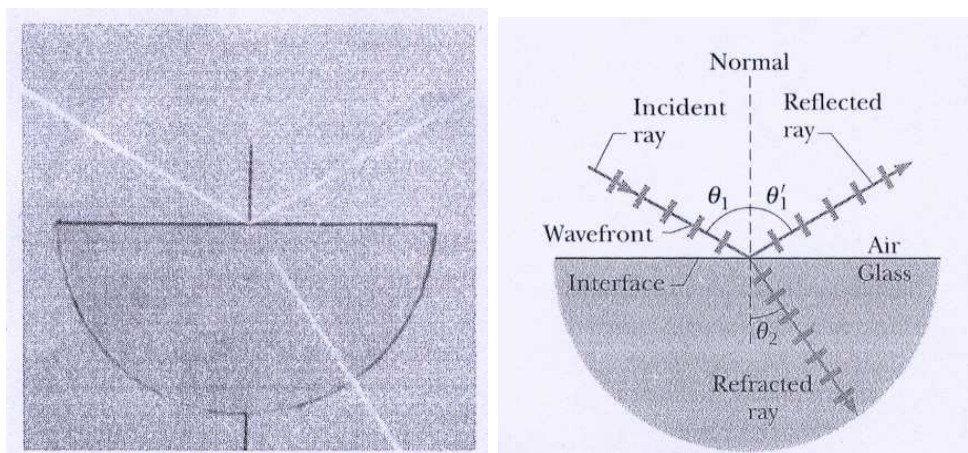


Figure 6.6: A ray of light in air (refractive index n_1) is incident upon a glass block (refractive index n_2). The ray is reflected at angle θ_1' and refracted at angle θ_2

The next figure (next page) shows the more complex case when the incident ray is in the higher refractive index medium, i.e. $n_1 > n_2$. The same Snell's law applies.

In this case, however, there is not always a refracted ray. To see why this is, think what happens when the incident angle θ_1 becomes large enough such that the refracted angle $\theta_2 = 90^\circ$. The refracted ray then runs exactly along the boundary. We call the value of θ_1 at which this happens the **critical** angle θ^{crit} .

For all values $\theta_1 > \theta^{crit}$ there is no refracted ray - it becomes unphysical. To see this mathematically look at Snell's law rearranged:

$$\sin \theta_2 = \frac{n_1}{n_2} \sin \theta_1 \quad (6.3)$$

we see that whenever $\sin \theta_1 > \frac{n_2}{n_1}$ the RHS becomes greater than 1 which means $\sin \theta_1$ is unphysical. Thus we see the critical angle is given by

$$\sin \theta^{crit} = \frac{n_2}{n_1} \quad (6.4)$$

The propagation of light in a long thin **optical fibre** is based on Total Internal Reflection. Thus our entire communications infrastructure depends upon this phenomenon.

6.5 Dispersion

Another way of looking at the refractive index of a particular material is that it tells us the speed of the light within that material. Specifically, if the refractive index of the medium is n then the wave speed inside the material is given by

