Modern Physics

DR MUHAMMAD SABIEH ANWAR
(LAHORE UNIVERSITY OF MANAGEMENT
SCIENCES)

Contents

Preface

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetuer.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

Dr Muhammad Sabieh Anwar

Chapter 1

An atomistic view of energy

We begin our discussion by looking at a fundamental concept in physics. This is energy. Energy and mass are the two tenets of universal existence. Since this is a course on modern physics and this necessarily deals with bits, atoms, photons or molecules, we would like to take energy to the atomistic realm. This chapter discusses the atomistic view of energy and through it, introduces how energy is distributed with atomic systems, how it changes and how these changes are measured in spectroscopic techniques. Furthermore, we will also motivate why energy is quantized. This may be our first rendezvous with quantum physics.

Before we proceed to an atomic description of energy, let's proceed with the energetics of some mechanical systems and how they are modeled.

1.1 What is a mathematical model?

In physics, systems are modeled mathematically. A model is an equation that describes the behavior of an object.

For instance, consider a second order differential equation of the form

$$m\frac{d^2x}{dt^2} + \beta\frac{dx}{dt} + kx = 0 \tag{1.1}$$

where m, β and k are constants, $x = x(t)$ is the position of some "object". This equation models a large variety of physical processes. A few examples are in order.

1. An object of mass m is attached to a string and placed on a frictionless floor. The other end of the spring is connected to a wall (Fig. (1.1 (a))). This physical system, almost present in every physics laboratory also abodes in our daily lives. It is modeled through Eq. (1.1) which predicts the behavior of the system.

2. The behavior of a resistor R connected with an inductor, a capacitor C, a switch and a battery of voltage V in a circuit can also be explained by Eq. (1.1). This may not be entirely evident but let me state the equivalent that is normally found in many books:

$$L\frac{d^2i}{dt^2} + R\frac{di}{dt} + \frac{i}{c} = 0 \tag{1.2}$$

   The structures of the equations (1.1) and (1.2) are the same. If one were to make the substitution $x \leftrightarrow q$, $L \leftrightarrow m$, $\beta \leftrightarrow R$ and $k \leftrightarrow \frac{1}{C}$, the equations are in fact identical. Two disparate physical systems are described by a model. That's the elegance and universality of models. They are generic. Let's look at a third example.

3. Consider a photon trapped inside a cavity made by two perfectly reflecting mirrors. The field of the photons comprises of electric and magnetic fields that are perpendicular to each other and change sinusoidaly with time.

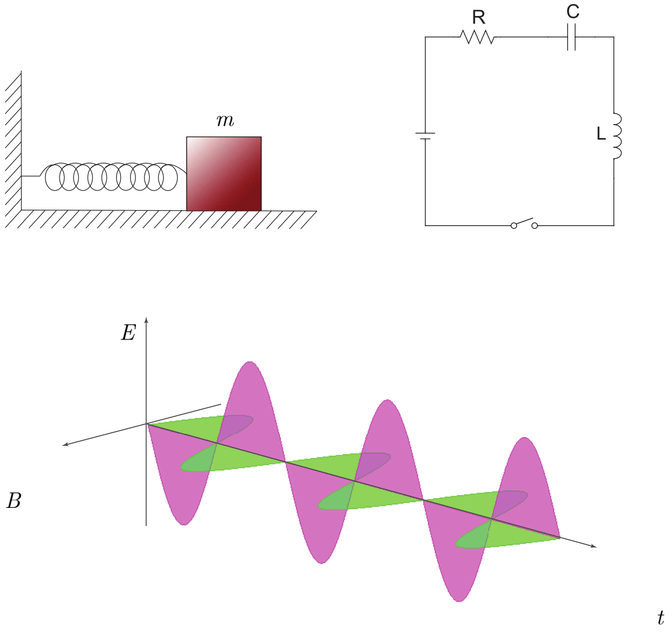


Figure 1.1: Examples of mathematical models

1.2 Harmonic oscillator

The harmonic oscillator is one of the most famous models in physics. Many the physical systems, such as the examples from the previous section (mass attached to spring, RLC circuits, and trapped light) can be modeled through the harmonic oscillator.

Revisiting a mass attached to a spring. Let's probe this a bit further. It will help us forage into the energetics of atomic systems where defining model is described in Eq. (1.1). This system is described by Hooke's law $F = -kx$, where k is the proportionality constant (spring constant) and x is the displacement from the equilibrium position. This law states that the larger the force, the larger will be the extension or compression[1].

DERIVATION

$$
\begin{aligned}
F &= -kx \\
ma &= -kx \\
ma + kx &= 0 \\
m\frac{d^2x}{dt^2} + kx &= 0
\end{aligned}
$$

If a spring is stretched, it will store some elastic energy but will remain at rest. Therefore, the spring does not possess kinetic energy, rather it only has elastic potential energy stored in it. Similarly, if the spring is compressed, it is still at rest and therefore possesses only potential energy. The potential energy is denoted by U.

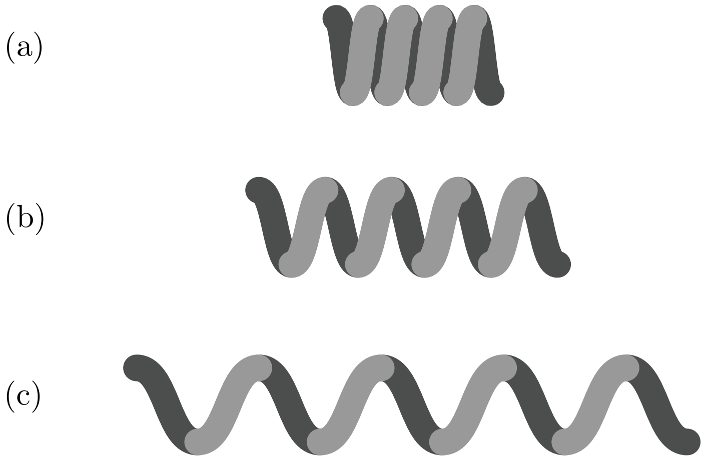


Figure 1.2: Spring compression, relaxation and extension

---

[1] Or, as Robert Hooke himself stated, *ut tensio, sic vis* ("as the extension, so the force")

The potential energy U and the resting force F are related as

$$F = -\frac{dU}{dx} \tag{1.3}$$

Therefore, the potential energy is found by integration.

$$\begin{aligned} U &= -\int F dx \\ &= -(-k)\int x dx \\ U &= k\frac{x^2}{2} + C \end{aligned} \tag{1.4}$$

where C is an integration constant. This is the relationship between the potential energy stored in the spring-mass system and the displacement x.

Let's plot U as a function of x. At $x = 0$, $U = 0$ which indicates the normal equilibrium state of the spring. If we stretch or compress the spring, the potential energy increases.

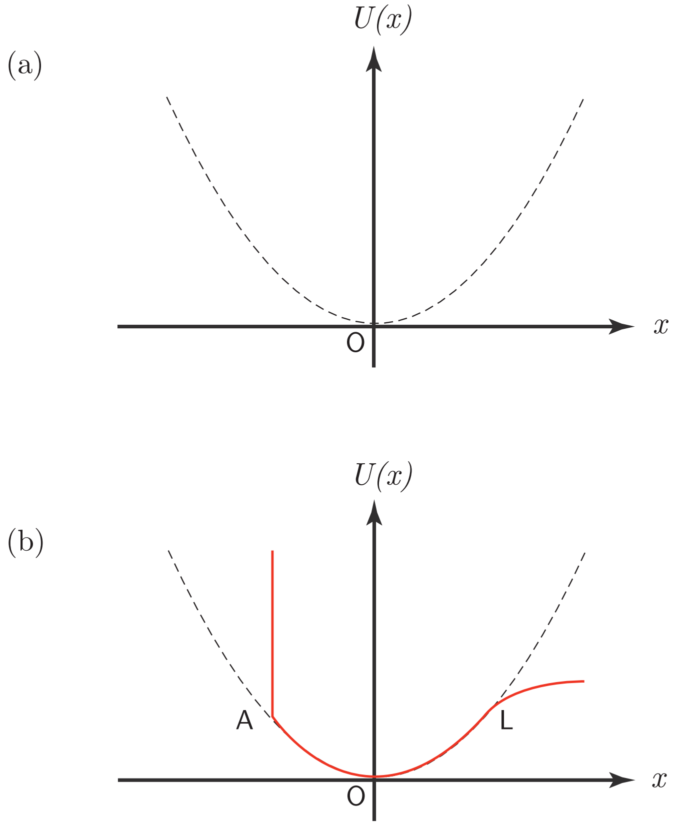


Figure 1.3: U-x plot for an ideal and non ideal harmonic oscillator

Fig. (1.3 (a)) is the potential energy versus position curve of an ideal harmonic oscillator. This curve is iconic and with modifications, form the basis of energy combinations inside atoms and in ??? of atoms too. In reality, however, no oscillator is ideal. For example, if we extend or compress the spring and remove the force, it will return to its original state. But if we keep on increasing the force, there will be a point from where the spring will not be able to restore its normal state; it will be permanently deformed and will no longer act as a harmonic oscillator. Thus, there is a limit to the harmonic oscillator beyond which it will deviate from ideality.

Fig. (1.3 (b)) shows that the real curve is very close to ideal when x is small. Once x becomes large, the curve deviates from ideality. So, in the small region AOL, we have a good approximation of the ideal harmonic oscillator when the displacement is small.

Potential and Kinetic Energy with respect to time

The mass-spring system in its idealized form is a harmonic oscillator. What if it is displaced from equilibrium and set free? It will ”oscillate”.

Let's now plot the potential energy $U(t)$ against time t as this undulating motion progresses. The potential energy oscillates between a minimum value 0 (at equilibrium $x = 0$) and a maximum value (at extreme positions $x = \pm x_0$). Of course, it cannot have a negative value.

The potential energy is not the only kind of energy possessed by the harmonic oscillator. By virtue of its motion, it also has some speed v yielding a kinetic energy.

$$\begin{aligned} K &= \frac{1}{2}mv^2 \\ &= \frac{1}{2}m(\frac{dx}{dt})^2 \qquad (1.5) \end{aligned}$$

The kinetic energy also oscillates between the minimum (at $x = +x_0$) and the maxima (at $x = 0$). When the potential energy is zero (at equilibrium position), the kinetic energy is maximum. When we reach a turning point (the maximum compression or extension), the mass comes to rest momentarily and becomes zero. Therefore, a graph similar to that of the potential energy can be drawn for the kinetic energy of the oscillator.

The total energy $E = U + K$ is conserved for the oscillator and we can draw

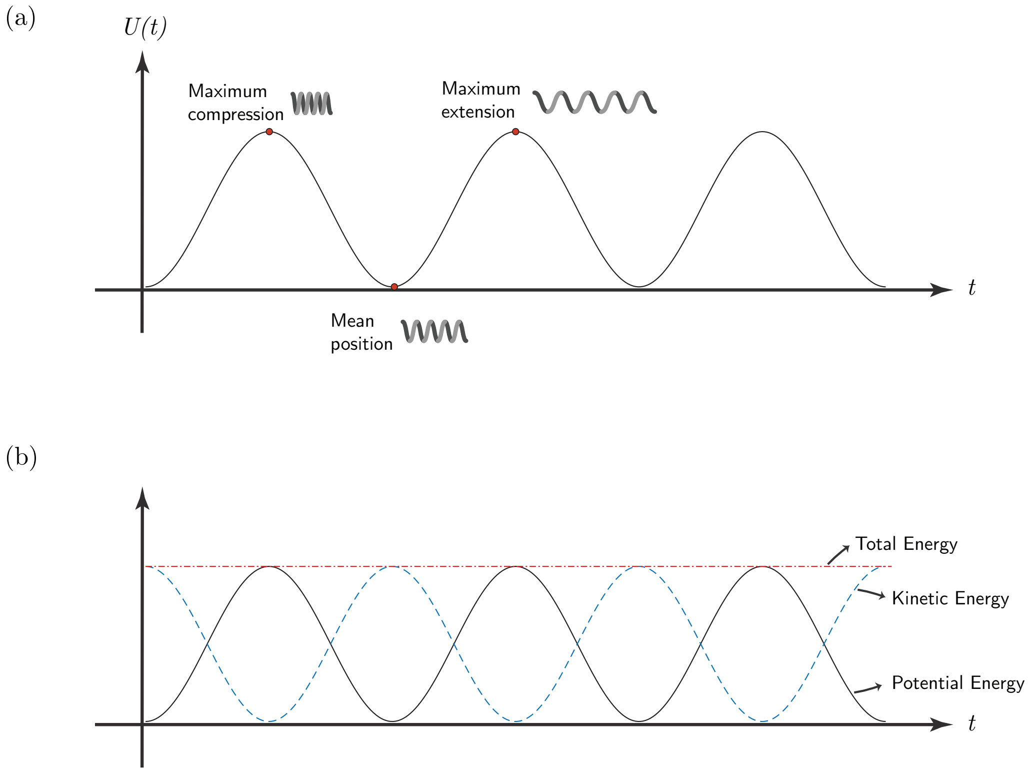


Figure 1.4: U-t plot and total energy for a harmonic oscillator

a straight line corresponding to the constant E. There is a never-ending interconversion between K and U. We can prove this mathematically as well.

A solution of Eq. (1.1) is

$$x(t) = x_0 \cos \omega t$$

Hence, U from Eq. (1.4) becomes

$$\begin{aligned} U &= \frac{1}{2}kx^2 \\ &= \frac{1}{2}k(x_0 \cos \omega t)^2 \\ &= \frac{1}{2}kx_0^2 \cos^2 \omega t \end{aligned}$$

as drawn in Fig. (1.5).

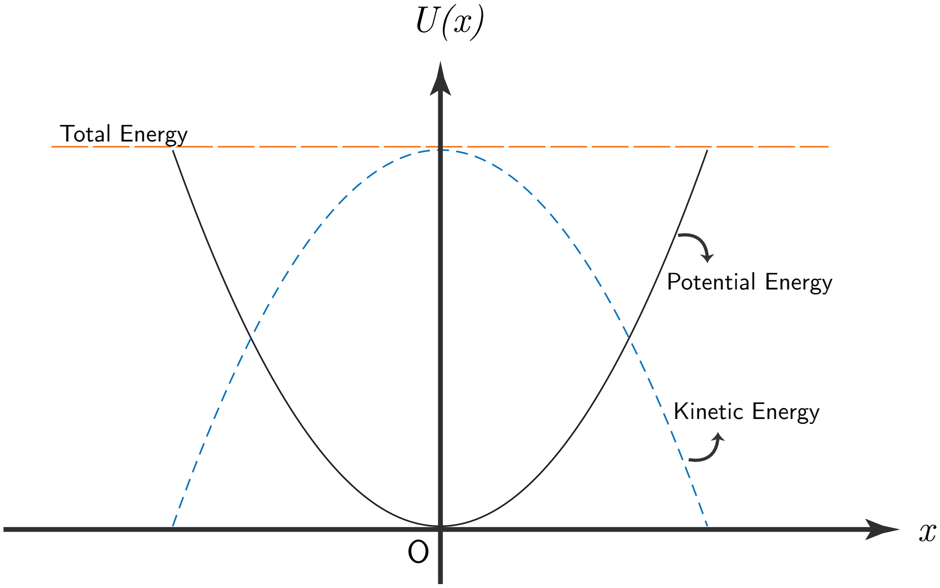


Figure 1.5: Plot of total energy

Similarly, K is

$$
\begin{aligned}
K =& \frac{1}{2}mv^2 \\
=& \frac{1}{2}m\left(\frac{dx}{dt}\right)^2 \\
=& \frac{1}{2}m(\omega x_0 \sin\omega t)^2 \\
=& \frac{1}{2}m\omega^2 x_0^2 \sin^2\omega t \\
=& \frac{1}{2}kx_0^2 \sin^2\omega t
\end{aligned}
$$

Adding these two kinds of energies

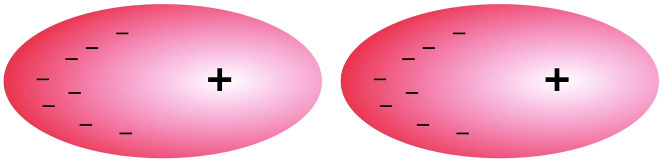
$$
\begin{aligned}
E =& K + U \\
=& \frac{1}{2}kx_0^2(\sin^2\omega t + \cos^2\omega t) \\
=& \frac{1}{2}kx_0^2 \\
=& \text{constant}
\end{aligned}
$$

This shows that the higher the amplitude of oscillation, higher will be the energy of motion. The energy breaks up like a see-saw between its kinetic and potential forms. This is characteristic of any harmonic oscillator.

1.2.1 Interaction between atoms

We go straight from a mass-spring system to two neutral atoms brought in close proximity to each other. This is quite a cataclysmic shift, but as I will show, this physical situation, under certain simplifications, resembles a harmonic oscillator, at least as far as the model is concerned. When two atoms are placed far apart , they will not interact with each other and will behave as two isolated, independent and identical atoms. Their potential energy of the system can be counted as zero, which is a datum energy.

(a) Two atoms far apart



(b) Two atoms close together

Figure 1.6: Two atoms far apart and close

As we move one atom closer to atom the other, its nucleus will attract the electron cloud of the other. As a result, the spherical electron cloud will slightly change its shape and become elongated. This charge redistribution is shown in Fig. (1.6 (b)) through a highly exaggerated form. Each atom becomes a temporary (induced) dipole and will result in nudging the atoms even closer. Therefore, if we bring atoms closer together, they take advantage and use innate electrostatic forces to facilitate the mutual attraction.

Let's plot the potential energy of the system U versus the distance r between atoms (Fig. (1.7)). Since the atoms are extended objects, we can consider r to be the distance between the nuclei.

At point r_1, the atoms are very far apart (r is very large), so the potential

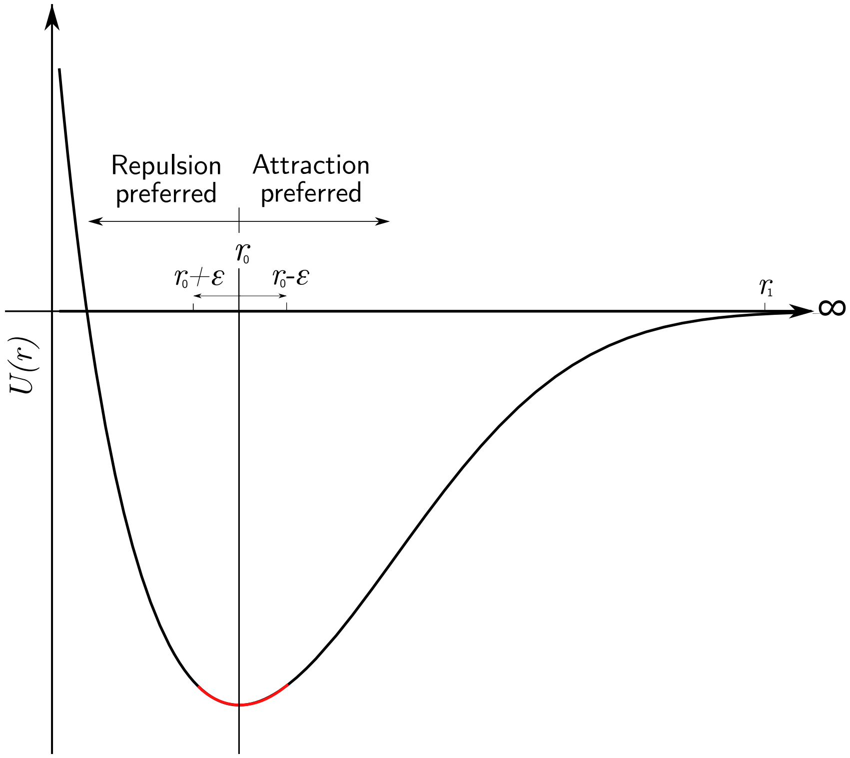


Figure 1.7: Morse potential

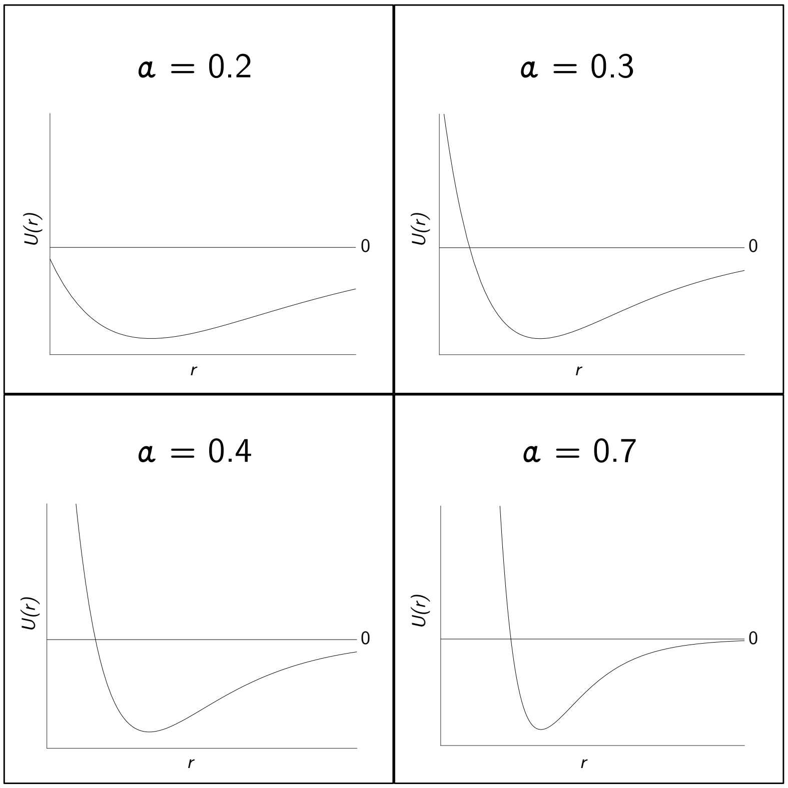
energy is zero. As they come closer, electrostatic attraction develops between them and the electron clouds and the atoms would like to come closer. Hence, the potential energy decreases until, at $r = r_0$, U reaches a minimum. If we were to squish the atoms closer together, both nuclei start to repel each other. Now the only way to further push them together is by pumping in more external energy. In other words, U takes a twist and blows up.