$$v = \frac{c}{n} \quad (6.5)$$

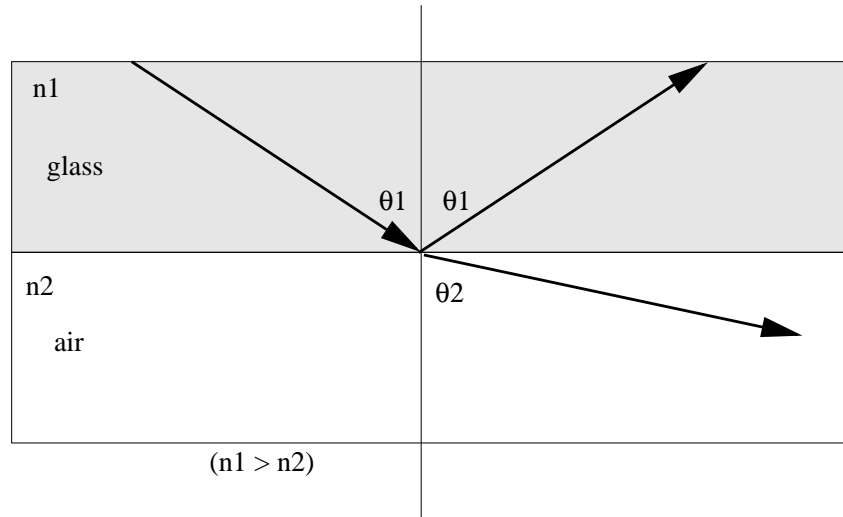


Figure 6.7: A ray of light in air (refractive index n_1) is incident at angle θ_1 upon a glass block (refractive index n_2). The ray is reflected at angle θ_1 and refracted at angle θ_2

For most materials, the refractive index has a different value for different frequencies of light. Such materials are known as *dispersive* materials: in this case, the simple relation $\omega = kv$ is no longer sufficient, and instead the angular frequency will depend on the wavenumber according to the *dispersion relation*, $\omega(k)$.

For example, in ordinary glass the refractive index is usually larger for blue light than red light. This means that blue light travels through glass more slowly than red light (as shown by equation 6.5). It is this which gives rise to the phenomenon of white light separating into a colourful continuous spectrum when it is refracted by a glass prism.

6.5.1 Phase and group velocity

One consequence of dispersion is that we have to be a little more careful about what we have been referring to as the ‘wave speed’. Strictly, so far this has been the *phase velocity*, which gives the rate of travel of the phase of the wave. If we define the phase velocity as v_p then

$$v_p = \frac{\omega}{k} \quad (6.6)$$

Waves of different frequency will travel with different phase velocities inside the medium. When these waves superpose, their combined amplitudes will be modulated by an overlying envelope, in a similar way as we saw with beating waves in section 5.4.

The envelope shape itself propagates with its own speed, known as the *group velocity*, v_g . The group velocity is the speed at which the wave transmits energy or information. For example, a light pulse in an optical fibre will transmit information at the group velocity.

The group velocity is obtained by differentiating the dispersion relation:

$$v_g = \frac{\partial \omega(k)}{\partial k} \quad (6.7)$$

In this course we will only consider in detail non-dispersive waves. These have a very simple dispersion relation:

$$\omega(k) = kv_p$$

and we see that

$$v_g = \frac{\partial(kv_p)}{\partial k} = v_p$$

i.e. for non-dispersive waves the group velocity and phase velocity are identical, and we restore the simple case we had before when we considered just the straightforward wave speed v , which will be sufficient for our purposes here.

7: Interference and Diffraction

7.1 Analysis: Young’s Double Slit Experiment

Young’s double slit experiment was touched upon in the last set of notes. The figure is repeated below. The diagram we use to analyse it is shown in the following figure. Make sure you fully understand this diagram. It is very important to the whole week.

The slit spacing is d and the distance to the screen is D . If $D \gg d$, the path length difference Δr between the two waves is:

$$\Delta r \approx r_1 - r_2 = d \sin \theta \quad (7.1)$$

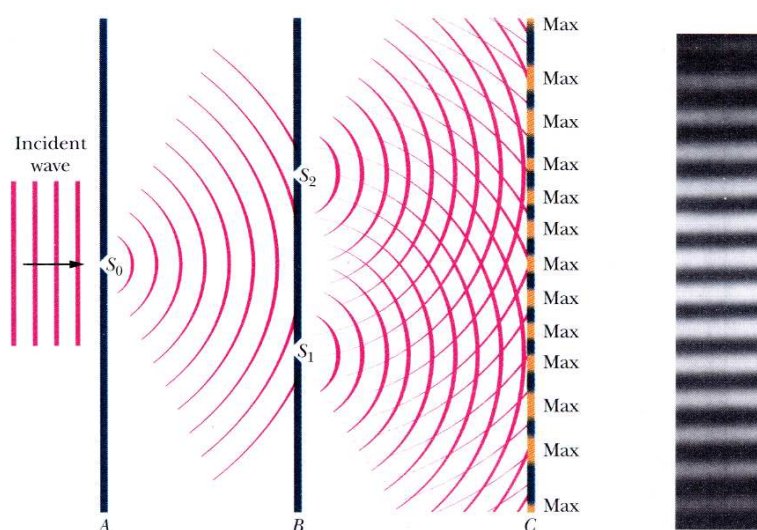
where r is the distance travelled by the ray and θ is the angle between the beam direction and the line from the halfway point between the slits to the point where we observe the pattern on the screen (at a distance y from the point opposite the slits). For small angles of θ (i.e. $y \ll D$) the following approximation is valid:

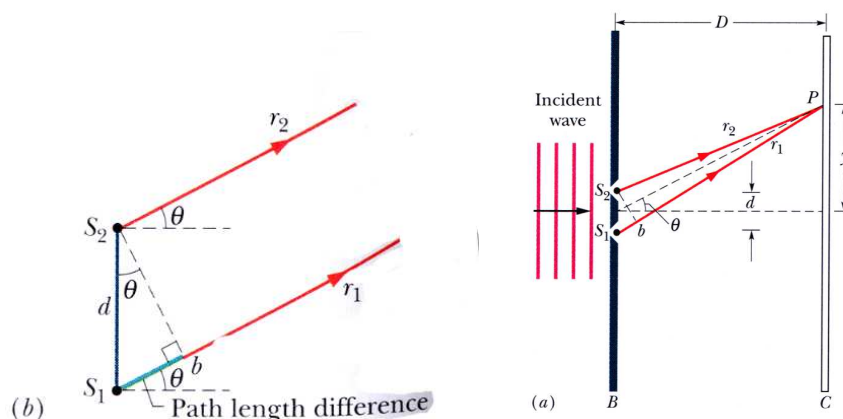
$$\sin \theta \approx \tan \theta \approx \theta \approx \frac{y}{r} \quad (7.2)$$

Bright fringes – constructive interference – are expected for phase differences $\Delta\phi = m2\pi$ ($m = 0, 1, 2, \dots$) which corresponds to path differences Δr equal to an integer multiple, m , of the wavelength λ of the light:

$$m\lambda = d \sin \theta \quad (7.3)$$

Please learn this well. It should become a mantra.





Dark fringes – destructive interference – are expected for half integer multiples of the wavelength λ :

$$(m + \frac{1}{2})\lambda = d \sin \theta \quad (7.4)$$

The pattern produced is that shown on the right hand side in the first figure. We will look at lots of examples of this as demonstrations during the lecture using water and lasers.

7.2 Diffraction Grating

A **diffraction grating** consists of a large number of equally spaced slits. They are usually made by drawing lines on a sheet of glass.

The arguments used for Young's double slits can be extended to show that a large number of slits, N , will produce **bright maxima** at the points satisfying:

$$m\lambda = d \sin \theta \quad (7.5)$$

If the number of slits is very large the intensity pattern is very simple. The maxima are rather narrow and well separated by large dark regions.

The reason for this is straightforward to see. For all points on the screen **other than the maxima** there will always be a large number of waves arriving each a little out of phase with the next. There will therefore always be pairs which cancel. The resulting aggregate is such that they all cancel. The only exception is for $m\lambda = d \sin \theta$ where all waves arrive exactly in phase. Hence this is the only condition for which they do not completely cancel out. This is depicted in the figure.

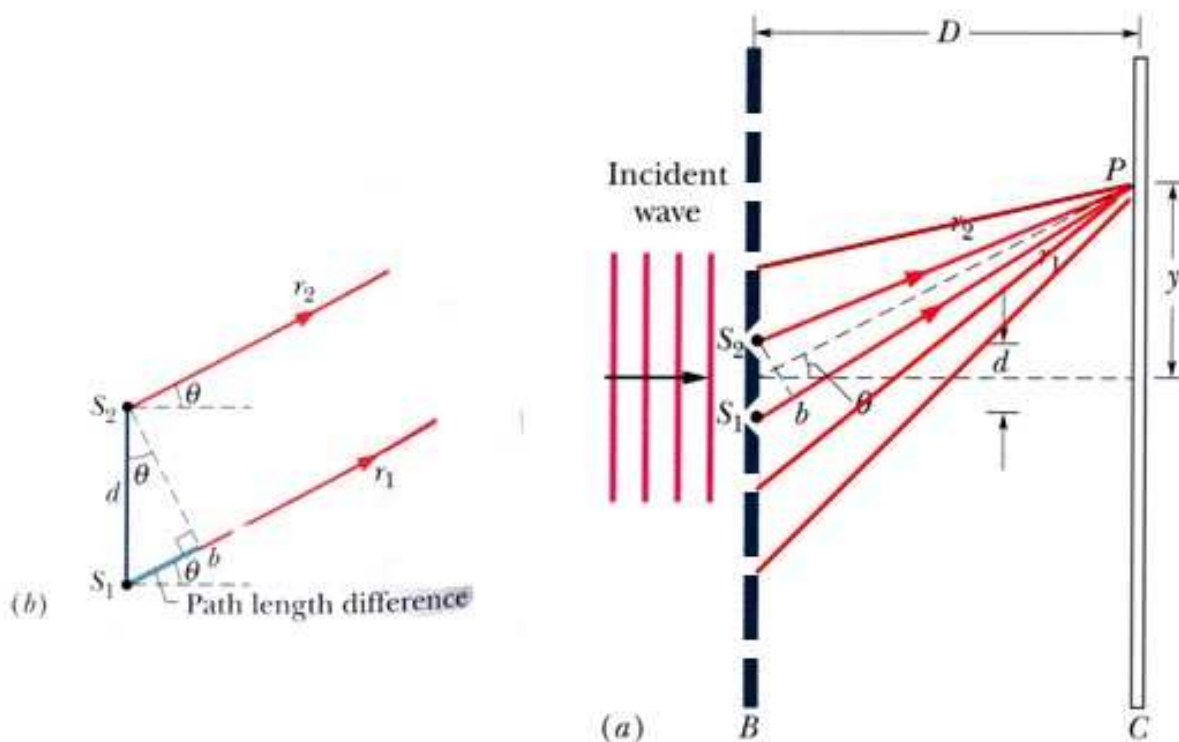


Figure 7.8: Transmission Diffraction Grating

The integer m defining the maxima is known as the **order number**. The straight through image is the zeroth order which is seen for all wavelengths. The first maxima on each side are the **first order lines**. The position of these lines depends on d and λ . There is a highest order number given by the limit $\sin \theta \leq 1$.

7.3 Reflection from diffraction gratings

The diffraction grating considered in the last section was implicitly used to transmit the incident wave through it.

Similar diffraction occurs if the incident wave is reflected off of the grating. This is depicted in the figure. Consideration of all of the path differences yields the result:

$$m\lambda = d(\cos \alpha - \cos \beta) \quad (7.6)$$

7.4 The Spectrometer

The spectrometer makes use of a diffraction grating to diffract the different coloured emission lines from a light source to different angles.