Even though it doesn't look so, this pair of atoms is quite similar to the harmonic oscillator. However, there are some notable differences. First, we observe regions of preferred attraction and preferred repulsion. We also found the U curve of the harmonic oscillator to be symmetric about $x = x_0$ (Fig. (1.7)). However, for the atoms, the curve is not symmetric; it tapers off at large r and goes to infinity at small r. It approximates the harmonic oscillator only in the region $(r - r_0) < \varepsilon$ where r is really close to r_0. It is within this narrow region that the inter-atomic interaction can be appreciated by a harmonic oscillator. Outside the ???, the system is ???.

This graphical description of P.E can also be written in form of an approximate equation, called *Morse potential*[1].

$$U(r) = E_0\left(1 - e^{-\alpha(r-r_0)}\right)^2 - E_0 \tag{1.6}$$

where α is some positive constant. By changing its value, the width of depression in the P.E curve changes. However, the critical points remain unchanged.



If r approaches infinity, the potential energy becomes

$$U(r) = E_0(1-0) - E_0 = 0$$

Whereas at $r = r_0$, the potential energy is

$$U(r) = E_0(1-1) - E_0 = -E_0$$

Sometimes r_0 is called the bond length and E_0 is called the bond energy.

Thus, the description of the behavior of harmonic oscillator is automatically carried over to the description of interaction between atoms, albeit in a small region.

1.2.2 Einstein Solid

An ideal solid has numerous atomic forms enmeshed in an intricate network called a crystal structure. A simple network is the Einstein solid which is a mesh of harmonic oscillators. For example, a polonium crystal. I use this rather strange example of polonium because it is a simple cubic crystal structure. Most other elemental solids have slightly more complicated structures (body-centered or face-centered cubic). can be modeled on a 3D network of iron atoms connected by springs. Each atom of the crystal is 'attached' to its neighboring atoms through these springs. These spring-like interactions are described by Morse-like potentials, which are in turn approximated by a harmonic oscillator. The polonium crystal is therefore a giant three-dimensional harmonic oscillator.

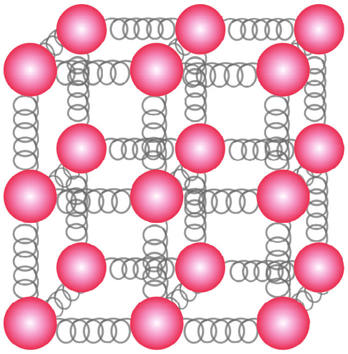


Figure 1.8: 3D polonium crystal

Thus, we can view any solid as a large network of harmonic oscillators, in turn allowing us to apply everything we have learned about harmonic oscillator to solids. In fact, this is how Einstein originally thought about the heat capacities of solids: modeling solids as a network of harmonic oscillators. We have now introduced the model called the harmonic oscillator and noticed how it applies to real solids. We can introduce the concept of the total energy of a solid.

---

[1] Named after the physicist Philip P. Morse

1.3 The Internal energy of atomic systems

1.3.1 Diatomic Molecules

The atomic pair of section 1.1.1 can exist independently. It is then called a diatomic molecule. It can be compressed or extended configurations(Fig. 1.9). The molecule, on a whole can also be moving with a speed v. Let's find out the molecular energies.

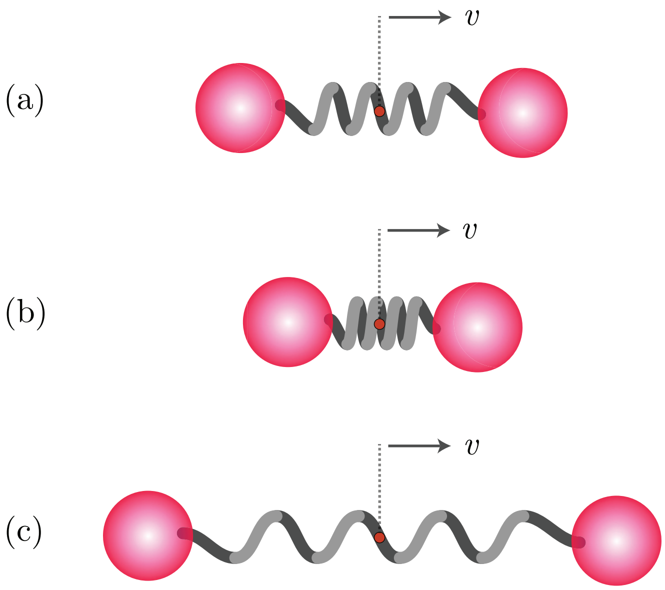


Figure 1.9: Diatomic molecule moving with velocity v

The most straightforward is the kinetic energy of the molecule as a whole totally irrespective of the internal compressional configuration. This energy is

$$K_{\mathrm{CM}} = \frac{1}{2}Mv^2 \tag{1.7}$$

where v is the velocity with which the center of mass (M) of the molecule is moving. Therefore, this kinetic energy is associated with the center of mass of molecule. All three cases shown in Fig (1.9) have the same center of mass and kinetic energy.

However, the molecule will also have some additional energy present *inside* the molecule. This energy is called *internal energy* and can be denoted by

U or E_{int}. Cases (b) and (c) will have higher internal energy E_{int} than (a), even though all three have the same K_{CM}. Hence, the total energy of each of these molecules is given by

$$E = K_{\text{CM}} + E_{\text{int}} \tag{1.8}$$

For the diatomic molecule that is free to oscillate just like any ideal harmonic oscillator, $E_{\text{int}} = \frac{1}{2}kx^2 + \frac{1}{2}mu^2$ as described in Eq (1.4), allowing us to write $E = K_{\text{CM}} + \frac{1}{2}kx^2 + \frac{1}{2}mu^2$ where $\frac{1}{2}mu^2$ is the effective kinetic energy inside the molecule. The variable m and u are effective masses and speeds. Note that x is measured from the equilibrium position relative to the CM. So x is internal to the body of the molecule. The body of the molecule may itself be moving with a velocity v which impacts K_{CM} to the overall energy.

”Internal energy of an object is the energy that is measured inside the frame of reference of that object (atom, molecule, etc.)”.

For example,

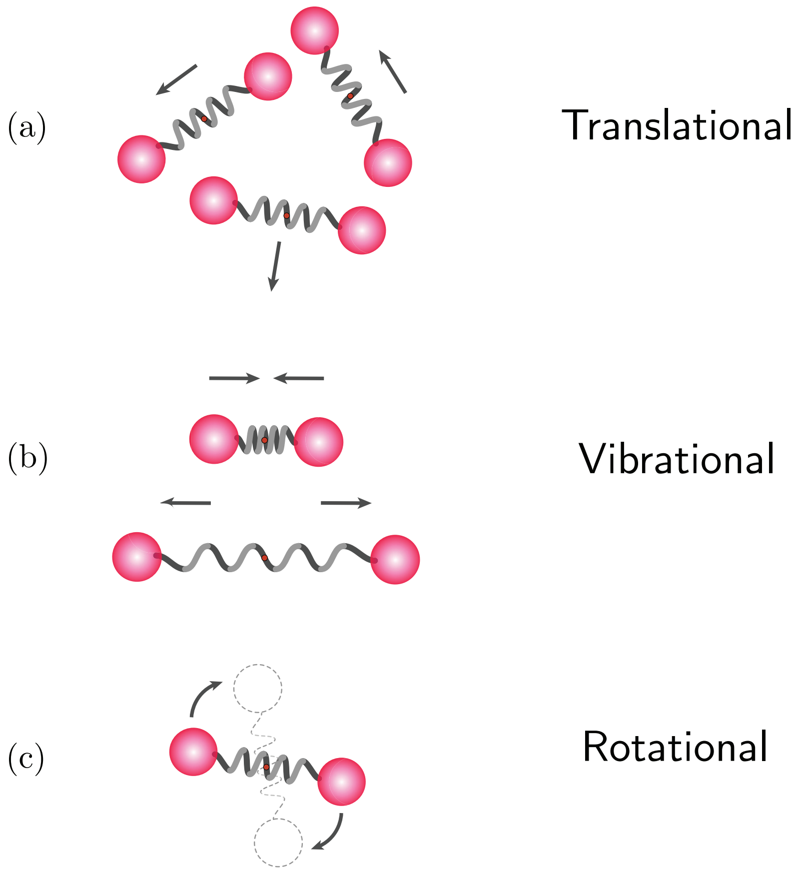


Figure 1.10: Translational, rotational and vibrational motion of a diatomic molecule

Components of Internal Energy and degrees of freedom

The example of the diatomic molecule highlights that the molecule is free to translate and is also free to oscillate. In physics parlance, we say that the molecule has two degrees of freedom, translational and vibrational. The latter is an internal degree of freedom.

Let's have a look at the different components of internal energy.

- A translating molecule has translational energy E_{trans} associated with it which adds to E_{int} of the system.
- The energy associated due to the vibratory motion of oscillators, like the compression and extension of a spring, is called **vibrational energy** E_{vib}.
- If we take the axis of rotation at the center, the molecule may also rotate about this axis of rotation. Hence, there is a pure kinetic energy associated with this rotation, called rotational kinetic energy E_{rot}.

  There could also be numerous other exotic forms of energy. For example, sunlight falls on chlorophyll and energy is absorbed inside the ATP molecules. There is no kinetic, rotational and vibrational energy, yet energy is still stored. This chemical energy E_{chem} also contributes to internal energy. In fact, E_{int} can also comprise other energies such as magnetic energy, electrostatic energy etc.

Mathematically, the internal energy budget is accounted by all these various forces.

$$E_{\text{int}} = E_{\text{trans}} + E_{\text{vib}} + E_{\text{rot}} + E_{\text{chem}} + E_{\text{mag}} + E_{\text{elect}} + \cdots \tag{1.9}$$

Finally, note that E_{vib} is both kinetic and potential while E_{trans} and E_{rot} are purely kinetic. We now have a fair idea how the energy of a system confining atoms, or solids with some internal structure, can be computed. Identify the various degrees of freedom and contemplate whether they have kinetic or potential energy components.

1.4 Thermal Energy

The components of internal energy can be of two kinds, ordered and disordered. The disordered component is called thermal energy.

$$E_{int} = E_{thermal} + E_{chem} + E_{mag} + E_{elect}$$

where

$$E_{thermal} = E_{trans} + E_{vib} + E_{rot} \tag{1.10}$$

This requires a bit of explanation. In an Einstein solid, as the equilibrium positions of the atoms are fixed inside the crystal structure, but the atoms undergo vibrational motion. If we had a solid in which all of the springs, or bonds, stretched and compressed at the same time in perfect unison, this synchronous motion would have qualified as ordered motion. In reality, these vibrations are completely out of sync (Fig. 1.11). Some of these springs undergo extension while others undergo compression at the same time. In ??? ???, the compressions and extensions are out of phase. This random vibrational motion contributes to the thermal energy of the system. The random jostling goes up with the temperature. Similarly, vibrational and translational motion are also degrees of freedom to which no particular order or directivity can be associated. They all contribute to thermal energy. Sometimes, ordered degrees of freedom are called 'waves' or 'modes' in physics terminology.

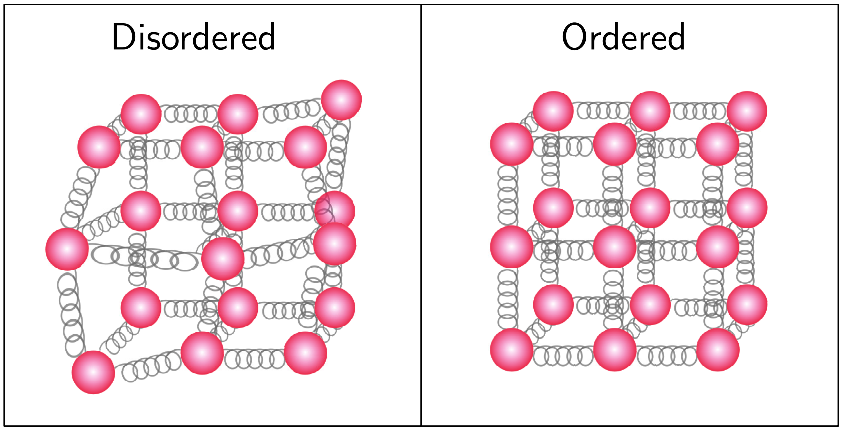


Figure 1.11: Einstein solids undergoing disordered and ordered motion

1.4.1 Measuring thermal energy

Suppose we have a volume of solid comprising 6×10^{23} atoms. This is one mole of a monoatomic solid with each atom connected to six neighboring atoms. The number of distinct harmonic oscillators in the solid is $N = 3 \times 10^{23}$. This means that to calculate the thermal energy of the solid, we will have to find 3×10^{23} values of energy. Furthermore, each energy has a kinetic and potential energy component.

A puritanist theorist would write the energy of each harmonic oscillator inside this solid and average over all these energies to obtain the thermal energy.

$$\frac{\sum_{i=1}^{N} E_i}{N} = E_{thermal}. \tag{1.11}$$

This task is evidently daunting and close to impossible. Fortunately, we can agree upon some definitions and directly measure the average thermal energy by using an appropriate instrument. This instrument is called a thermometer and the average thermal energy is called the temperature. One of the most well-known thermometers is of the mercury kind.

Suppose we have a solid, in which a hole is drilled and a glass bulb thermometer containing mercury is placed into it.

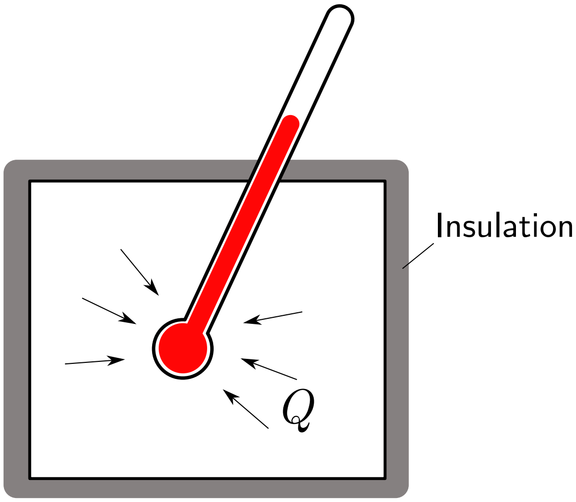


Figure 1.12: A mercury thermometer is immersed into a hole drilled into an insulated solid

This solid is thermally insulated from its surroundings. The mercury level will either rise or drop depending on the temperature of the solid. The atoms inside the solid are undergoing random motion. They have thermal energy and as they come in contact with the glass thermometer or with the air molecules in the intervening air,they will transfer their energy to the mercury's or air's atoms. As a result, there is energy transfer from the solid to the mercury. This is denoted as Q in Fig (1.12).

If one were to taker a closer look at the energy transfer process, in this transfer of energy, there is work done on air molecules by the solid atoms but this work is done only on the microscopic scale. There is no evidence of macroscopic work. *This energy transfer Q entails microscopic work, but no macroscopic work.* Eventually, there is a transfer of thermal energy to the mercury atoms and the mercury atoms expands.

1.4.2 Explaining thermal expansion

Let's look into the potential-energy curve to understand the expansion of mercury. The harmonic oscillator model and its anharmonic upgrade. The Morse Potential curve comes in real handy here.

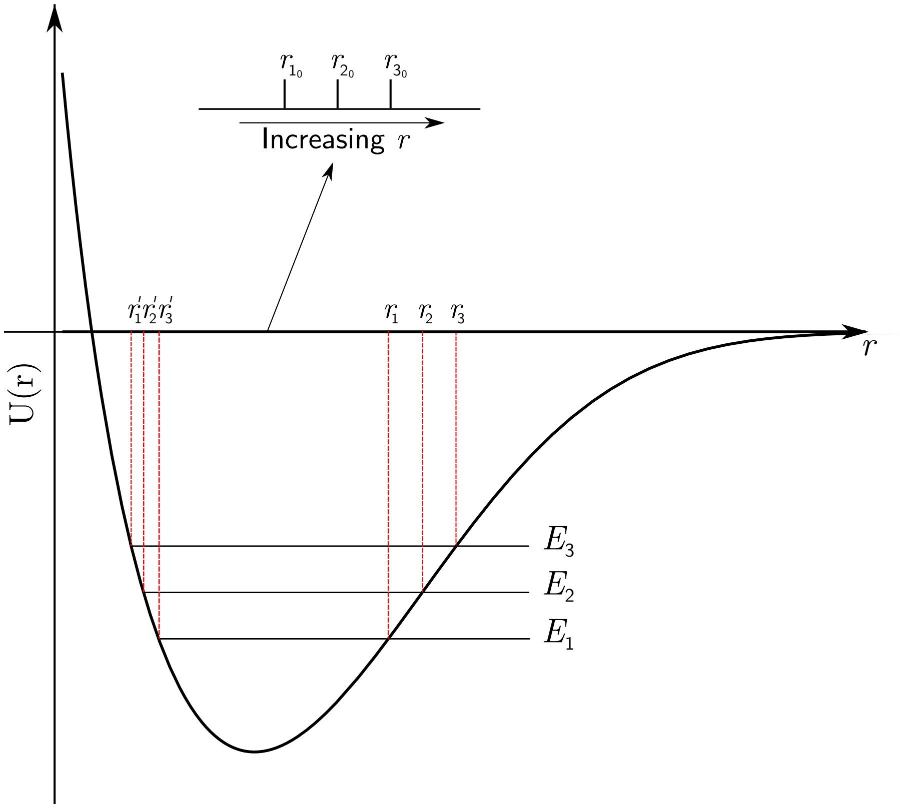


Figure 1.13: Morse potential for mercury

Let's follow the Morse potential curve for mercury atoms. Here r_0 is the equilibrium distance between the two mercury atoms. As the mercury atoms are brought closer to each other, both attractive and repulsive interactions come into play. At a particular value of thermal energy, we can represent the total energy E by a straight line, before the thermometer comes in contact with solid. Consider three energies, E_1, E_2 and E_3 where $E_1 < E_2 < E_3$. For E_1, the U curve intersects this straight line at $r = r_1$ and $r = r_1'$, therefore the range of $r's$ associated this straight line of E_1 is (r_1, r_1'). Since everything is in equilibrium, the average distance between the two atoms r_{1_0} must be at the center of r_1 and r_1'. As the thermometer comes in contact with the solid, the thermal energy of the mercury column increases. Thus, the total energy must increase. Say it goes up to E_2. With the increase in the total energy E, the average distance between the two atoms increases, i.e. if we interpolate the points r_2 and r_2' on distance curve, the average of these points has moved to the right r_0', which means the average distance between the atoms has increased. Viola! This is thermal expansion. If the solid is at a higher temperature, more thermal energy is transferred to the mercury atoms and we will get another energy line, E_3, further up. The average distance between the two atoms also increases further and the volume of mercury will increase as a consequence. Therefore, as the thermal energy of mercury atoms goes up, volume expansion takes place. It is our fundamental harmonic oscillator model and its variants that allows one to adequately explain thermal energy expansion.

1.4.3 Basic meaning of temperature

The height of the mercury column inside the glass capillary indicates the thermal energy inside the mercury column and the glass walls of thermometer. This, in turn, indicates the thermal energy of the air molecules, which bespeaks the thermal energy of the solid itself.

Note that if this device, called a thermometer, has to work, there has to be good contact between the solid and mercury, and mercury must have smaller mass as compared to the solid so that we can calibrate the thermometer scale in terms of thermal energy. Thermal energy has units of joule (J) but conventionally, we calibrate thermometers in terms of temperature. Hence, there is one to one correspondence between thermal energy and the temperature. It means temperature is a measure of the thermal energy of the system. The thermal energy is the disordered part of the internal energy.

1.4.4 Voltage and thermal energy

There are a host of clever ways to measure thermal energy and temperature. Consider a solid with a grove in it (Fig. (1.14)). A battery and an ammeter are connected in series with two resistors R_1 and R_2. The resistor R_1 is studded into the solid with which it makes thermal contact. A voltmeter is used to measure the voltage across R. The passage of current heats R_1 which increases the thermal energy of the solid, but as the thermal energy of atoms inside the solid changes, the resistance R_1 changes accordingly. This change in resistance will be registered as a change in voltage across R_1 and will be measured by the voltmeter. Hence, this voltage is a measure of the thermal energy inside the solid. This scheme is an example of thermal energy measurement using an electric circuit.

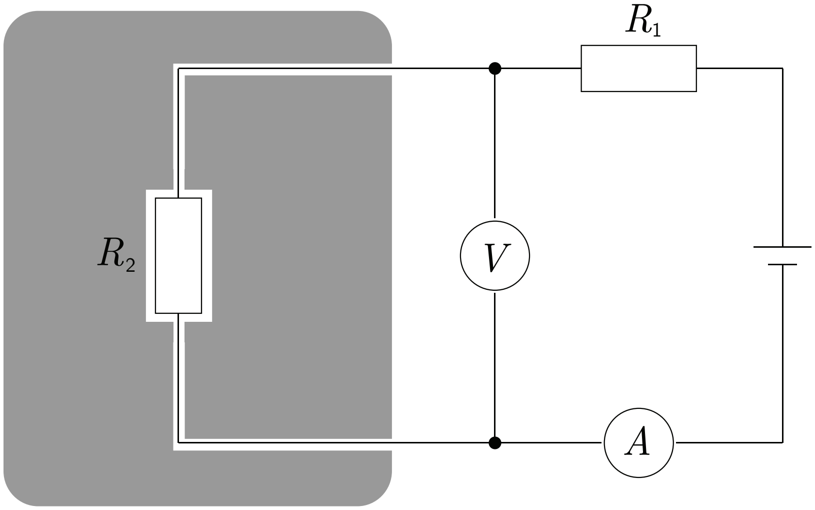


Figure 1.14: A device to measure thermal energy

In the previous examples, we've seen how thermal energy transfers from one solid to another, such as from the resistor to the surrounding block. The transit of thermal energy is accomplished by microscopic work.

1.4.5 Flow of thermal energy

Let's take two solids of unequal size lying separately from one another. The larger is hotter and the smaller is colder. We establish contact between the two atoms and place the complete system inside thermal insulation.

As the solids are in contact with each other, There will be a flow of energy from the hot object to the cold one.The atoms in the hot object undergo random motion and transfer a part of their energy to the atoms in the cold object. This energy flow involves microscopic work only and no macroscopic work is involved because the center of mass of the cold object does not move. This energy flow is generally denoted by Q and can be regarded as energy in transit. The process is also called heat transfer.

Note that there is no such thing as heat *inside* an object. It is the thermal

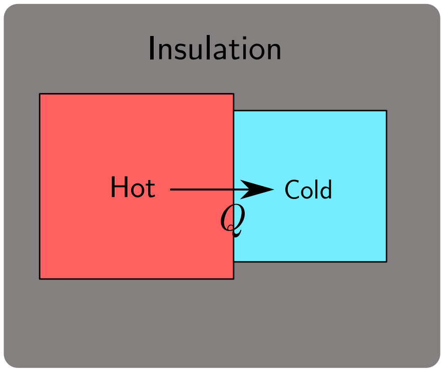


Figure 1.15: Flow of heat from a hot body to a cold body

energy which is present inside an object due to the disordered motion of atoms and which can be increased by various means such as transfer of energy.

1.5 The First Law of Thermodynamics

So far we have talked about the energy flowing into objects through microscopic work thus raising the internal energy but the increase in internal energy does not only depend on the flow of heat. There are other means too. We open up the discussion by a simple example.

Consider a reservoir of water such as the Mangla Lake. River Jehlum flows into Mangla Lake. Consequently, the volume of water in the lake increases. Pressures gradients cause the river to flow into the reservoir. The flowing stream of water can do mechanical work. This work is of the microscopic kind.

However, it is quite possible to stop the flow of river into the lake by using floodgates. In this case, there is no macroscopic work done on the lake. Even in this case, the volume of water may not remain preserved. This is because of *condensation* and *evaporation* – evaporation will decrease the volume of water in the reservoir while condensation will increase it. These local phenomena can change the volume. The processes of evaporation and condensation entail microscopic work done between the water and the air molecules and are accompanied by heat transfer. The volume of water in the lake roughly represents its internal energy. This energy can increase or decrease depending upon whether one does macroscopic or microscopic work. Macroscopic work is generally denoted by W and microscopic work by Q. Earlier we have associated Q with energy transfer as well. Hence, we

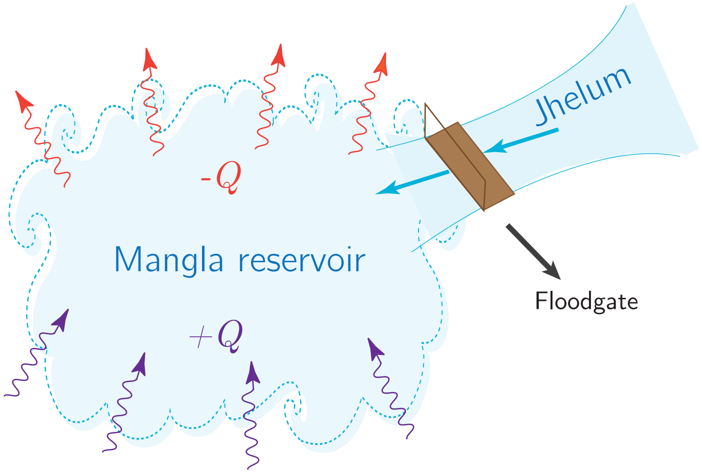


Figure 1.16: Jhelum river flowing into Mangla reservoir

can say that the change in E_{int} is due to multiple components.

$$\Delta E_{int} = Q + W + \text{other forms of energy transfer} \tag{1.12}$$

This statement is called the *first law of thermodynamics* and is merely a manifestation of energy conservation. This principle is a cornerstone concept and is like a tenet of nature. The "other forms of energy" in Eq. (1.12) encompasses electrical, chemical etc. However, looking at these processes at the atomic scale, they are all of the microscopic or macroscopic kind.

Suppose we have an electric heater that is drawing 1 kW power. The heater is simply a coil of resistive wire like nichrome connected to a power supply. Since the nichrome wire has high resistance, the mobile electrons and atoms frequently collide with each other. As a result, the vibratory motion of the atoms inside the nichrome wire increases and consequently the internal energy of the wire also increases.

Suppose that we have reached a heater is operating such that the temperature T remains constant. Furthermore, nothing moves or expands or contracts: there is no macroscopic work involved. Since the wire has a constant temperature, there will be no change in E_{int}.

$$\Delta E_{\text{int}} = 0,$$

In a time period of 1 second, involving the first law of thermodynamics,

$$\Delta E_{\text{int}} = Q + W + 1\text{ kW} \times 1\text{s},$$
$$0 = Q + 0 + 1\text{kW} \times 1\text{s}$$

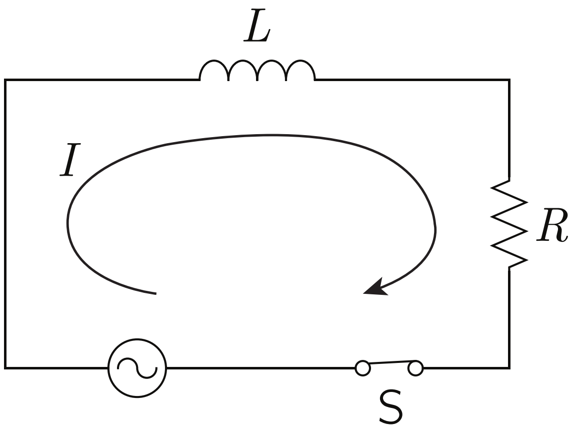


Figure 1.17: Electric Heater

resulting in the microscopic energy transfer $Q = -1kJ$, meaning that all the electrical energy that enters the nichrome wire is transferred out of the heater through microscopic work only. The energy is radiated out to heat up the room.

Let's plot the temperature of the wire against time and discuss how the temperature T varies with time and reaches a steady value T in house hold heaters.

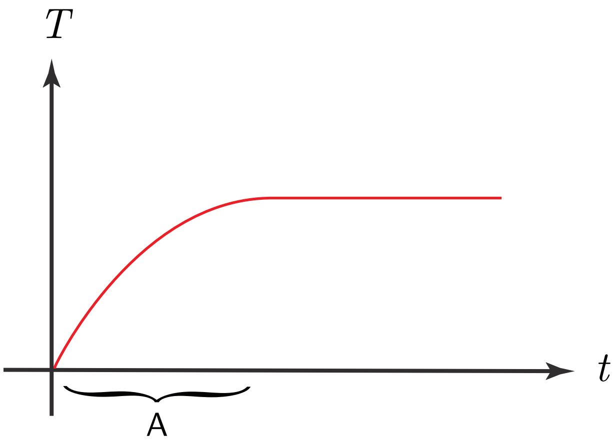


Figure 1.18: Temperature-time curve of an electric heater

When we plug the heater into the mains, the temperature of the heater starts from room temperature T_{room} and increases until it reaches a steady value T_{steady}. If the electrical energy were to keep on flowing, E_{int} would perpetually increase. Soon we will have unbearably hot temperatures in the room - a sun in our vicinity! Fortunately for us, the temperature stabilizes because energy is transferred to the environment due to the large temperature gradient between the nichrome wire and the atmosphere. Although the internal energy and thermal energy also increase in this region (region A in Fig. (1.18)), but once the whole environment (heater+ room) reaches a steady temperature T_{steady}, the internal energy remains constant.

Chapter 2

Quantization of Internal energy

The previous chapter introduced the concept of internal energy. One of the most important advances of the twentieth century physics is the quantization of energy. This allows one to associate energy levels inside atomic or molecular systems. The current chapter will slowly lay the foundations of this concept. Quantization of energy just like conservation can be regarded as a cornerstone principle. However, before discussing the quantization of internal energy, we need to introduce the particle called the photon. Where do photons come from? Photons are important because their energies truly reveal the quantization of energy i.e energy comes in discrete packets.

2.1 Photon

A photon is a packet of energy.The amount of energy is determined by its frequency f. The energy of the photonic packet is

$$E_{photon} = hf, \tag{2.1}$$

where $h = 6.63x10^-34Js$ is called the Planck constant.

In vacuum, this packet moves with the phenomenal speed of light $c = 2.99795x10^8 ms^{-1}$. It is not possible to change the speed of a photon while it freely propagates in vacuum.

When moving with a speed c, the photon must also have linear momentum given by

$$p_{photon} = \frac{hf}{c}. \tag{2.2}$$

This is easy to remember if we write $p_{photon}c = hf$. In classical physics, the momentum of any object is the product of its mass and its speed. Therefore, the momentum of photon can also be written as

$$p_{photon} = mc. \tag{2.3}$$

Comparing Eq. (2.2) and Eq. (2.3), we obtain $hf = mc^2 = E_{photon}$ which is Einstein's famous energy-mass equation. The mass m of the photon becomes

$$m = \frac{hf}{c^2}. \tag{2.4}$$

This is called the relativistic mass of the photon. A photon therefore moves with the speed c in vacuum and has a mass $\frac{hf}{c^2}$. This mass and speed mc qualifies as its momentum and can exert a force on objects. This force (x area) is sometimes called the radiation pressure.

How are photons affected by gravity?

The earth is a massive object and it attracts any other ??? object. Photons are no exception. The mass of the photon manifests in various interesting ways.

Hold a ball and throw it against an opposite facing wall. What happens? Unless the ball is filled with helium, it is pulled downward due to Earth's gravity and eventually drops. In a similar fashion, consider we shine a laser to the wall. If the earth were heavy enough to attract photons and the effective mass of the photon were appreciable, we should actually see the beam of light falling back to the earth just like a projectile falling down to the ground. In fact, the extreme case of pulling a photon happens near a black hole. The black hole is such a massive object that light falls into it very easily and cannot escape it.

Consider a star far away from our sun. When the light from this star moves out, which means photons are being emitted by the star, they enter our solar system and come closer to the sun. Now the sun has mass and the photons have relativistic mass too, so the beam of light is bent by the sun and as it reaches us, we see only the star's apparent position. The sun acts like a lens which bends light and the phenomenon is, quite aptly, called gravitational lensing.

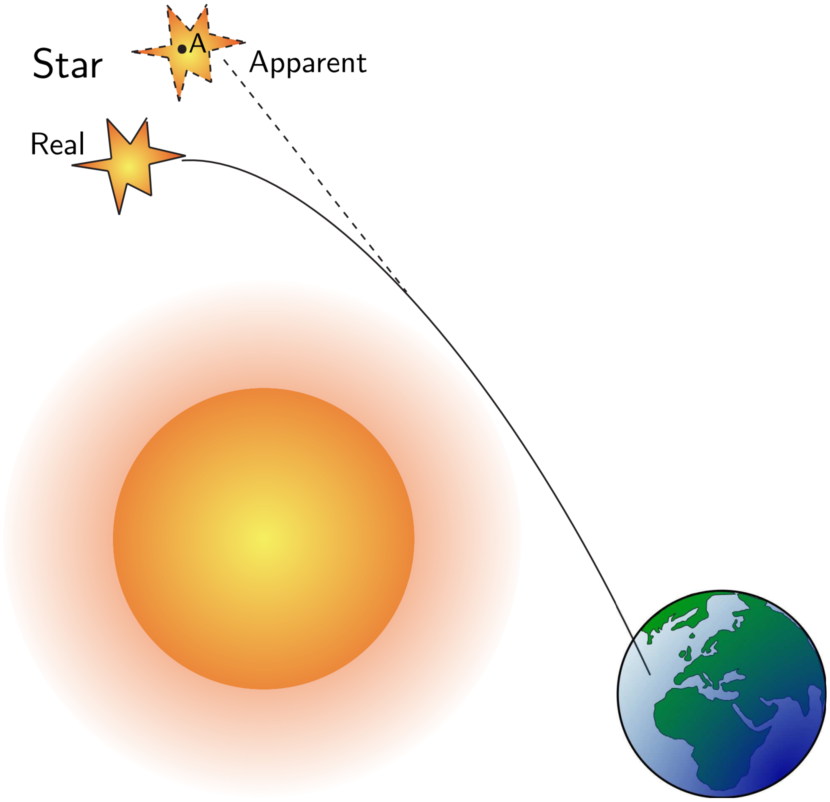


Figure 2.1: Gravitational lensing

2.1.1 Where do photons come from?

Photons may come in streams, one after the other. If the flux is appreciable, we call this stream simply light or radiation. Radiation has a wave property associated to it, which defines a frequency f and wavelength $\lambda = \frac{c}{f}$. Hence, the energy of photon in terms of λ is

$$E_{\text{photon}} = \frac{hc}{\lambda}. \tag{2.5}$$

Photons thus carry energy and are emitted by objects. These emitters could be atoms, molecules, solids and so on. Since the energy of a system as a whole must be conserved, the emitter of photons must lose its energy in the emission process. So, if an emitter is firing out photons of energy $hf = \frac{hc}{\lambda}$, the energy of the emitter must continually decrease in steps of hf to conserve energy. After n emissions, the number of emitted energy would have gone down by nhf.

The energy harbored by the photon may stem from either the internal energy

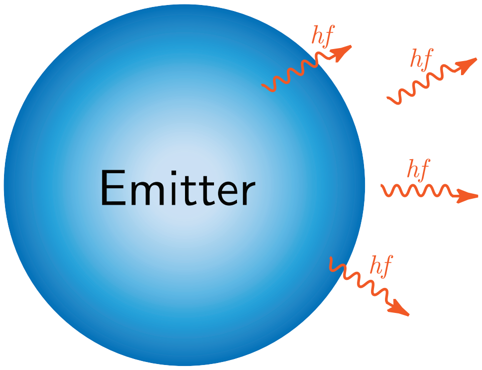


Figure 2.2: Emission of photons

or the center of mass energy ???. Eventually, with time the energy of the emitter diminishes. If we consider photons emitted by a hot object, a process called thermal radiation, the object is sent to ??? with time. The stepwise drop in energy hints that the (internal) energy may be quantized.

2.1.2 An Electron inside a Quantum Dot

The advent of nanotechnology has given us countless new atomic toys to play with. An example is a quantum dot which can be considered as an artificial atom, which can harbor only a few electrons, possibly even one. Suppose we have a single electron in the quantum dot. The electron will have some energy and the system, in turn, will have some internal energy. According to quantum mechanics, the electron can only have precise values of energy due to its confinement inside the dot.

Let's have a look at the potential-energy curve of this electron (Fig. 2.3). Here, r is the distance between the electron and the center of atom. As r approaches infinity, the potential energy $U(r)$ approaches zero. When the atom is ionized, the electron escapes the atom and is free to wander anywhere. When the electron is not within the attraction of nucleus, its energy will be zero and when it is inside the atom, its energy will be smaller than zero. In this curve, the straight green line drawn on the potential-energy curve AB represents one choice of the total energy. If the electron is oscillating back and forth (just like in the harmonic oscillator) between two points A and B, its E_{int} is constant and fixed at this green line even though kinetic and potential energies are interchanging. At points A

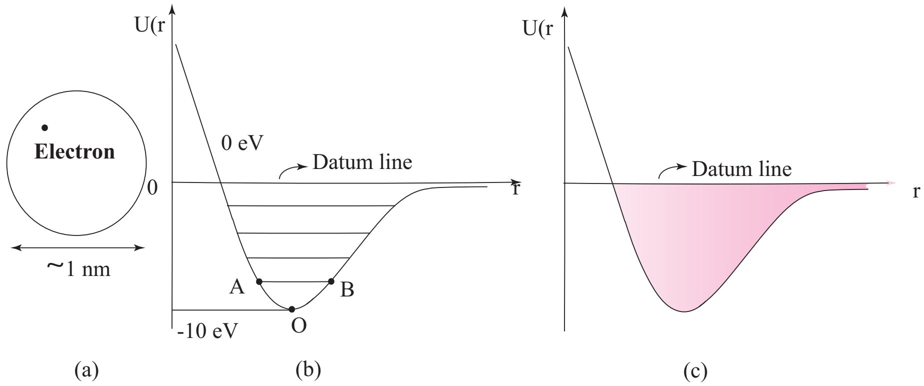


Figure 2.3: A quantum dot with a trapped electron

and B, all of its energy is solely potential whereas at point O, the energy is purely kinetic. We can draw other straight lines on the potential-energy curve. In fact, we can draw an infinite number of straight lines in this region of negative energy up to the *datum line.* This datum line represents a boundary between bound electrons and free electrons.

If we view the quantum dot through classical mechanics, we can draw a shaded region instead of drawing infinite number of lines in bound region because all these energies are possible. This is the classical picture: the energy of the electron inside the atom is continuous and can take any value.