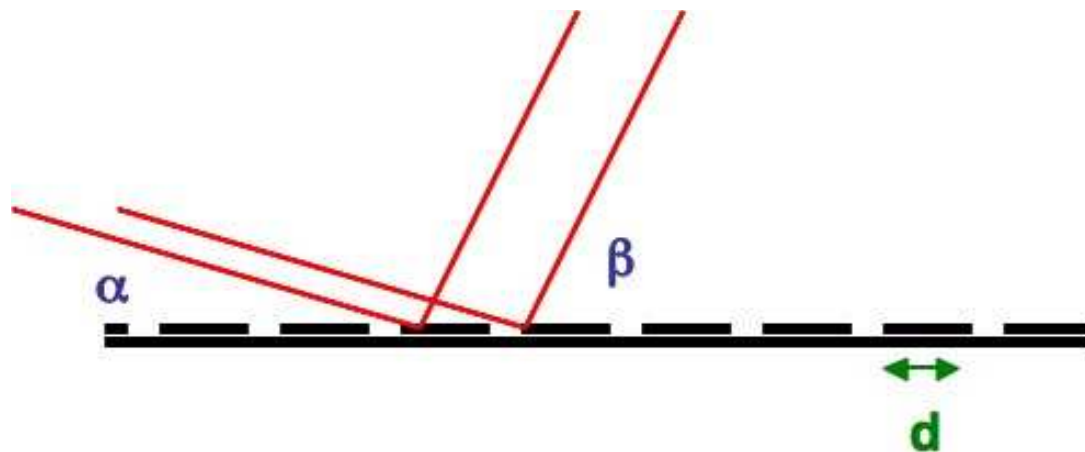


Figure 7.9: Reflection Diffraction Grating

The figure shows what one sees in such a spectrometer. For each emission line there are a series of very well defined bright lines corresponding to $m = 1, 2, 3, \dots$. The sets of lines have different positions because λ is different.

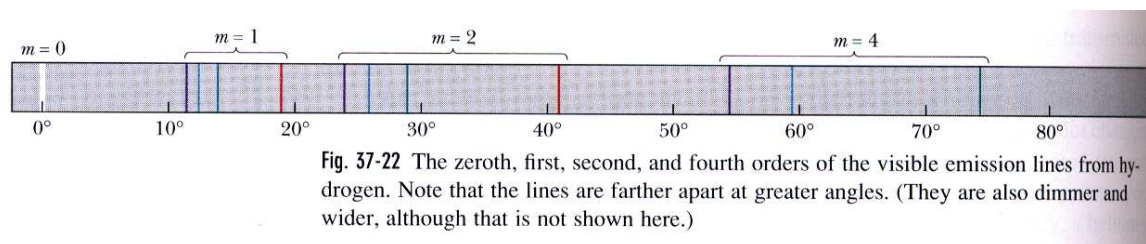


Figure 7.10: Spectrometer output

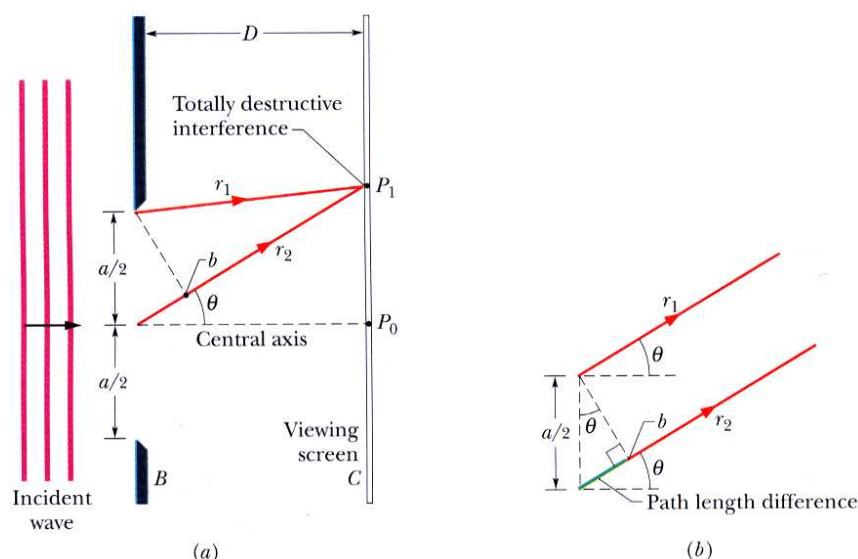
8: More Diffraction

8.1 Diffraction of Waves

Diffraction occurs when a wave encounters an object whose **dimensions** a are **comparable** to its **wavelength** λ . This leads to patterns which would be totally unexpected in naive geometrical optics.