However, this notion really contradicts the quantum picture, according to which the energy of a bound electron cannot be continuous. *We can only have discrete (quantized) energy levels in a bound region.*

On the other hand, the quantum mechanical picture, while appearing to be daunting, is no more complex than a child taking a slide at a park. A child climbs up the stairs of a slide through discrete steps and comes back down sliding. While climbing, the child can only step on individual steps of stairs, he can only have discrete values of height and hence discrete values of potential energy. On the other hand, no such restriction is present while the child slides down to the ground. Therefore, while his upward motion is discrete, his downward motion is continuous.

In a similar fashion, we can imagine the energy of an electron to be like that of the child: its energy is quantized when it is bound and continuous when it is unbound.

In the potential energy curve, we can represent the energy levels in bound regime with discrete values E_1, E_2,..... and so on (Fig. 2.5). We cannot

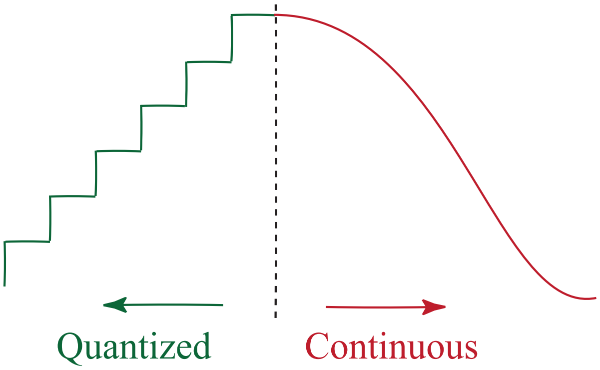


Figure 2.4: Playground as an example of discrete and continuous motion.

have any energy value in between the two consecutive energy levels. As electron escapes nuclear attraction above the datum line, it does not have any potential energy and only kinetic energy which can take up any value.

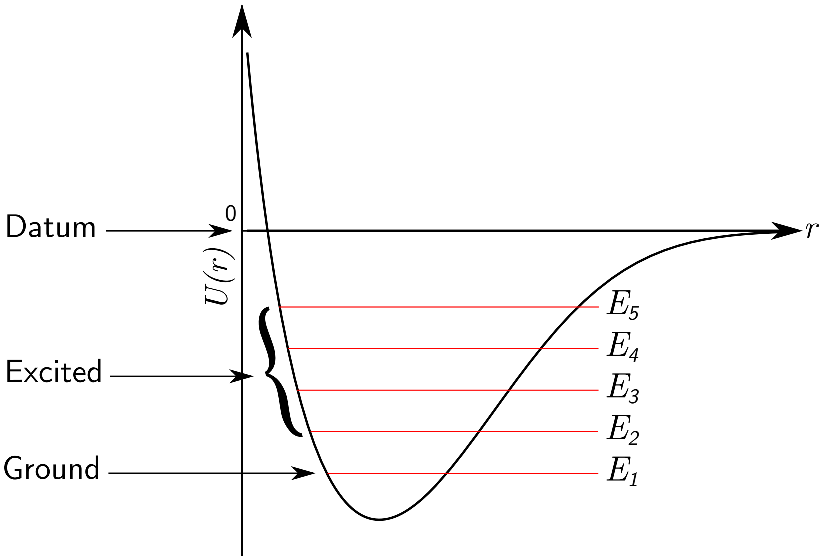


Figure 2.5: Energy levels in a quantum dot

Suppose there is only one electron in the quantum dot. It can exist in any of these five discrete energy levels with energies E_1, E_2, E_3, E_4 and E_5. Let's consider the electron to be residing in energy level E_3. As E_3 is an excited state. It cannot stay there forever. It can either make a transition to E_2 or it could fall directly to E_1. As it makes a transition from E_3 to E_2, E_{int} of the system decreases. However, the center of mass energy of the system does not change because the whole system is not moving. This decrease in energy is accompanied with the emission of a photon such that the total energy of the system remains conserved. The energy of this photon will be equal to the difference between E_2 and E_3

$$E_{\text{photon}} = E_3 - E_2 = hf.$$

The frequency of the emitted photon depends on the energy difference between two quantized levels.

Now, consider four quantum dots in our system (Fig. (2.6)). An electron is present in the first excited state in each quantum dot. They can either fall down together to the ground level or at different times. In any case, four photons will be emitted, each with an energy around 1 eV. These photons can be detected by a detector. By comparing the four-quantum-dot system with two-quantum-dot system, we can easily see that the four-quantum-dot system may produce a higher intensity of photons for the same kinds of transitions. This also shows that intensity (or brightness) depends directly on the number of photons emitted, but higher brightness does not mean higher energy per photon. Each photon has the same energy hf. Quantized energy levels lead to the concept of photons carrying discrete packs of energy.

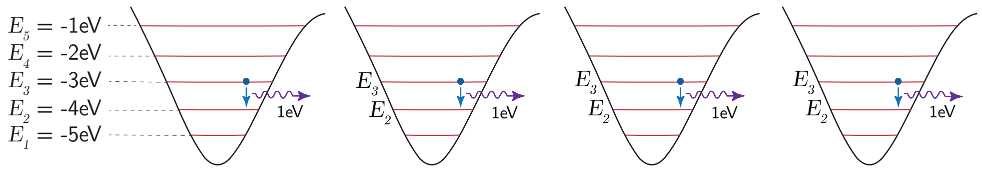


Figure 2.6: A four-quantum-dot system

2.2 The Electromagnetic Spectrum

The electromagnetic spectrum (EM) covers a wide range of wavelengths and photon energies. The EM spectrum is merely a classification of photons based on wavelength or frequency or energy. For example, blue light is a photon with wavelength $\lambda_{\text{blue}} \sim 400$ nm and red light is a photon of wavelength $\lambda_{\text{red}} \sim 630$ nm. Similarly, $\lambda_{\text{green}} \sim 550$ nm. There are no white photons, but if a large number of photons of different frequencies are correctly mixed in the right proportion, white light can be obtained.

The visible range in the EM spectrum is exceedingly small. The photons of the observable frequency enter the eye and strike the retina. The retina, in turn, is connected to the optical nerve that carries the information to the brain. Our brain processes these photons and we see different colors as a result. The eye is a very sensitive device, as only 30 to 40 photons are enough to register a signal. This limit, however, is being pushed by some recent work (REF TO WORK) which shows that or eye could in fact be sensitive to even single photons.

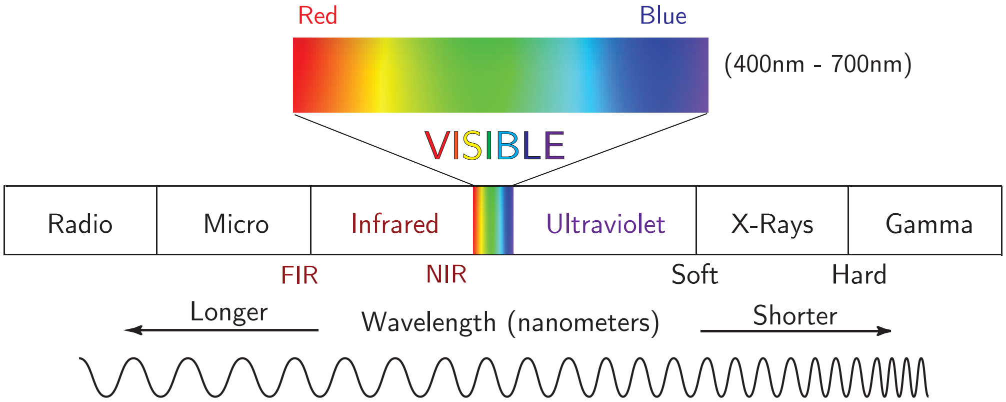


Figure 2.7: The electromagnetic spectrum

Due to biological evolution, the human eye can only detect light in the visible region but not in other regions. For example, We cannot detect infrared photons, but the *camera in the mobile phones* can though certain *animals in marine environment*, like eels and octopuses, can detect infrared photons. *Blind Dolphin* in Indus River cannot detect visible photons, but can detect photons in radio frequency regime. Furthermore, certain *insects* can detect ultraviolet photons.

Fig. (2.7) shows the electromagnetic spectrum in order of increasing frequency (or decreasing wavelength). In the regions of higher frequencies beyond blue, there are different kinds of waves. All of them have photons moving with the same speed and differing only in their frequencies. On the other side of spectrum (on the left of red) lies the infrared regime (near infrared, far infrared, etc), microwaves regime and then radio waves, in decreasing order of frequency. Most telecommunications ??? take place around 1550 nm in the infrared regime.

2.2.1 Types of Spectrums: an eye into quantization

We now have an idea of what a photon is and that quantized levels lead to photons of precise wavelengths. We build upon this concept further. In the seventeenth century, Newton performed an experiment, among countless others. He made a hole in the wall such that a beam of sunlight could pass through it. He then placed a prism in the path of the beam and used a screen to observe the beam coming out of it. He saw different colors visible on the screen. Actually, the prism had resolved the white light into a spectrum of colors. A spectrum is any distribution of photons resolved with respect to wavelength.

What Newton observed was an example of *continuous spectrum*. In fact, every photon emitted has its origin in an atomic transition and carries a discrete value of energy but since the atomic levels are distributed across a wide range, a continuous flux of energies and wavelengths is seen. I will talk more about this later.

Let's modify his experiment a bit. This time we use a tube filled with hydrogen gas and having two electrodes (anode and cathode) connected to a power supply. The remainder of the equipment is the same. We'll place a prism in front of tube and a screen to observe the spectrum.

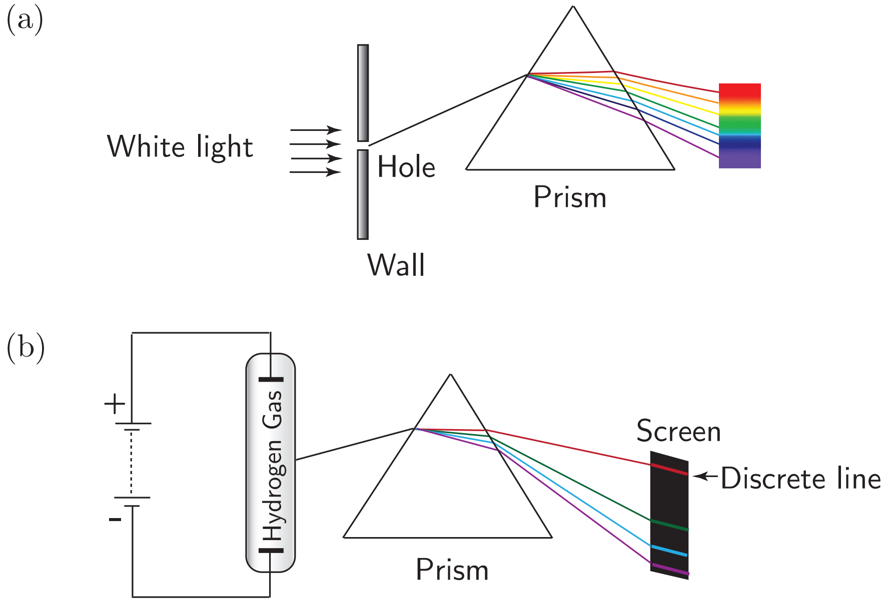


Figure 2.8: Continuous spectrum

Different lines of various colors will be observed on the screen. We will only be able to observe the visible, obtaining discrete lines instead of a continuous spectrum. This is a direct proof of the quantization of energy levels inside two hydrogen.

There is only one electron in hydrogen atom which experiences attraction towards the nucleus. The energy of electron inside the atom is quantized and is given by

$$E_n = -\frac{13.6}{n^2}\ \text{eV} \tag{2.6}$$

where $n = 1, 2, 3 \cdots$ is called a quantum number. Eq. (2.6) is a quantization formula. This is similar to the energy quantization of an electron in a quantum dot. Returning to Fig. (2.9), suppose the quantized energy levels inside the quantum dot obey the formula. Hence, we can formulate its energy as

$$E_n = (-6 + n)\ \text{eV} \tag{2.7}$$

where the quantized levels are labeled by quantum numbers $n = 1, 2, 3, 4, 5$. This will yield an energy level diagram given in Fig. (2.9).

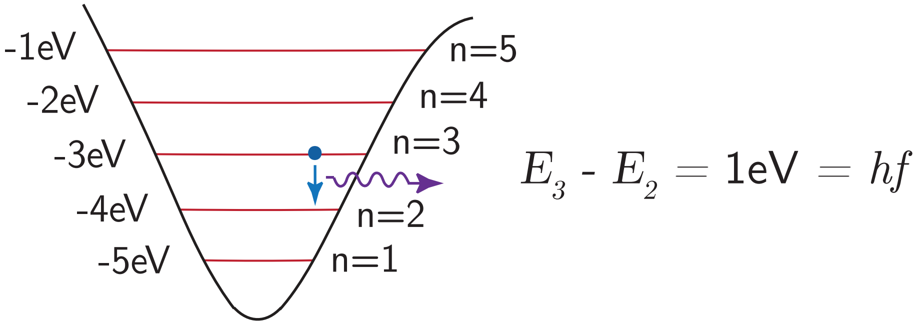


Figure 2.9: Quantized energy levels inside the dot

Let's analyze a similar energy level diagram of hydrogen atom.

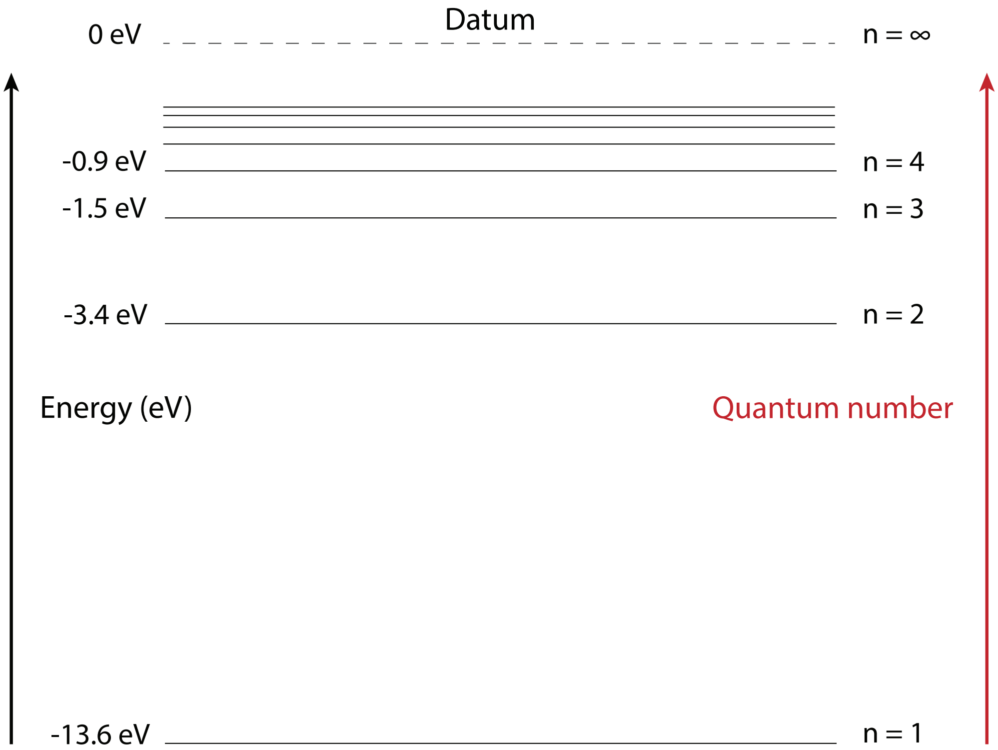


Figure 2.10: Energy level diagram of hydrogen atom

The ground level energy is given by

$$E_1 = -\frac{13.6}{1}\ \text{eV} = -13.6\ \text{eV}, \tag{2.8}$$

and then excited energy levels have energies

$$E_2 = -\frac{13.6}{2^2}\ \text{eV} = -3.4\ \text{eV} \qquad \text{,and} \tag{2.9}$$

$$E_3 = -\frac{13.6}{3^2}\ \text{eV} = -1.5\ \text{eV} \qquad \text{(second excited state)} \tag{2.10}$$

$$E_4 = -\frac{13.6}{4^2}\ \text{eV} = -0.9\ \text{eV} \qquad \text{(third excited state)} \tag{2.11}$$

and so on.

Suppose that initially the hydrogen atom is in its ground state. In this case, it will not emit any photon. Let's excite the hydrogen atom by promoting its electron to a higher level, say E_2, and wait long enough so that this electron eventually comes down to E_1. It does so by emitting a photon having a particular frequency and a particular energy given by

$$E_{\text{photon}} = E_2 - E_1 = 10.2 \text{ eV}. \tag{2.12}$$

We now provide enough energy to the electron to jump to a higher level E_3. If it is in the second excited state E_3, it could fall to E_2 by emitting a photon of energy

$$E_{\text{photon}} = E_3 - E_2 = 1.9 \text{ eV}, \tag{2.13}$$

and then fall back to E_1 by emitting another photon of energy

$$E_{\text{photon}} = E_2 - E_1 = 10.2 \text{ eV}, \tag{2.14}$$

or it could fall down directly to E_1 by emitting a photon of higher energy

$$E_{\text{photon}} = E_3 - E_1 = 12.1 \text{ eV}. \tag{2.15}$$

As a result of one transition, a single photon is emitted. When the electron reaches state E_1 in two steps, light becomes brighter doe to the emission of two photons $E_3 \rightarrow E_2 and E_2 \rightarrow E_1$. Each photon has a characteristic frequency and wavelength. However, the overall energy does not change. These photons of precise energies and frequencies give rise to a line spectrum.

The photon emitted in E_3 to E_2 transition has a wavelength

$$\begin{aligned} \lambda &= \frac{hc}{E_3 - E_2} \\ &\approx 650 \text{ nm}. \end{aligned}$$

Similarly, in the single transition from E_4 to E_2

$$E_{\text{photon}} = E_4 - E_2 = 2.5 \text{ eV}$$

which corresponds to the wavelength 495 nm.

In Fig. (2.10), peak 3 is nearly at $\lambda = 656$ nm, which corresponds to an electron that has jumped up the level E_3 by electrical discharge and then fallen back to the level E_2. While peak 2 is at $\lambda = 470$ nm, it is actually due to a transition that is taking place from E_4 to E_2. The smaller peak 1 is due to the transition from E_5 to E_2. This is the direct proof that the energy levels inside the hydrogen atoms are quantized. The humps in graph are due to background light.

From Fig. (2.10), it is clear that peak 3 which corresponds to the transition from E_3 to E_2 is stronger than peak 2 which corresponds to the transition E_4 to E_2. Actually, since E_3 is closer to the ground state than E_4, it is more populated than E_4. More electrons are available to de-excite in E_2, thus emitting more number of photons. Therefore, the intensity corresponding to $\lambda \sim 656$ nm (E_3 to E_2 transition) will be higher than that corresponding to $\lambda \sim 470$ nm (E_4 to E_2 transition). However, when we increase the temperature of hydrogen atom, more electrons are excited to E_4 and peak 2 will start to increase. Since E_3 is at lower energy, there is a higher probability of the electron to excite to E_3 than to E_4. In transition from any of the higher levels to E_1, more energetic photons are emitted (having shorter wavelengths). The spectrometer, on the other hand, works only in the visible regime.

The photon emitted during the transition from E_2 to E_1 has wavelength

$$\lambda = \frac{hc}{E_2 - E_1} \sim 100 \text{ nm}, \tag{2.16}$$

which is in ultraviolet regime and cannot be detected by this spectrometer. If we had a spectrometer that could detect light beyond the shorter frequencies, we would have seen this stronger peak as well.

Modification of Newton's experiment

Another experiment is performed using an incandescent tungsten lamp as a source of white light and continuous spectrum is observed. This is Newton's experiment, with the only difference that we have an incandescent lamp instead of sunlight. The bulb is emitting photons of different frequencies. As there are different kinds of atoms in the incandescent lamp, each atom is emitting its characteristic light spectrum. Hence, this continuous spectrum has almost all these wavelengths.

We can filter the light of desired wavelengths using different filters. Suppose we place a filter in front of an incandescent lamp which allows light of wavelength $\lambda \sim 530$nm (green light) to pass and blocks all other photons. Hence, it lets photons between 520 nm and 540 nm to pass.

Spectrum of LEDs

In this experiment, four kinds of LEDs are used: red, white, blue, and green. Suppose we turn on the blue LED having a wavelength $\lambda \sim 490$ nm. If we put a filter of wavelength $\lambda \sim 430$ nm, the peaks for the rest of the wavelengths (including blue) disappears. This verifies that the blue LED is indeed emitting photons of wavelength $\lambda \sim 485$ nm.

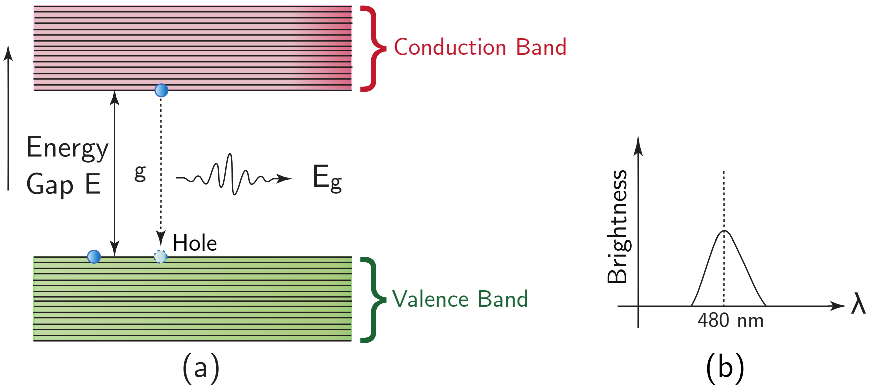


Figure 2.11: (a) Blue LED spectrum (b) Blue LED spectrum with filter

The band gap of blue LED is E_g. If we apply a voltage equal to E_g/e to the electron in the valence band, it will go to a quantized level in the conduction band. As it leaves the valence band, a *hole* is created. The electron has to eventually come back to its origin (hole) by emitting a photon of energy $E_{photon} = E_g$ and frequency $f = E_g/h$. The LED is made up of such a material that the gap corresponds to blue color.

If a higher voltage is applied, the electron will excite to a higher level in the conduction band, and will jumps back by emitting a photon of different frequency. Hence, all of the photons are not precisely of the frequency $f = E_g/h$ and we don't obtain a straight line in the spectrum of blue LED.

Line spectra of various mediums

The spectra of various mediums is shown in Fig. (2.12). Hydrogen has only a single electron so it can jump to a few different energy levels (which are quantized, as seen by the separate peaks). Compared to this, a mercury-filled tube has a more complicated line spectrum due to its many electrons which can jump to many different energy levels. This creates many combinations for the electrons' path. The same could be said about a neon-filled tube.

On the other hand, the bulb and tubelight spectra show a continuous graph of wavelengths.

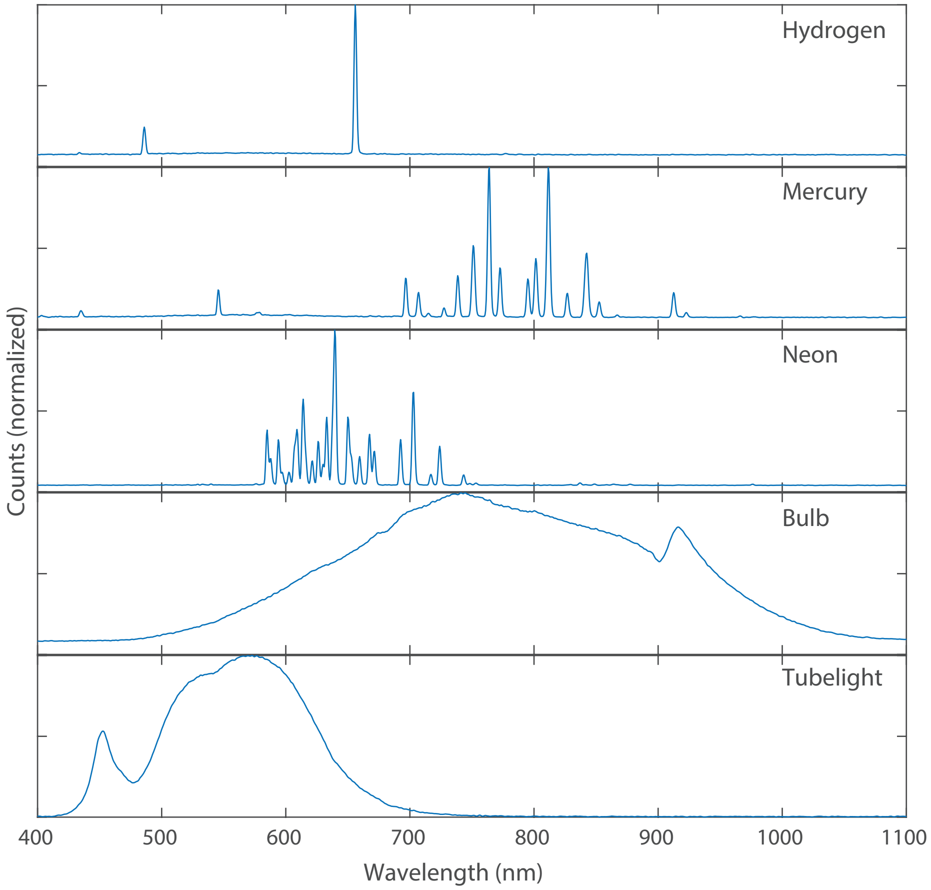


Figure 2.12: Hydrogen, mercury, neon, bulb and tubelight spectrum

2.3 Absorption and emission of photons

2.3.1 Absorption and spontaneous emission

Consider a quantum dot comprising of three discrete energy levels labeled E_1, E_2 and E_3. The electron can exist in any of the three levels. Suppose that an electron is in the ground state E_1 and a photon of just the correct amount of energy $E_{\text{photon}} = E_2 - E_1$ comes in. The electron will be excited to the first excited state E_2. This process is called **absorption**, and is shown in Fig. (2.13).

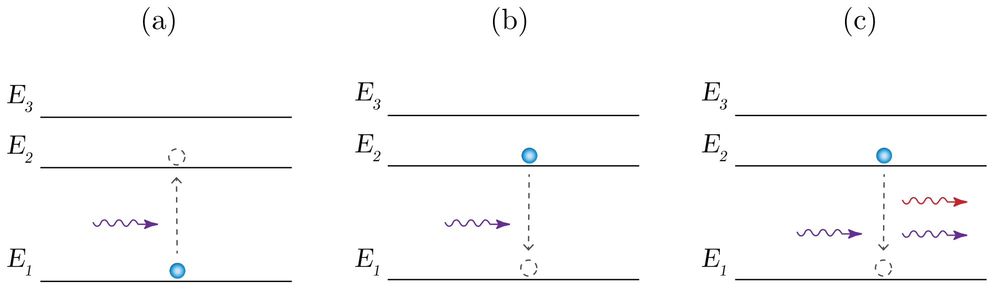


Figure 2.13: (a)Absorption, (b) spontaneous emission and (c) stimulated emission.

Note that if the frequency of the incoming photon does not match either of the differences $E_3 - E_1$ or $E_2 - E_1$, it will pass straight through the quantum dot. Such a photon cannot be absorbed and no transition will takes place. The transparency of the quantum dot depends on the incoming photon's frequency. A few specific examples might help.

- If the incoming photon has energy $E_{\text{photon}} = E_3 - E_2$ and the system is in E_1 in the three-level system, the photon will pass straight through (Fig. 2.14(a)).
- A photon having energy $E_{\text{photon}} < (E_2 - E_1)$ cannot excite the electron from E_1 to E_2 as there is no level below E_2. The electron does not get promoted and will remain in the lower state with the specific energy E_1 (Fig. 2.14(b)).

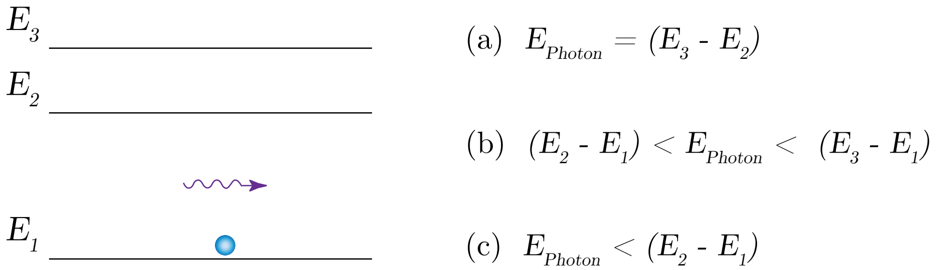


Figure 2.14: Three examples when atom is transparent to the photons

- Similarly, for a photon of energy $(E_2 - E_1) < E_{\text{photon}} < (E_3 - E_1)$ no transition will take place. No doubt, it has sufficient energy to excite the electron to level E_2, but the electron will not be excited

> as there is no level between E_2 and E_3 that can serve as a receptacle for the excited electron. Actually, photon is a particle which can only be destroyed while in vacuum, it can neither share its energy with anything else nor can it split into two particles wherein one daughter photon is providing enough energy to excite to level E_2 and other daughter photon carries away the remainder of energy (Fig. 2.14(c)).

However, if through some process, the electron has been excited to the state E_3 it will stay there on average for a certain amount of time (roughly of the order of nanoseconds) and will decay to the ground state E_1, emitting a photon of energy $(E_2 - E_1)$, Fig. (2.13(b)). This process is called **spontaneous emission**. Hence, the excited state lifetime can be around $\tau \approx 10^{-9}$s.

Before Einstein, only two processes were known; absorption and spontaneous emission. A third process namely **stimulated emission** was presented by Einstein and forms the basis of lasers.

2.3.2 Stimulated emission and lasers

Suppose the electron is in a state E_2 and a photon of energy $E_{\text{photon}} = (E_2 - E_1)$ comes in and stimulates the downward transition from E_2 to E_1, Fig. (2.13(c)). To conserve the energy of atom, a photon with energy $(E_2 - E_1)$ must emit in downward transition of electron from E_2 to E_1, to balance the decrease in internal energy of this atom. So the incident photon will pass straight through, after doing the business of stimulating the emission of a photon. In this process, we input one photon of the correct energy and two photons come out. One is the incident photon and the second is a photon whose emission has been triggered by the incident photon. This process is called **stimulated emission**. The wonderful thing that is happening in this process is that these two photons are moving in same direction and are in phase. They are *absolutely coherent*.
Stimulated emission is really an instantaneous process, as soon as the photons comes in and interact with the system, other photons are emitted. The two photons have an associated wave that is exactly in phase, totally synchronous, so even there could be a delay but the phase could actually match itself (Fig.2.13(c)).

In summary, three processes can happen, when an electron inside an atom or a quantum dot sees a incoming photon. Each process occurs with specific rates and probability.
In fact, spontaneous emission can be received as stimulated emission triggered by photons coming from nowhere, i.e vacuum. Huh! Strange isn't it.

Vacuum has some residual, zero-point energy which is manifest by virtual vacuum photons. So the major difference between spontaneous and stimulated emission, is that the two photons coming out in the stimulated emission process are phase coherent.

2.4 Lasers

Now that we know about the types of emission, we are ready to take our discussion towards lasers. The word **LASER** means Light Amplification by Stimulated Emission of Radiation. This abbreviation gives us an outline for our discussion on lasers in this section. All we need to make a laser is a system that takes in light and amplifies it using stimulated emission.

2.4.1 Two-level system

Let us begin by analyzing a system with only two energy levels. The lower energy level has an electron population of N_1 electrons, and the higher energy level has an electron population of N_2 electrons.

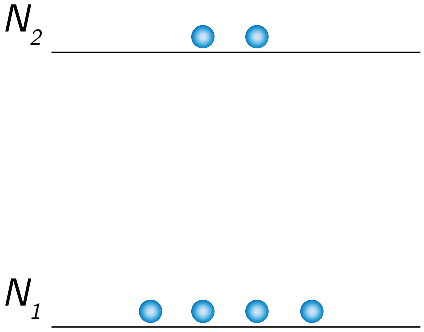


Figure 2.15: Two level system

Electrons in the lower level are photon absorbers while the electrons in the upper level are electron emitters as we have seen in the previous sections. We know that under normal circumstances ,when everything is at equilibrium, electron population in energy levels follows Boltzmann distribution i.e. there are more electrons in the lower level (absorbers) than in the higher level (emitters). This means that the number of photons absorbed is greater than the number of photons absorbed. In other words, the rate of emission and absorption is proportional to the electron population in energy levels.

$$\frac{R_{Emission}}{R_{Absorption}} = \frac{N_2}{N_1} \approx e^{\frac{-\Delta E}{K_B T}} < 1 \tag{2.17}$$

This is the opposite to what we need to make a laser. Hence, to make a laser, we somehow need to invert this general trend of electron population and get more electrons in the upper energy level than there are in the lower energy level. We need to get *population inversion.*

$$\frac{N_2}{N_1} > 1 \tag{2.18}$$

The most obvious solution might be to simply pump in energy corresponding to the energy gap. However, there are a couple of complications with that

- **Spontaneous emission:** As we pump in energy, the number of electrons in the higher energy level will increase. These electrons will tend to go back to the lower energy state. Hence, making this state difficult to maintain.
- **Simulated emission:** As an electron transitions to the lower level spontaneously, the emitted photon will trigger stimulated emission of other electrons present in the higher energy level; this would make the situation even less favorable for us.

Therefore, the best we can achieve by just pumping in energy (even if we pump in an infinite amount of energy), is an equal number of electrons in both energy levels. It is a limit imposed by the laws of nature.

Observing the expression of the Boltzmann distribution, one might suggest that another way to make $\frac{N_2}{N_1} > 1$ is to somehow make the power of exponential $(\frac{-E}{K_B T})$ a positive number. Since ΔE and K_B are constants, the only way to make that happen is to get negative temperatures. That, yet again, is not permitted by nature. In conclusion, population inversion cannot be achieved in a two-level system.

2.4.2 Three level system laser

The time for which an electron stays in an excited state, due to the uncertainty principle, is known as the *lifetime* of that excited state (denoted by the Greek symbol τ). Let's say, for the two level system we analyzed in the previous section, the lifetime of the excited electrons in N_2 was of the order 10^{-15}s. We now introduce another energy level having less energy than N_2, with N_3 number of electrons. This energy level is more stable than the N_2 level and has a lifetime of 10^-10s (10^5 times longer than the N_2 level). These finite lifetimes give us finite line width in the energy spectrum. The longer the lifetime, the sharper the peak.

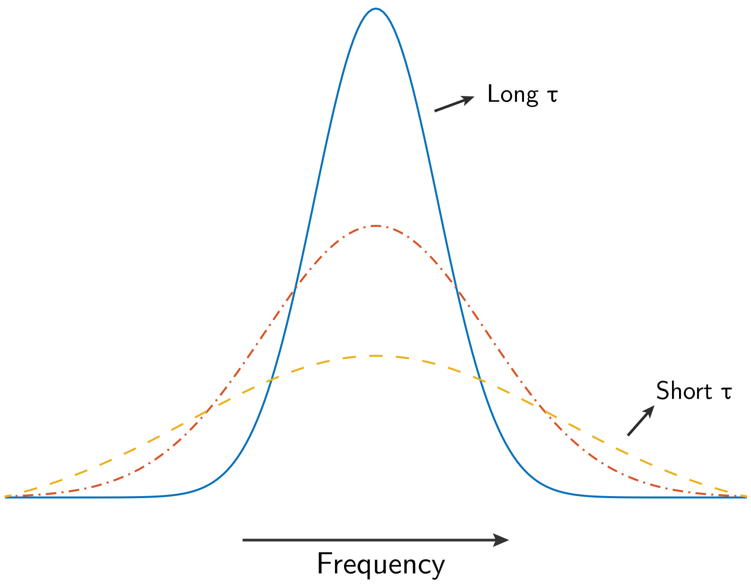


Figure 2.16: Lifetime of different levels

Such a three level system is very hard to make ourselves but, luckily, nature already gave us one such system in the form of Ruby crystals (the gemstone). A Ruby crystal is just an Alumina crystal Al_2O_3 with one of its Aluminum atoms replaced with a Chromium atom $Cr : AlO_3$. The Chromate ion Cr^{+3} causes a defect in the Alumina crystal and makes it have a three-level energy system. One of the first lasers ever made were made using Ruby crystals.

We start off by pumping in photons using a Xeon flash lamp (just like the one used in cameras) (Fig. (2.18)). The photons that we provide, excite the electrons the ground level to the N_2 level just like before. Level N_2 quickly starts to lose these electrons because it has a very short lifetime (10^{-15}s). But this time, when the electrons de-excite, they move to level N_3. The transition from N_2 to N_3 releases photons but they are all of very long wavelengths, giving a very broad peak in the energy spectrum. Since the level N_3 it has a lot longer lifetime than N_2, it starts hoarding these electrons and makes a *meta-stable* state.

After some time, we see that, because the electrons are getting stuck in the

Figure 2.17: Three level system

meta-stable state, the electrons are leaving N_1 faster than they are returning. This results in a greater number of electrons in N_3 (higher energy level) than there are in N_1 (lower energy level). In short, we have achieved population inversion by tricking nature using a meta-stable state that was provided to us by nature itself. For Ruby, when an electron transitions from N_3 to N_1, it will release a photon of wavelength 693 nm, and these photons will constitute the laser light. Until now, we have only seen spontaneous emission and it is not robust enough to make a laser. We will need stimulated emission to kick-in for that.

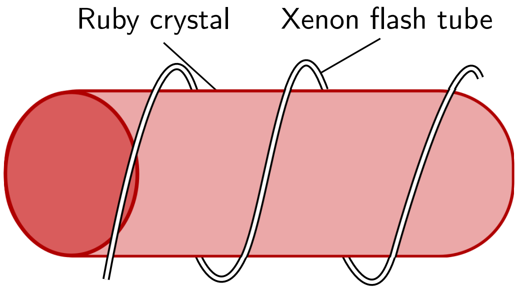


Figure 2.18: Ruby crystal

In a Ruby crystal, there isn't a sole Cr^{+3} ion making a three-level system, rather an entire ensemble of them. When one of the electrons decays and releases a photon, this photon will act as a trigger to stimulate another photon of the same frequency. Now these two photons will trigger the emission of two new photons and the effect will go on; this is known as *the snowball effect* (Fig. (2.20)). We now have an avalanche of coherent photons inside the crystal, all synced up. It is like a giant wave-function of many photons. This is a macroscopic quantum phenomenon; millions of photons moving in the same quantum state. For an instance of time, we have

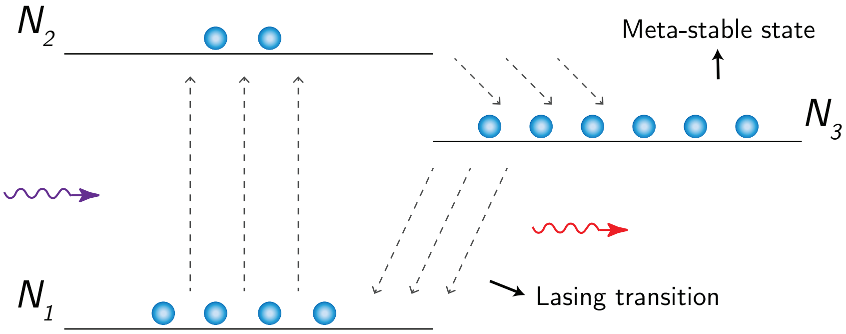


Figure 2.19: Population inversion

more emission than absorption of radiation (photons) and we have achieved that using stimulated emission. The only thing left is to throw these photons out of the crystal in the form of a beam. This completes our definition of a Laser that we defined in the beginning of this section.

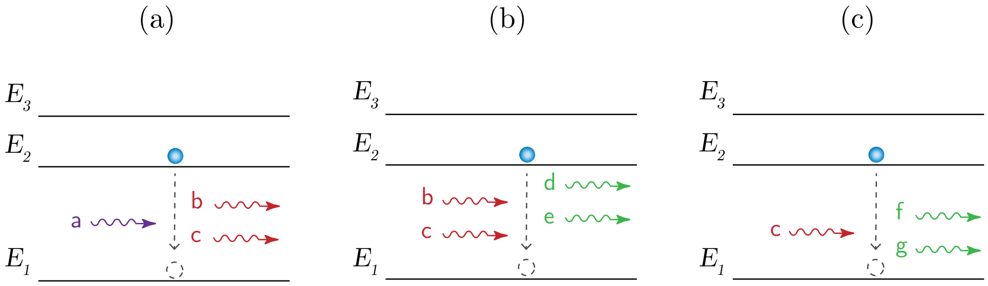


Figure 2.20: A snow ball: Effect of coherent photons

2.4.3 Extraction of produced photons

The medium that emits the photons for the laser is called a lasing medium. We enclose this medium with two perfectly reflecting mirrors. The coherent wave of photons reflects from these mirrors and makes many trips back and forth between them. While doing so, the photons will interfere with one another and make standing waves inside this cavity. To sustain this wavelength of light we need constructive interference, which is only possible if the length of the cavity L is an integral multiple of half the wavelength, i.e $L = n\frac{\lambda}{2}$

We get many many peaks inside this cavity corresponding to different wavelengths (Fig. (2.21)). Some of them get amplified, while the rest get absorbed.