8.2 Diffraction by a Single Slit

If the width of a single slit a is comparable to the wavelength λ of the light, diffraction occurs due to interference between waves coming from different parts of the slit.



To understand this we will use Huygens’ principle. In this context it is going to mean that all points along the slit act as sources of a spherical wave. **You should read and learn about Huygens’ principle from the course text.**

First, we mentally divide the slit into two zones of equal widths $a/2$. Using Huygen’s principle, a wavelet from the top of the upper half of the slit can interfere with a wavelet from the top of the lower half of the slit. The path difference between these waves is the distance from the centre of the slit to point b . If the distance D between slits and screen is much larger than the slit width a we can approximate rays r_1 and r_2 to be parallel. The path difference is then:

$$r_1 - r_2 = \frac{a}{2} \sin \theta$$

The first dark fringe of a single slit diffraction pattern is produced when the path difference is $\lambda/2$. Considering the whole slit, every point in the upper half of the slit interferes destructively with a point $a/2$ below it, so at P_1 the intensity is zero. Thus the waves

will interfere destructively if:

$$\frac{\lambda}{2} = \frac{a}{2} \sin \theta$$

$$\lambda = a \sin \theta$$

This is remarkably similar to the expression for interference, but with a replacing d **except it is for the first dark point - not the first bright fringe**. Be careful!

To calculate the position of the next dark fringe is a little more elaborate than you might think. It is explained in the course text book. Here we only give the result:

A **single slit diffraction** pattern has **dark fringes** above and below the central axis when:

$$m\lambda = a \sin \theta \quad \text{for } m = 1, 2, 3, \dots \quad (8.1)$$

8.2.1 Full intensity pattern of single slit diffraction

In the last section we looked at the bit which is simple to calculate, the dark fringes. To obtain the expression for the actual observed intensity pattern across the whole screen takes a little more work, so we are not going to derive it in full.

The result is that the **intensity** distribution of the diffraction pattern of a single slit is:

$$I(\theta) = I_0 \frac{\sin^2 \alpha}{\alpha^2} \quad \text{where} \quad \alpha = \frac{\pi a}{\lambda} \sin \theta \quad (8.2)$$

It can be seen that this expression is zero for the condition given in 8.1 (and is not zero at $\theta = 0$).

The shape of this expression will be illustrated in the lectures. The shape is the envelope shown in the figure below.

8.3 What pattern do you observe on the screen for a 2 slit interference?

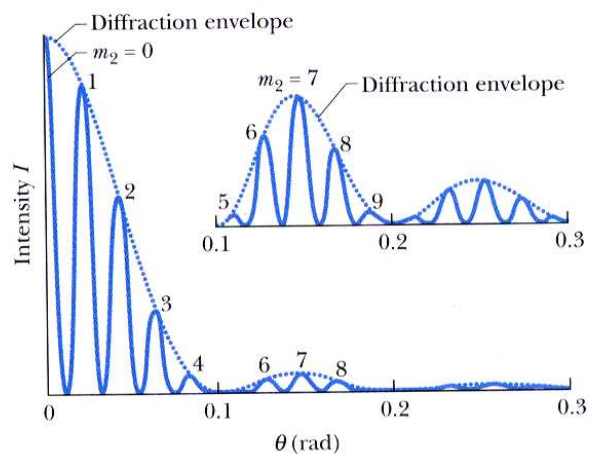
It can easily be the case that you have a 2 slit interference experiment with a slit spacing d and with a slit width a . As we have learnt (and seen) the 2 slits will produce a set of bright fringes with separation:

$$m\lambda = d \sin \theta \quad (8.3)$$

Similarly the width of the slits gives rise to a pattern whose dark fringes are at:

$$m\lambda = a \sin \theta \quad (8.4)$$

It is most likely to be the case that $a \ll d$ thus the interference pattern width will be much narrower than the diffraction pattern width. This is illustrated in the figure below. We see a fast oscillation due to the 2 slit spacing d , modulated by a wider envelope due to the slit width a .