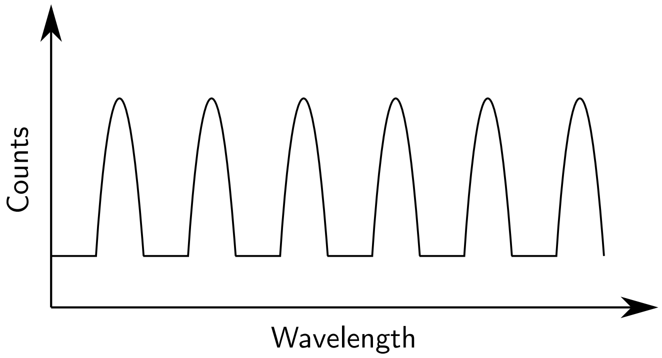


Figure 2.21: Cavity peaks

The giant wave of photons is oscillating inside the cavity and all that we need now is to get it out in the form of a beam of light. For this purpose, instead of using two perfectly mirrors, we replace one of the mirrors with a partially reflecting mirror (Fig. (2.22)). Some of the coherent light will escape the cavity in form of a narrow beam of light (though with some divergence).

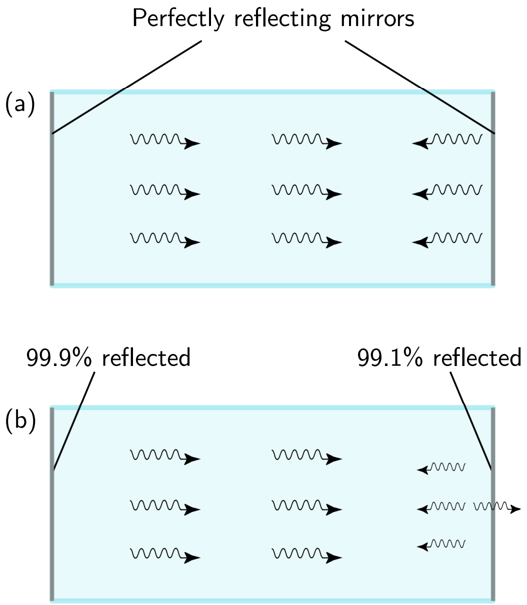


Figure 2.22: a) Photons in a perfectly reflecting glass tube b) Photons in a partially reflecting glass tube

2.4.4 Helium-Neon Laser

Another type of laser is a Helium-Neon laser. It is made by trapping Helium and Neon gases inside a vacuum tube. In this case, we get a four-level energy system unlike the three-level energy system of Ruby lasers.

We start this laser by providing energy to the He atom and excite its electrons from N_1 to N_2 ($2p$ energy level of He in this case). Neon happens to have an energy level only slightly lower than the $2p$ energy level of He. When the He and Ne atoms come close, they collide mechanically. This collision transfers electrons from He to Ne by a non-radiative process. The particles that get emitted are phonons instead of photons in this case (for now, you can consider a phonon as a packet of heat).

When the electrons transfer to the Ne atom, they end up in a meta-stable state. Beneath this meta-stable state we have got an empty energy level and when electrons decay to this lower level, lasing transition takes place. From here the electrons move back to their ground state energy level N_1 in the He atom. Throughout this process, all the transitions are fast except the one with a meta-stable state and only that state is responsible for giving the lasing transition.

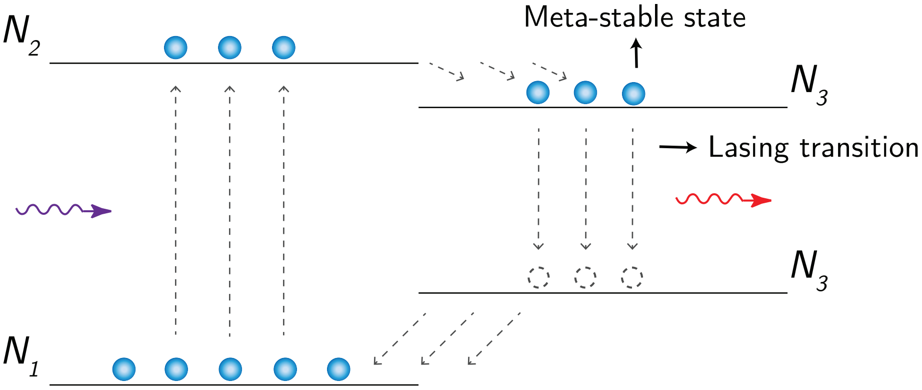


Figure 2.23: Four level system

A $He-Ne$ laser is more efficient than the Ruby laser for the following reasons:

- In this case, the lasing transition is taking place when we go back to a higher energy level which has little or no electrons. It is easy to establish population inversion in this case. Previously, the lasing transition led the electrons to the ground state which was already heavily populated.

- A $He-Ne$ laser has a lower input requirement as compared to a Ruby laser.
- In a Ruby laser, the laser took some time to charge up and we got output in the form of laser pulses only. On the other hand, in a $He-Ne$ laser, we get a continuous wave.

2.4.5 Semiconductor laser

Semiconductor lasers are the type of lasers we often see in our daily lives. To make a semiconductor laser, we combine an n-type semiconductor (with excess electrons) with a p-type semiconductor (with electron deficiencies known as holes) and make a diode (Fig. (2.24)).

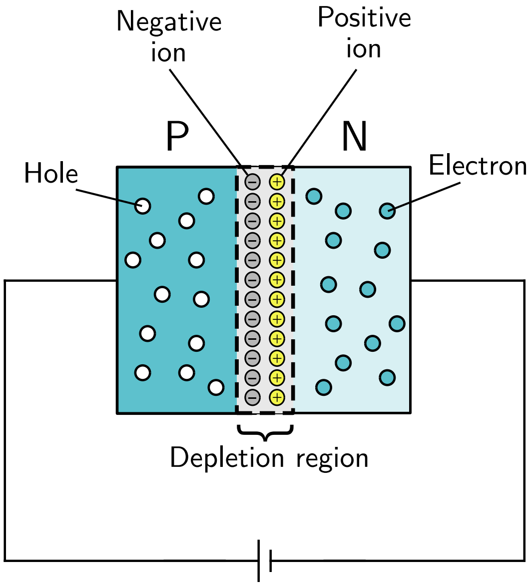


Figure 2.24: Diode

The electrons and holes have a tendency to succumb one another. They diffuse into the junction and recombine, due to which energy is released and there is no overall charge in this region. This junction where the charge gets depleted is known as the *depletion region*.

We then connect a battery to this diode. Negative terminal to the n-type material side and positive terminal to the p-type material side. Battery's positive terminal pushes the holes and the negative terminal pushes the electrons into the depletion region where they combine to release a packet of energy (a photon). At this point the photons released are incoherent and this is just a normal LED (light emitting diode). However, if we supply a large enough voltage (more than the threshold voltage), population inversion can be achieved and we get a laser light (coherent photons) (Fig. (2.25)). To make a light trapping cavity just like before, we can use a perfectly reflecting mirror on one side of the junction and a partially reflecting mirror on the other side.

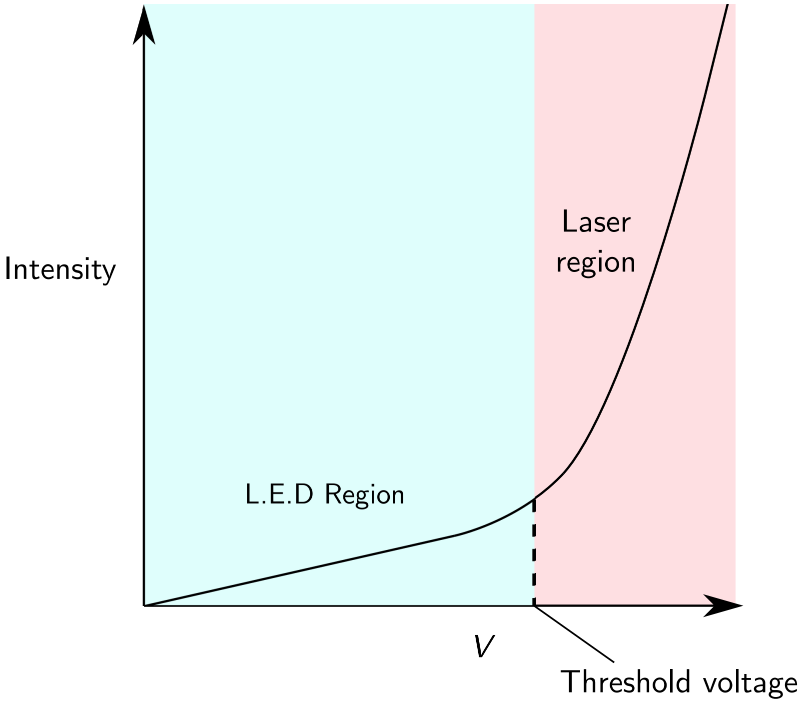


Figure 2.25: Threshold Voltage

2.4.6 Applications of lasers

Lasers are a very handy invention and have tons of applications. Lasers carry power, can trigger fusion reactions and can even cut through metal. In general, lasers are a source of energy and it's interesting to know that this source of energy can also be used to cool things down; to a level that none of the other cooling methods can. Inventing such fascinating applications get you Nobel prizes, just like the one awarded in 1997 to a group of scientists who invented this cooling method.

Laser Cooling

There are a lot of methods to cool things. We can use a refrigerator, put the object in dry ice, use the solvent bath of acetone and isopropanole, and putting salt on ice also causes cooling. Some of the extreme cooling methods we know of are cooling using liquid Nitrogen or liquid Helium. Liquid He can get us to temperatures as low as 4.2K and that was the limit of cooling temperature before laser cooling was invented.

Temperature is related to the average energy of atoms and molecules in a system. The higher the kinetic energy, the higher the temperature and vice versa. In laser cooling, we target these atoms and molecules and drain the energy out of them.

Consider a soup of atoms all moving in different directions. The kinetic energy of an object constrained to move in x-direction is $\frac{1}{2}mv_x^2$. *Equipartition theorem* tells us that this energy, on average, is equal to $\frac{1}{2}k_BT$ and since the atoms in our soup are free to move in all three directions (three degrees of freedom), the total energy will be $3 \times \frac{1}{2}k_BT$.

$$\frac{3}{2}k_BT = \frac{1}{2}mv^2 = \frac{1}{2}m(v_x^2 + v_y^2 + v_z^2) \tag{2.19}$$

We begin by shining a laser light from the x-direction onto gaseous atoms. The frequency of laser light used (f_L) is lower than what is needed (f_o) to cause an electron transition to the higher energy level. However, assuming that the atom is moving towards the laser, the frequency being received by the atom (f') is actually higher than what was supplied. The laser light has Doppler shifted.

$$f' = \frac{f_L}{1 - \frac{v_x}{c}}$$
$$f' = f_L(1 - \frac{v_x}{c})^{-1}$$

Using binomial expansion we get,

$$
\begin{aligned}
f' &= f_L(1 + \frac{v_x}{c}) \\
f' &= (f_o + \Delta f)(1 + \frac{v_x}{c}) \\
f' &= f_o + f_o\frac{v_x}{c} + \Delta f + \Delta f\frac{v_x}{c}
\end{aligned}
$$

In the final term of the last equation we are getting two small variables being multiplied which will turn out to be a negligible value that can be ignored. So our final equation becomes

$$f' = f_o + f_o\frac{v_x}{c} + \Delta f \tag{2.20}$$

The frequency observed by the atom f' must match f_o for the transition to take place. For this purpose, we detune our laser such that we get $\Delta f = -f_o\frac{v_x}{c}$. Substituting this value in eq.(2.20) we get,

$$
\begin{aligned}
f' &= f_o + f_o\frac{v_x}{c} - f_o\frac{v_x}{c} \\
f' &= f_o
\end{aligned}
$$

Now the atom absorbs this photon and goes to an excited state. The absorbed photon had a momentum $-\frac{h}{\lambda}$ and so our atom gets a kick of $\Delta p_x = m\Delta v_x = -\frac{h}{\lambda}$. Following the law of conservation of momentum, the atom slows down(it decelerates).

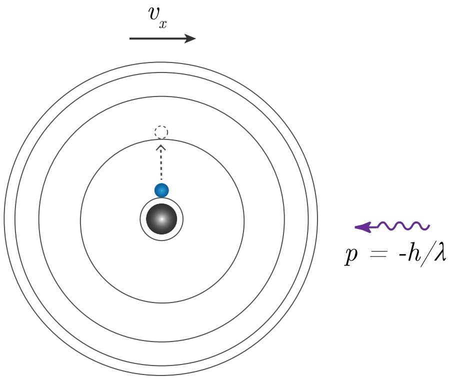


Figure 2.26: Photon absorbed by the atom

After the atom has slowed down, the photon gets released through spontaneous emission. The emission is in a totally random direction and for a large number of atoms, the momentum stays conserved. This emission takes even more energy out of the atom. The atom has lost a tremendous amount of energy and hence gets cooled down. With such laser beams coming from all six directions, the atoms get a backward kick from wherever they move and lose energy. We can achieve temperatures as low as 10μK using this method.

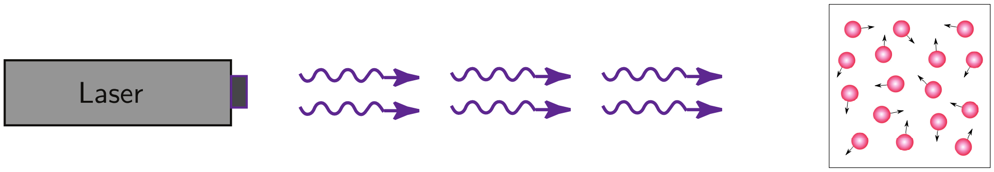


Figure 2.27: Cooling laser

2.4.7 Scattering Processes

Now we discuss how scattering of incoming photon with an atom at rest, may occur obeying the first law of thermodynamics. Consider an atom or a quantum dot which comprises of four energy levels labeled E_1, E_2, E_3, E_4 and furthermore there is only one electron in this atom. There are precise energy gaps between these quantized levels. The atom as a whole could be moving, which means the center of mass could have exhibit certain kind of motion, so it could have a kinetic energy due to its center of mass, but this motion has nothing to do with internal energy. If a photon of precise energy equaling energy gap, $E_{\text{photon}} = (E_2 - E_1)$ negotiates an atom and the electron is indeed present at the lower level E_1, then absorption process will occur. This excites the electron to E_2. Similarly, if the electron is initially in level E_2 and a photon of the exact energy $E_2 - E_1$ hits the atom then stimulated emission may take place resulting in the emission of another photon. If $E_3 - E_2 = E_2 - E_1$, then the incoming photon may either trigger stimulated emission or promote the electron to E_3. There is a probability associated with each process. Atomic physics describes how to calculate these probabilities.

If the electron is initially in level E_3 and a photon of energy $E_{\text{photon}} = (E_2 - E_1) = \hbar\omega$ impinges on this atom, then there could be different possibilities. First of all, the photon will not be absorbed and will pass straight through. No change in internal energy takes place. Secondly, if the atom is at rest then this photon may collide with the atom, lose some of its energy, and could

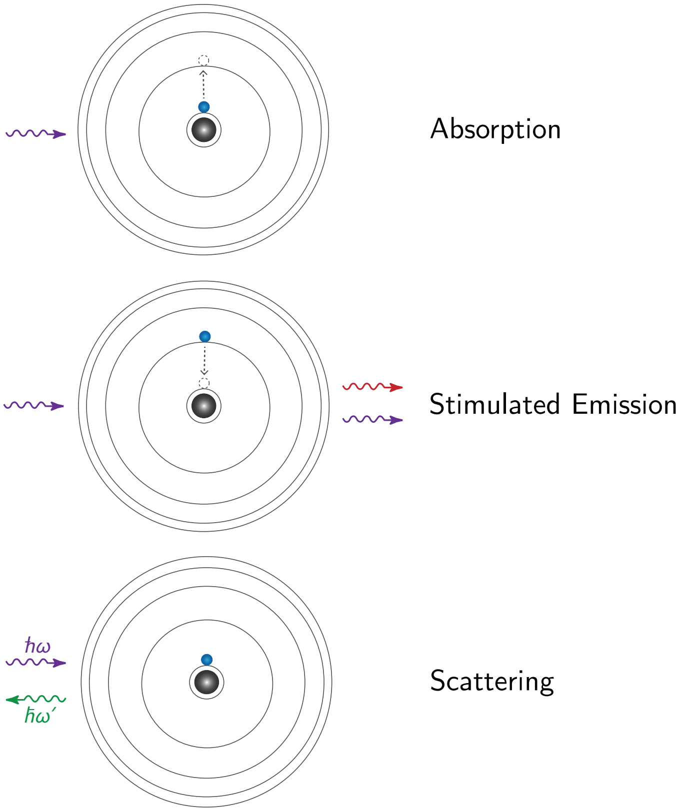


Figure 2.28: (a)Absorption, (b) stimulated emission (c) scattering

be scattered off at a certain angle with new energy $E' = \hbar\omega' < E_{photon}$. So the scattered photon will have frequency ω' smaller than incident one. For example, a blue photon may get scattered off as a red photon, an ultraviolet photon may scatter off as a visible photon, a γ-ray incident on an atom at rest may get scattered as X-ray photon and a microwave photon incident on an atom (at rest) may scatter off as a radio frequency photon.

The remainder of the incident photonic energy goes into the overall kinetic energy of the atom, conserving the total energy. This remainder energy will not change the internal energy of the atom. The atom as a whole will take a kick in a direction, such that the final momentum remains the same as the initial momentum. This means that the scattered photon must have a direction such that it vectorially add to the momentum of atom, to give the same net momentum as before the collision. Such a photon-atom interaction is an elastic collision. On the contrary, when the incident photon has an energy exactly matching an energy gap, a change in internal energy exists ??? This is called inelastic scattering. (This process in which a high energy photon collide elastically with an atom at rest, is named as **Compton scattering**.)

2.4.8 Electron colliding with an atom

Let's now consider the scenario the scenario in which an electron with kinetic energy $\mathrm{K} = \frac{1}{2}mv^2$ (ignoring the relativistic effects), hits an atom. There is only one electron in this atom that can exist in three energy levels i.e., E_1, E_2, E_3.

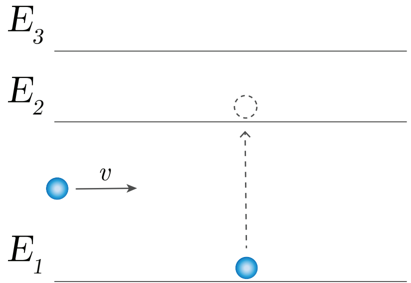


Figure 2.29: An electron hits an atom

Suppose that the kinetic energy of the incident electron is

$$(E_2 - E_1) < \mathrm{K_e} < (E_3 - E_1), \tag{2.21}$$

and the electron in the atom is initially at level E_1. Unlike the photon, the electron is happy in sharing its energy between the internal energy and center of mass kinetic energy of the atom. The incident electron imparts some of its energy with the internal energy, exciting the electron to level E_2, while the remainder of the energy $K_e - (E_2 - E_1)$ is further shared between the kinetic energy of the atom and the electron,

$$K_e - (E_2 - E_1) = K_e' + K_{CM}, \tag{2.22}$$

where K_e' is the kinetic energy of the electron after striking the atom and K_{CM} is the center of mass kinetic energy of atom. Thus, energy and momentum will surely be conserved in this process.

2.5 Quantum harmonic oscillator

A mass attached to a spring is a harmonic oscillator where potential energy $U(x)$ varies quadratically with x, i.e $U(x) = \frac{1}{2}kx^2$ (ADD REF). If this energy is associated with an electron in a ??? confined environment, the energy levels becomes quantized at E_4. We have seen ample examples of these quantized levels. The energy levels are identified by the quantum number n which takes up integer values $0, 1, 2, 3....$ For a quantum harmonic oscillator, these energy levels are all equally spaced. The spacing between any two consecutive levels is equal to $\hbar\omega_0$, where $\hbar = h/2\pi$, h is Plank's constant and and ω_0 is the (natural) frequency of oscillation. For a mass spring system, we have figured out that

$$\omega_0 = \sqrt{\frac{k}{m}}, \tag{2.23}$$

Therefore, if we take the energy of the ground state to be E_0, the first excited state will have an energy $E_1 = E_0 + \hbar\omega_0$, the second excited state will have $E_2 = E_0 + 2\hbar\omega_0$ and so on. Note that E_0 is never equal to zero, it has some finite value $E_0 = \hbar\omega_0/2$. This is the vacuum energy (Section 2.3.2). We can write a relationship between the quantized energy E_n and the quantum number n as

$$E_n = \big(n + \frac{1}{2}\big)\hbar\omega_0 \qquad\qquad \text{with} \;\; n = 0, 1, 2... \tag{2.24}$$

Suppose there is only one electron in energy level E_3. It has to come down to ground state as it cannot stay in excited state forever. So there are different possible ways that an electron can follow to come down. For example, the electron first comes from energy level E_3 to E_2, emitting a photon of energy $E_0 = \hbar\omega_0$, then de-excite from level E_2 to E_1 emitting another photon of the same energy. Finally, it jumps down to ground state E_0 from E_1, emitting another photon of same energy. In this case, three photons in total, each of energy $E_0 = \hbar\omega_0$ are emitted. It may well ??? happen that electron first de-excite from energy level E_3 to E_2 emitting a photon of energy $E_0 = \hbar\omega_0$ and then from E_2 it directly jumps down in E_0, emitting a photon of energy $E_0 = 2\hbar\omega_0$.

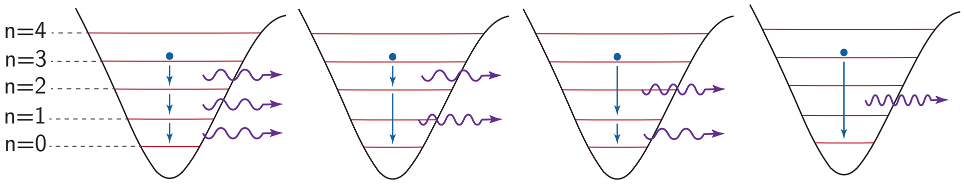


Figure 2.30: Four possible ways of transition from n=3 to n=0 level

Yet a third possibility is that the electron jumps down directly from E_3 to E_1 emitting a photon of energy $E_0 = 2\hbar\omega_0$ and then it falls down into E_0 from E_1 emitting a photon of energy $E_0 = \hbar\omega_0$. Finally, the electron may directly jump from E_3 to E_0 emitting a single photon of energy $E_0 = 3\hbar\omega$. In summary, the atom can deexcite by emitting four, three, two, or a single photon.

We have discussed the modeling of a two-atom system to a classical harmonic oscillator ???. In a similar fashion, we can take a step further and model it to a quantum harmonic oscillator.

Consider Fig. (2.31) wherein r_0 is the most favorable distance between the atoms. In the region close to r_0, we have good approximation of an ideal harmonic oscillator which corresponds to a parabolic energy curve. However, the energy levels are quantized in the true spirit of quantum mechanics. In the harmonic regions, the quantized energy levels will be equally spaced. The potential-energy curve is totally symmetric (parabola) for harmonic oscillator. But as we go to longer inter-nuclear distances, the potential-energy curve becomes asymmetric, and the equal spacing rule breaks down. We now have energy levels that are unequally spaced. They get closer and closer together up to datum line. Above this, the energy becomes continuous, because the atoms are no longer interacting with one another.

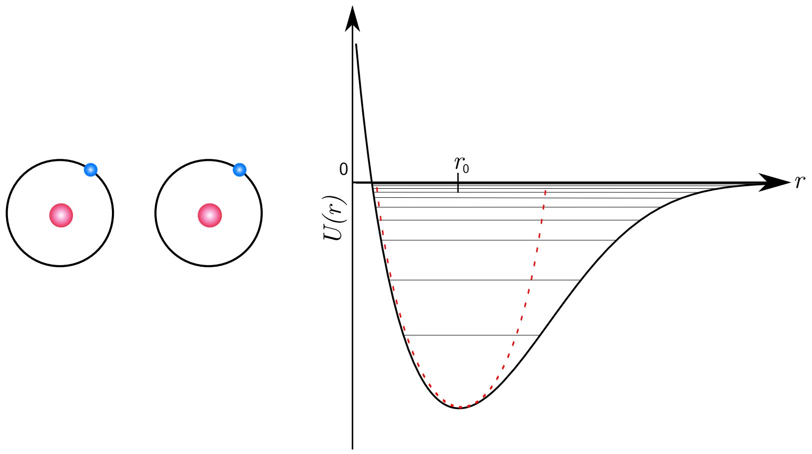


Figure 2.31: Two atoms modeled as quantum harmonic oscillator

2.5.1 Quantized energy levels in the hydrogen atom

A hydrogen atom consists of a nucleus (single proton) of charge $+e$ and an electron with charge $-e$ revolving around it. The force of attraction between electron and proton is given by Coulomb's force

$$\overrightarrow{F} = -\frac{1}{4\pi\varepsilon_0}\frac{e^2}{r^2}. \tag{2.25}$$

The magnitude of this force is

$$\begin{aligned} |\overrightarrow{F}| &= \frac{1}{4\pi\varepsilon_0}\frac{Ze^2}{r^2} \\ &= \frac{ke^2}{r^2} \end{aligned}$$

Looking at this equation, one can deduce that if the electron is very far away from the nucleus ($r \to \infty$), the force of attraction between the electron and the proton becomes zero. Thus, its potential energy is also zero. When the electron is brought close to the nucleus, some work is being done by the nucleus to attract the electron which is stored in the electron as a non-zero potential energy. This potential energy is given by

$$\begin{aligned} |\overrightarrow{U}| &= \frac{kZe^2}{r^2}r \\ &= \frac{kZe^2}{r} \end{aligned} \tag{2.26}$$

However, the electron does not simply 'fall into' the nucleas. Thus, there must be an equal and opposite force keeping the electron in orbit. This is the electron's centrifugal force which is given by $\frac{mv^2}{r}$. Equating the two equations gives us

$$\begin{aligned}\frac{mv^2}{r} &= \frac{kZe^2}{r^2}\\ mv^2 &= \frac{kZe^2}{r} \qquad (2.27)\end{aligned}$$

The total energy of the electron is given by its kinetic energy and potential energy.

$$E = \frac{1}{2}mv^2 - \frac{kZe^2}{r} \qquad (2.28)$$

Inputting Eq. (2.27)

$$\begin{aligned}E &= \frac{1}{2}\frac{kZe^2}{r} - \frac{kZe^2}{r}\\ &= \frac{kZe^2}{r}(\frac{1}{2} - 1)\\ &= -\frac{kZe^2}{2r} \qquad (2.29)\end{aligned}$$

If we input Bohr's equation for the atomic radius in place of r

$$\begin{aligned}E &= -\frac{kZe^2}{2}\frac{Ze^2\pi m}{n^2h^2\varepsilon_0}\\ &= \frac{-Z^2e^4m}{8\varepsilon_0^2n^2h^2} \qquad (2.30)\end{aligned}$$

where n is the energy level of the electron. For a hydrogen atom, $Z = 1$. Hence the equation becomes

$$E = \frac{-e^4m}{8\varepsilon_0^2n^2h^2}$$

and the difference between energy levels n_1 and n_2 is given by

$$\triangle E = \frac{-e^4m}{8\varepsilon_0^2h^2}(\frac{1}{n_1} - \frac{1}{n_2})$$

2.6 Further examples of quantization

Now that we understand the quantization of atomic energy, we can delve into some examples.

2.6.1 Franck-Hertz experiment

The Franck-Hertz experiment was designed by two scientists, James Franck and Gustav Hertz, to represent the quantum nature of atoms by demonstrating the excited states of mercury atoms. It won them the Nobel Prize in 1923. The setup consisted of a Tungsten filament within a vacuum chamber filled with mercury atoms. Current is passed through the filament to raise its temperature and hence, electrons are emitted through thermionic emission. These electrons are attracted to a positively charged perforated plate within the chamber. The collector at the end of the chamber has a negative charge to 'collect' the electrons (if they have sufficient energy to reach this point). The ammeter then registers the current flowing through the wire connected to the collector. This setup is shown in Fig. (2.32)

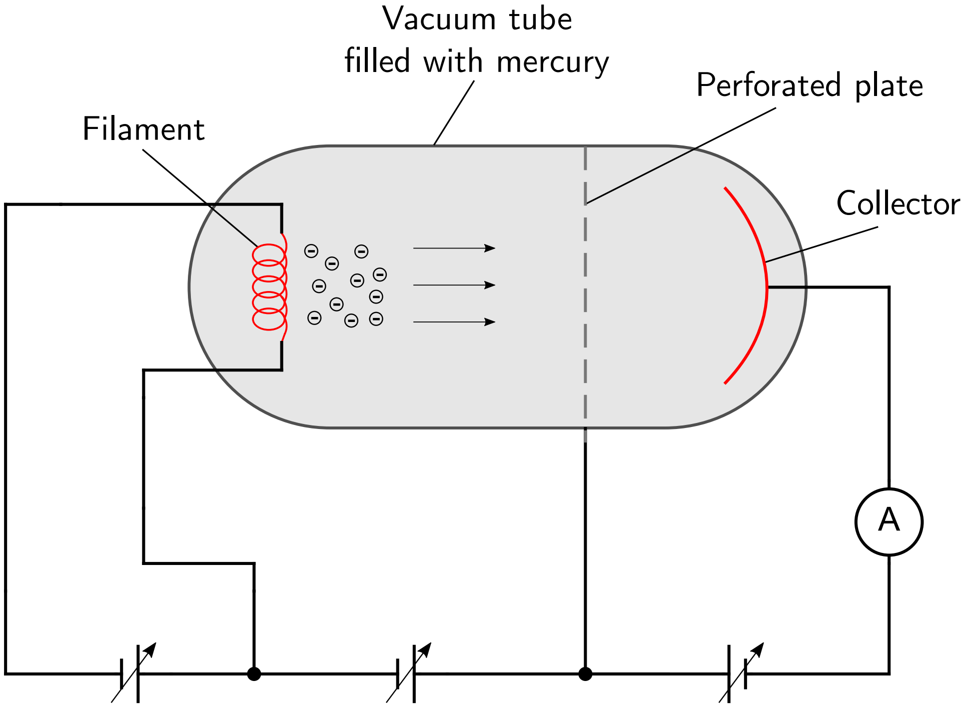


Figure 2.32: Setup for the Franck-Hertz experiment

As the electrons are bombarded, the dense mercury atoms may sometimes come in the way. This would lead to collisions between them, and hence, a transfer of energy from the electrons to the mercury atoms. This collision is inelastic as the external kinetic energy of the electron is transferred to the mercury atom in the form of internal energy. If enough internal energy is provided to the electrons of the mercury atom, they may jump to a higher energy level. This is the reflected in the equidistant 'dips' of current in the graphs in Fig. (2.33), which are the original results of the Franck Hertz experiment.

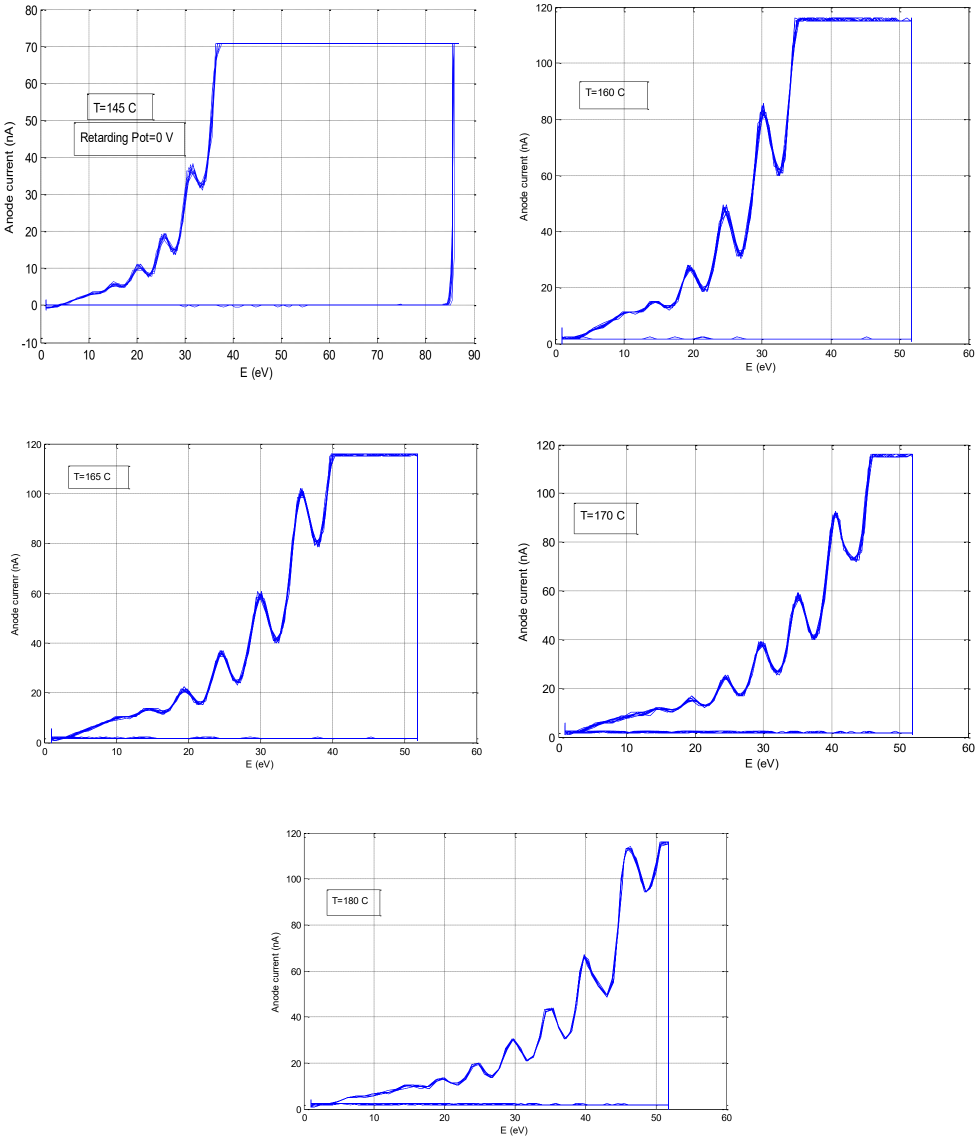


Figure 2.33: Original results of the Franck-Hertz experiment

These 'dips' occur as the bombarded electrons lose substantial energy to the electrons within the mercury atom, causing them to jump energy levels. The fact that these dips are equidistant show the quantized energy levels of the mercury atom.

As the voltage applied to the filament is increased, the temperature of the filament increases and the energy in these electrons increases as well.

This causes more 'dips' in the graph as the electrons within the mercury atom overcome more and more energy levels. On the other hand, if the temperature decreases, a point may come when the dips disappear and the plot shows a flat line. At this point, the energy provided to the Tungsten is lower than its work function, which is the minimum energy required for thermal emission.

Chapter 3

Schrodinger Equation and Quantum Systems

There are two types of waves in nature: **mechanical** waves like sound and water waves, and **electromagnetic** waves like gamma waves and visible light. The speed of these classical waves is given by:

$$v = f\lambda \tag{3.1}$$

where f is the frequency of the wave and λ is its wavelength.

We have seen that the wavelength of de Broglie's wave is given by $\hbar/mc$. This matter wave is denoted by $\psi(x,t)$ or by $\psi(x,y,z,t)$ in three dimensions. ψ holds the information about the evolution of matter wave. But how can we extract this information? The answer lies in classical wave equation, given by

$$\frac{\partial^2}{\partial x^2}\, y(x,t) = \frac{1}{v^2}\frac{\partial^2}{\partial t^2}\, y(x,t) \tag{3.2}$$

where y(x,t) is a classical wave given by $y(x,t) = A\sin(kx - \omega t)$ with A being the magnitude of the wave.

Plugging it in Eq. (4.2) gives us

$$k^2 A\sin(kx - \omega t) = \frac{1}{v^2}\omega^2 A\sin(kx - \omega t)$$

or

$$k = \frac{\omega}{v} \tag{3.3}$$

Notice that if we plug in the values of k and ω, we arrive at Eq. (4.1) once again.

3.1 Schrödinger Wave Equation

The behavior of a free non-relativistic particle is described by Schrödinger's wave equation:

$$\frac{\hbar}{\iota}\frac{\partial\Psi(x,t)}{\partial t}=\frac{\hbar^2}{2m}\frac{\partial^2\Psi(x,t)}{\partial x^2}+V(x,t)\Psi(x,t) \tag{3.4}$$

where $\Psi(x,t)$ is the wave function of the wave associated with the particle. This equation states that the space part is equal to the temporal part.

Let me propose the solution of this equation to be $\Psi(x,t)=A\exp i(kx-\omega t)$. By plugging it in Eq. (4.4) and considering V to be zero (as is the case with a free particle), we arrive at

$$\frac{\hbar^2k^2}{2m}=\hbar\omega.$$

As $\hbar k=p$ and $\hbar\omega=E$, the above equation becomes

$$\frac{p^2}{2m}=E \tag{3.5}$$

Let's make the equation a bit complex by placing the particle in a nonzero potential V. However, to simplify the situation, let's take $V=V(x)$ so that the Schrödinger equations becomes

$$\frac{\hbar}{\iota}\frac{\partial\Psi(x,t)}{\partial t}=\frac{\hbar^2}{2m}\frac{\partial^2\Psi(x,t)}{\partial x^2}+V(x)\Psi(x,t) \tag{3.6}$$

Can you come up with a way to solve this equation? (Hint: The left-hand side is entirely time-dependent while the right-hand side entirely space-dependent)

3.2 Time-Independent Scrödinger's Equation

By using variable separable method to solve the equation, let's consider $\Psi(x,t)=\psi(x)\phi(t)$.

Plugging it in Eq. (4.6)

$$\frac{\hbar}{\iota}\psi(x)\frac{d\phi(t)}{dt}=\frac{\hbar^2}{2m}\phi(t)\frac{d^2\psi(x)}{dx^2}+V(x)\psi(x)\phi(t)$$

Further dividing by $\psi(x)\phi(t)$

$$\frac{\hbar}{\iota}\frac{1}{\phi(t)}\frac{d\phi(t)}{dt}=\frac{\hbar^2}{2m}\frac{1}{\psi(x)}\frac{d^2\psi(x)}{dx^2}+\frac{V(x)\psi(x)}{\psi(x)}$$

I have multiplied and divided $V(x)$ with $\psi(x)$ to show that it is acting on $\psi(x)$.

Notice that the left-hand side is entirely time-dependent while the right-hand side entirely space-dependent. This can only be true if both the sides are equal to some constant. Let's call that constant E (the reasons of this choice will be obvious in the coming sections.)

Schrödinger equation becomes

$$\frac{\hbar}{\iota}\frac{1}{\phi(t)}\frac{d\phi(t)}{dt} = \frac{\hbar^2}{2m}\frac{1}{\psi(x)}\frac{d^2\psi(x)}{dx^2} + \frac{V(x)\psi(x)}{\psi(x)} = E \tag{3.7}$$

Solving the left-hand side of this equation gives us

$$\frac{\hbar}{\iota}\frac{1}{\phi(t)}\frac{d\phi(t)}{dt} = E$$

or

$$\frac{1}{\phi(t)}\frac{d\phi(t)}{dt} = \frac{\iota E}{\hbar}$$

By rearranging and integrating

$$\int \frac{1}{\phi(t)} d\phi(t) = \frac{\iota E}{\hbar}\int dt$$

or

$$\phi(t) = Ae^{\frac{\iota E}{\hbar}t} \tag{3.8}$$

Eq. (4.8) determines the evolution of the wave function $\Psi(x,t)$ in time.

Coming back to Eq. (4.7) and separating the spatial part gives us

$$\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x) \tag{3.9}$$

Here, $\frac{\hbar^2}{2m}\frac{d^2}{dx^2} = \frac{p^2}{2m}$ is the kinetic energy while $V(x)$ is the potential energy of the particle. It implies that E is the total energy of the system. Hence, Eq.(4.7) is telling us that for separable solutions, the energy of the system becomes constant.