9: X-ray diffraction

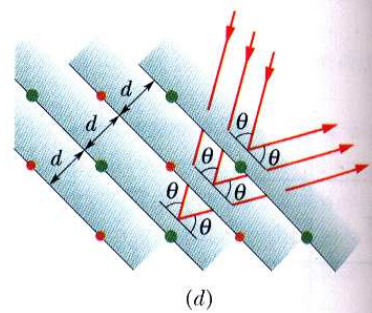
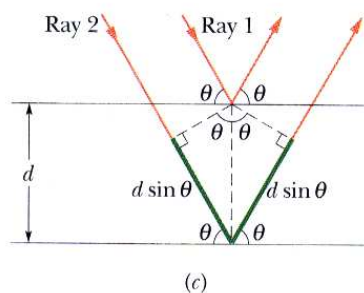
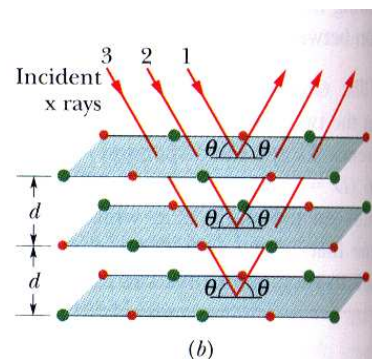
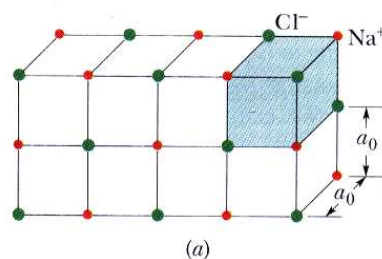
9.1 X-ray Diffraction from a single crystal

X-rays with wavelengths of order $1\text{\AA} = 10^{-10}m$, are used to study the structure of materials. This wavelength is comparable to the spacing between the atoms.

In **crystalline solids** the atoms are regularly spaced. They can be grouped into successive planes, separated by a distance, d , which is characteristic of the structure of the material. Note that for each solid there are several different sets of planes with different orientations and spacings.

The X-rays are reflected by successive planes of atoms to create a pattern which is similar to that of a diffraction grating (but note the crucial extra factor of 2). The maxima are described by **Bragg's Law**:

$$m\lambda = 2d \sin \theta \quad (9.1)$$



The really important result to keep in mind is:

- The end effect is as if the X-rays **reflect** off the planes as if they were a mirror.
- This reflection can only happen for specific angles given by Bragg's law.

- There is no reflection for any other angle.

In a more likely situation the X-ray beam is fixed (it probably comes from a big X-ray generator) and the crystal is oriented in different directions until the Bragg condition is satisfied. The series of pictures illustrate this.

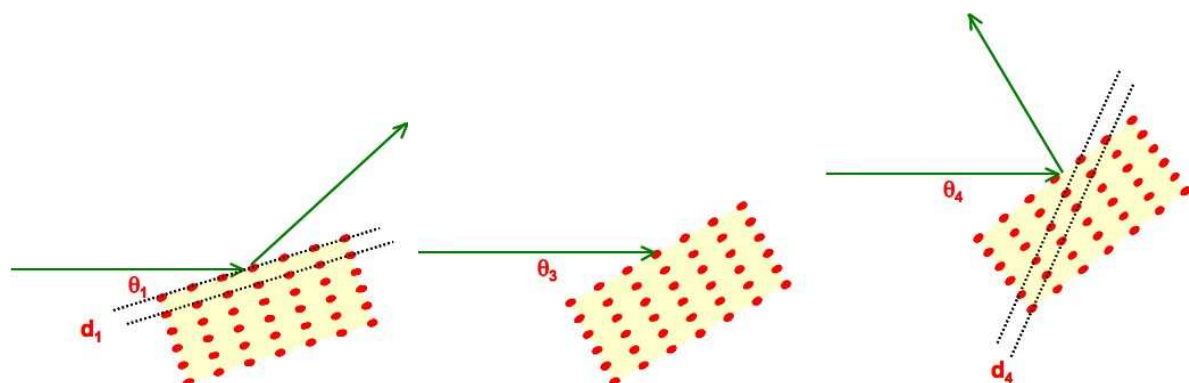


Figure 9.11: Successive orientations of a crystal in a fixed X-ray beam. Only certain orientations satisfy the Bragg condition.

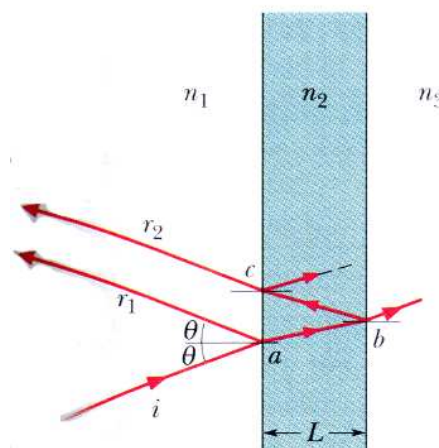
9.2 X-ray Diffraction from a powder

Normally one does not have a single crystal. Instead it is more likely one has a powdered crystal. A powder can be thought of as containing many thousands of mini-crystals. They will all be randomly oriented. Therefore there will always be some mini-crystals oriented in the right direction for every possible Bragg condition. This has two consequences:

- There will be a many diffracted rays, corresponding to all possible sets of plane spacings.
- Since the mini-crystals will be oriented randomly around the axis of the incoming X-ray beam then the observed pattern will be concentric circles around this axis.

9.3 Interference due to Thin Films

Consider a light wave incident normal to a thin transparent film (e.g. soapy water) with a refractive index n . One wave is produced by a reflection from the upper surface of the film. A second wave comes from light transmitted through film and then reflected from the lower surface of the film.



The **phase difference** between two waves in a thin film can change by:

- The **path length**, $2L$, through the thin film.
- The **lower wavelength**, $\lambda_{\text{film}} = \lambda/n$, and speed of light, $v_{\text{film}} = c/n$, inside the film with **refractive index** n .
- A **phase shift of half a wavelength**, π , when light is reflected at the surface of a **higher refractive index** material.

Note that there is **no phase shift** when light is reflected at the surface of a **lower refractive index** material.

The light wave reflected off the upper surface of the soap film gets a phase shift π ($n_1 = 1 < n_2$). The light wave that reflects off the lower surface has no phase shift from the reflection ($n_2 > n_3 = 1$), but gets a phase shift due to the path length and the refractive index $n = n_2$, which results in a change to the wavenumber, k_{film} . The phase difference between the two waves is:

$$\Delta\phi = k_{\text{film}}2L - \pi = \frac{2\pi n}{\lambda}2L - \pi = \frac{4\pi nL}{\lambda} - \pi$$

We have assumed that the light rays are almost perpendicular to the film, i.e. $\theta \approx 0$. There is **constructive interference** – **bright maxima** for:

$$2L = (m + \frac{1}{2})\frac{\lambda}{n} \quad (9.2)$$

where the factor $\frac{1}{2}$ accounts for the π phase shift from the reflection at the upper surface.

There is **destructive interference** – **dark minima** for:

$$2L = m\frac{\lambda}{n} \quad (9.3)$$

Note that zero film thickness, $L = 0$, corresponds to destructive interference!

When white light is reflected off a thin film it produces **coloured bands**. Other examples of thin film interference are **non-reflective coatings** for glass, and **Newton’s rings** which are observed when there is a small air gap between a curved and a flat piece of glass.