The probability of finding the wave function $\Psi(x,t)$ can be calculated as

$$\begin{aligned}|\Psi(x,t)|^2 &= |\psi(x)\phi(t)|^2 \\ &= (|\psi(x)| \cdot |\phi(t)|)^2 \\ &= |\psi(x)|^2 |Ae^{\frac{\iota E}{\hbar}t}|^2 \\ &= |A|^2 |\psi(x)|^2\end{aligned}$$

Notice that the probability density is independent of time. Hence, Eq. (4.9) is called the **time-independent** Scrödinger equation.

3.3 Quantum Systems

3.3.1 Free Particle

A free particle will have zero force acting on it. As force is given by $F = -{}^{dV}/_{dt}$, this means that a constant potential V is acting on it. This makes TISE quite simple i.e.

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} = E\psi(x)$$

or

$$\frac{d^2\psi(x)}{dx^2} + \frac{2mE}{\hbar^2}\psi(x) = 0$$

The characteristic equation will be

$$p^2 + \frac{2mE}{\hbar^2} = 0$$

or

$$p = \pm\frac{\iota\sqrt{2mE}}{\hbar}$$

(p should not be confused with the momentum operator.) Thus, the solution comes out to be

$$\psi(x) = Ae^{\frac{\iota\sqrt{2mE}}{\hbar}x} + Be^{-\frac{\iota\sqrt{2mE}}{\hbar}x} \tag{3.10}$$

We know that

$$E = \frac{p^2}{2m} = \frac{k^2\hbar^2}{2m}$$

or

$$k = \frac{\sqrt{2mE}}{\hbar}$$

Thus, Eq. (4.10) becomes

$$\psi(x) = Ae^{\iota kx} + Be^{-\iota kx}$$

Therefore,

$$\begin{aligned}\Psi(x,t) &= \psi(x)\phi(t) \\ &= (Ae^{\iota kx} + Be^{-\iota kx})(e^{-\frac{\iota E}{\hbar}t})\end{aligned}$$

Since $E = \hbar\omega$,

$$\Psi(x,t) = Ae^{-\iota(kx+\omega t)} + Be^{\iota(kx-\omega t)}$$

The first term is a sine wave that propagates to the right while the second term is a sine wave that propagates to the left. Intuitively, this is what we should expect: a wave function that propagates in both directions.

3.3.2 Particle in a well

Suppose we have a particle confined to a one-dimensional well of height h. It is doomed to remain inside this well. It is just like a rabbit trapped between two mountains, each of potential energy $V = mgh$. The rabbit needs an energy E higher than this potential energy to escape from the mountains. If $E > V$, the rabbit, or the electron in our case, has enough energy to 'climb' the mountain and go outside. If the electron can never go outside of the well, the height h is practically ∞. Inside the well, we assume that the potential energy is zero.

Let's define three regions: I, II and III. Region I and III have infinite potential energy while region II has zero potential energy. Now, let's calculate the wave function of the particle. The wave function in regions I and III is zero because the potential energy is ∞. In region II, the Schrödinger's equation is

$$\begin{aligned}\frac{-\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x) &= E\psi(x) \\ \frac{-\hbar^2}{2m}\frac{d^2}{dx^2} - E\psi(x) &= 0\end{aligned}$$

Upon rearranging,

$$\begin{aligned}\frac{d^2\psi(x)}{dx^2} + \frac{2mE}{\hbar^2}\psi(x) &= 0 \\ \frac{d^2\psi(x)}{dx^2} + \frac{2m}{\hbar^2}\frac{p^2}{2m}\psi(x) &= 0 \\ \frac{d^2\psi(x)}{dx^2} + k^2\psi(x) &= 0\end{aligned}$$

Since P.E =0, the energy is only kinetic. Therefore, the solution to this equation is

$$\psi(x) = Ae^{ikx} + Be^{-ikx}$$

where A and B are arbitrary constants. We can find out their values by applying boundary conditions. We know that $\psi(x) = 0$ at $x = 0$ and $x = L$.

At $x = 0$

$$\psi_{II}(x=0) = A + B = 0 \text{ or } B = -A\,,$$

The original wave function in region II becomes

$$\psi_{II}(x) = A(e^{ikx} - e^{-ikx}) = 2iA\sin kx$$

this is the wave function inside the well.

At $x = L$

$$\begin{aligned} \psi_{II}(x=L) = 2iA\sin kL &= 0 \\ \text{or} \sin kL &= 0 \\ kL = n\pi \text{ or } k &= \frac{n\pi}{L} \end{aligned}$$

The possible values of $k = \pi/L, 2\pi/L, 3\pi/L, ...,$. i.e. k is quantized. So the energy

$$E = \frac{\hbar^2 k^2}{2m} = n^2 \frac{\hbar^2 \pi^2}{2mL^2}\,,$$

is quantized. The energy of the confined particle takes up special values

$$\begin{aligned} E_n = \frac{\hbar^2}{2mL^2}, \frac{4\hbar^2}{2mL^2}, \frac{9\hbar^2}{2mL^2} \cdots \\ \therefore E_n \propto n^2 \end{aligned}$$

The wave function can also be written as

$$\psi_{II}(x) = 2\iota A\sin\left(\frac{n\pi x}{L}\right) \tag{3.11}$$

The particle is present everywhere inside the well. In the ground state, it has maximum probability at the center of the well. In the first excited state, it will never be found at the center of the well. Moreover, if we increase the length L of the well, the energies will decrease.

There is some minimum, non-zero energy E_1 possessed by the electron even at the absolute zero temperature. This is called **zero point energy**. At 0 K, only the ground state is populated.

Consider 4 electrons, each of them in a potential well. At 0 K, all the electrons will be in their ground states. By increasing the temperature, there is a possibility that higher states will be occupied. It is the same as having 4 electrons inside a single potential well. When the electrons are excited to a higher state then upon decaying they will emit photon with a certain frequency. The electron before measurement is everywhere but the measurement restricts it and the wave function collapses to find a particular position for the electron. The hydrogen energy levels were derived from Bohr's model but they can also be calculated from Schrödinger equation. By luck, or unsurprisingly, the results of both of these calculations is the same. The old quantum theory fails because it cannot give the position of the electron. The energy levels for the hydrogen atom can be calculated if we use the correct potential i.e; $V(r) = -1/4\pi\epsilon_0$. If you consider a harmonic oscillator, then the energy levels are equally spaced with energy gap of $\hbar\omega$. The potential energy for the harmonic oscillator is $V(x) = 1/2kx^2$.

Any electron that is not bound to the nucleus in a hydrogen atom has continuous, non-quantized energy. Let's plot the wave functions for the infinite well. The ratio between their energies is

$$\frac{N_2}{N_1} = \exp\left(- \frac{E_2 - E_1}{k_B T} \right)$$

The wave function inside the well is given by Eq. 4.11. There is only one quantum number in this case because we are dealing with one-dimensional case. For the hydrogen atom there are three quantum numbers n, l, m_l because it is in 3 dimensions. Every quantum number quantizes something. In the case of a potential well, it quantizes energy. Suppose we take the wave function of the ground state

$$\psi_1 = 2iA \sin\left(\frac{\pi x}{L}\right),$$

the electron must exist somewhere in space otherwise its non physical. It is trapped inside the well so the probability of finding electron here is 1, i.e.

$$\int_0^L 4A^2 \sin^2\left(\frac{\pi x}{L}\right) dx = 1,$$

By solving this to find $A = \sqrt{2/L}$ and putting in the wave function, we get

$$\psi_n = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right), \tag{3.12}$$

Physically achievable wells

Infinite wells are an idealization and do not really exist in real life. If we make the well such that the floor of this well is sloping, it means that the potential energy is higher at one end and lower at the other. What if there a hump inside the well? What if electron is moving inside teh well encounters an obstacle such that the potential energy suddenly goes up by some value V_0? Once again, we imagine a rabbit in between the two mountains with a certain height. If this rabbit takes food and gathers up energy, if it wants to cross the mountain, it needs energy more than the potential energy mgh. This is a classical picture. However if this was a quantum rabbit, it could escape the mountains with lesser energy than that of the potential energy. In other words the *kinetic energy can be negative in quantum mechanics.*

Suppose we have two parallel plate capacitors placed at some distance from each other. If we apply a voltage V_0 to both capacitors, one of their plates becomes positively charged and the other negatively charged. In each of the plates, we make a hole such that if an electron impinges them with energy E, it can pass through the hole. Once the electron passes through, the hole is plugged so that the electron is trapped inside the parallel plate capacitor for eternity.

If the centrepoint of the first capacitor is $x = 0$ and $x = L$ of the second capacitor, what will the potential energy for this configuration look like? The potential for the electron is given by

$$V = \int_{x=0}^{a} E_{electric} dx \tag{3.13}$$

As $E_{electric}$ is constant, the potential depends linearly on x. We can draw the potential energy curve for such a configuration. However, if we reverse the polarity of plates, the potential energy curve will be the y-mirror of the former curve.

Let's bring the plates closer and see what happens. In this case, the potential energy curve will become steeper. If we keep on decreasing the distance between the plates to the point where this distance becomes infinitesimal, the setup will predominantly become an infinite potential well. If the voltages of the plates are not equal, how will the potential energy look like? It will be curved.

Let's consider another kind of potential well with a "sloping" floor. In this case, the potential will be lower at $x = 0$ and higher at $x = L$. The negative plate in the second capacitor is more negative as compared to the first one. If the ejected electron has energy E, do you think it can exist outside the well? It cannot, because the potential energy outside the well is ∞. The electron's energy needs to be conserved. What kind of wave function do we get inside this sloping potential well?

If we look at the time independent Schrödinger's equation

$$\frac{-\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x) + V(x)\psi(x) = E\psi(x)$$
$$\frac{-\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x) = (E - V(x))\psi(x)$$

by simplifying it a little more, we get

$$\frac{d^2}{dx^2}\psi(x) = -\left[\frac{2m}{\hbar^2}(E - V(x))\right]\psi(x)\,, \tag{3.14}$$

Since $E > V(x)$, the term to the right of the equation is positive. $E - V(x)$ is simply the kinetic energy of the particle. The term in the square bracket in Eq. 4.14 is $[2m/\hbar^2(E - V(x))] = k^2$. By putting this in Eq. (4.14) we get

$$\frac{d^2}{dx^2}\psi(x) = -k^2\psi(x)$$

k^2 will be less than that in the case of a straight floored potential well. As k decreases, the wavelength should increase. The solution should be the same mathematically except now the wavelength changes as we move from lower to higher potential. Since the potential energy is still infinite, the wave function must die at the walls. As k is decreasing when we move from left to right in the well, the momentum and velocity will also decrease. It means that while travelling in the potential well, the electron will spend more time in the proximity of the right wall. Hence, the amplitude of the wave function also increases. The probability density function also increases towards the right.

Consider another region where the potential energy is zero and an electron moving inside this region. The electron has some kinetic energy, and the solution of the Schrödinger's equation will look like

$$\psi(x) = Ae^{ikx} + Be^{-ikx} \tag{3.15}$$

If we inject the electron in this region, it will move from left to right. In Eq. (4.15), e^{ikx} represents its movement from left to right and e^{-ikx} represent its movement from right to left. The two counter-propagating matter waves interfere with each other and produce standing waves. Normally, we need confinement to produce standing waves. However, in this case, we are considering that the particle has no chance of reflection because there is no obstacle. For a free particle, there is only a field propagating in a single direction. If we include time dependence as well

$$\begin{aligned}\psi(x,t) &= (Ae^{ikx} + Be^{-ikx})e^{-i\omega t}\\ &= Ae^{i(kx-\omega t)} + Be^{-i(kx+\omega t)}\\ &= Ae^{ik(x-\frac{\omega}{k}t)} + Be^{-ik(x+\frac{\omega}{k}t)}\,,\end{aligned} \tag{3.16}$$

where $v_p = \omega/k$. For a free particle, the phase velocity is the same as group velocity. Since the signs of the velocity terms in this equation are different, the first term represents a forward propagation and the second term represents a backward propagation. Hence for a free particle, the wave function will be

$$\psi(x) = Ae^{ikx}$$

If we plot it, the real part will simply be a cos function and the imaginary part will be a sin function.

Now suppose that the potential energy initially zero and it takes a step of V_0. I we inject a particle with energy E into this region where $E > V_0$ and plot the wave function, the wave will have some wavelength initially. When it reaches the step, the difference between the energy and potential decreases, which decreases k. Consequently, the wavelength increases. The wavelength beyond the step will remain constant.

Now, instead of positive step, consider a valley where the potential energy is negative and the particle is again injected from the left. The wavelength will initially be smaller but when it reaches the valley, the wavelength decreases with an increase in k. Remember that we are only talking about $E > V$.

Suppose we have an atomic mirror placed on earth and an atom at a certain height h above it. The atom hits the mirror and rebounds. The atom is having an oscillatory motion in this region. The potential energy is zero on the ground and it increases linearly as the atom moves up. At the top, the potential energy is mgh – this is like the slopping potential with $E > V$. The difference between energy and potential is small such that k is small. Hence, the wavelength will be longer at the higher points and it decreases when we move down.

3.3.3 Quantum dot

A quantum dot is a nanoparticle with quantized energy levels. The smaller the quantum dot, the larger the energy difference between the energy levels.

We know that when an electron decays from E_1 to E_2, it will emit a photon. If we reduce the size of the quantum dot, the emitted photon will have a higher frequency. If we keep on decreasing the size of the quantum dot, the energy gaps will increase and the emitted photons will have higher frequency. In this way, we can tell the size of the quantum dot by looking at the emitted photon. This process is called fluorescence.

Let's now demonstrate. We have prepared a solution of quantum dots. These are semiconductor quantum dots made up of Cd-Se of varying sizes. The solutions are of different colors to differentiate between the sizes of the quantum dots. This phenomenon is shown in the demonstration.

3.4 Time dependence of wave functions

Let's continue our discussion of the particle in the box. If we want to find out the realm of possibilities for the position of particle i.e. Δx, this would be the **order** of L (the dimensions of the well). If $\Delta x \sim L$, $p \sim \Delta p = \hbar/2L$. In this situation, the kinetic energy becomes

$$E = \frac{p^2}{2m} = \frac{\hbar^2}{2mL^2}$$

This is almost of the order of minimum energy. Thus, without solving the Schrödinger's equation, we can find te minimum energy just by looking at the uncertainty principle. If the well shrinks in space, the minimum energy goes up.

We have discussed the dependence of wave function on space but we have not discussed it on time. If the wave function evolves with time, it means that $\Psi(x, 0)$ evolves to a new wave function $\Psi(x, t)$ after some time t. We have the time-dependent part is the Schrödinger's equation as

$$\begin{aligned} \frac{ih}{f(t)}\frac{d}{dt}f(t) = E &\Rightarrow \iota h\frac{d}{dt}f(t) = Ef(t) \\ ih\frac{df(t)}{f(t)} = Edt &\Rightarrow \frac{df(t)}{f(t)} = -\frac{\iota Edt}{\hbar} \end{aligned} \tag{3.17}$$

On solving this differential equation, we get

$$\begin{aligned} \ln f(t) &= -\frac{\iota Et}{\hbar} \\ f(t) &= \exp\left(- \frac{\iota Et}{\hbar}\right) \end{aligned} \tag{3.18}$$

Hence, the total wave function becomes

$$\Psi_n(x,t) = \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi x}{L}\right)\exp\left(-\frac{\iota E t}{\hbar}\right). \tag{3.19}$$

If we write the complete wave function for the first state

$$\begin{aligned}\Psi_{n=1}(x,t) &= \sqrt{\frac{2}{L}}\sin\left(\frac{\pi x}{L}\right)\left[\cos\left(\frac{Et}{\hbar}\right) - i\sin\left(\frac{iEt}{\hbar}\right)\right]\\ &= \sqrt{\frac{2}{L}}\sin\left(\frac{\pi x}{L}\right)\cos\left(\frac{Et}{\hbar}\right) - i\sqrt{\frac{2}{L}}\sin\left(\frac{\pi x}{L}\right)\sin\left(\frac{Et}{\hbar}\right) \end{aligned} \tag{3.20}$$

Let's have a look at the plot of the real part of Eq. (4.20) at different times starting from $t = 0$, which will look like this:

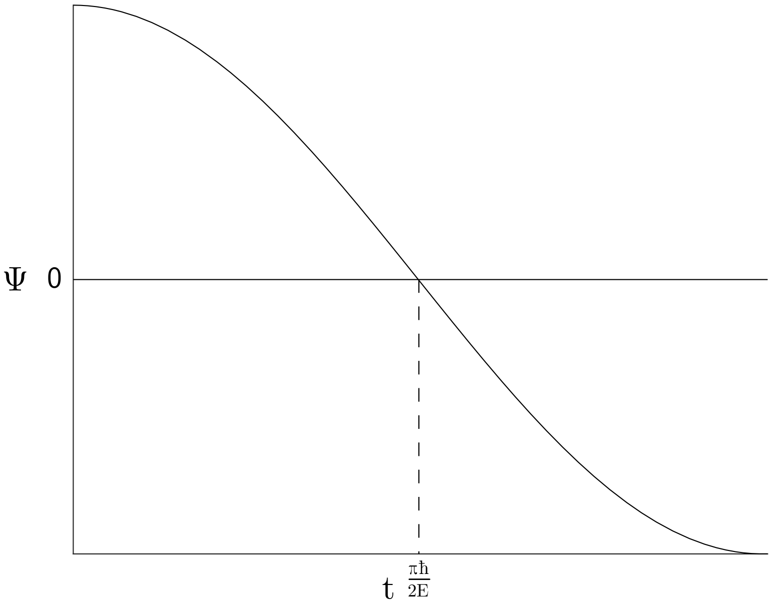


Figure 3.1: Plot of the real part of Eq. (4.20) at different times starting from $t = 0$

As we progress in time, the real part will decrease because of the decrease in cos function. There is a specific time at which the real part becomes zero.

$$\frac{Et}{\hbar} = \frac{\pi}{2} \Rightarrow t_0 = \frac{\pi\hbar}{2E} \tag{3.21}$$

If the time increases any further, the cos function becomes negative and the wave function starts increasing in the negative direction till it reaches a maximum point. This is a static wave –a wave which start from a maximum, goes all the way to the negative maximum and then starts oscillating.

Now let's plot the imaginary part of Eq. (4.20). At $t = 0$, the wave function is zero because $\sin(0) = 0$. As time increases, the sin function starts increasing but in the negative direction because of the negative sign. It is the converse of the real part i.e. once it reaches the maximum point in the negative direction, it starts increasing upon moving further in time till it reaches a positive maximum. There is a flip-flop of the real and imaginary parts; as one goes up, the other goes down and vice versa.

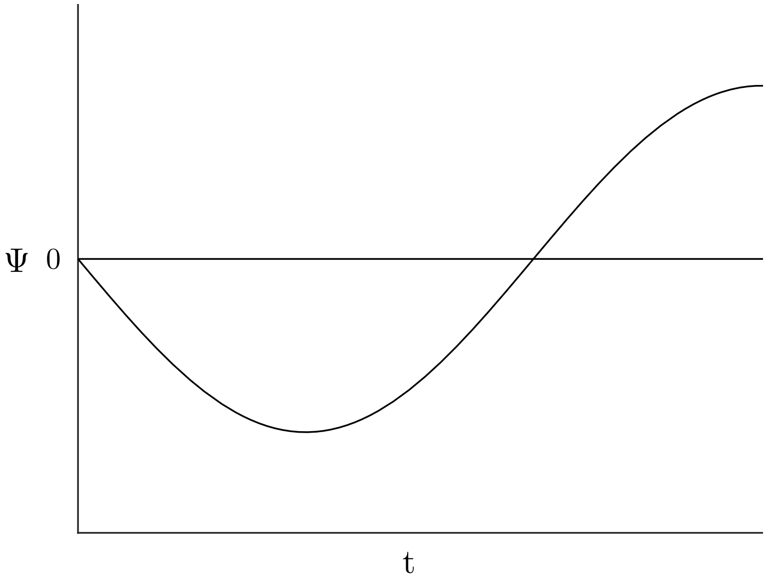


Figure 3.2: Plot of the imaginary part of Eq. (4.20) at different times starting from $t = 0$

In real experiment, we do not measure the wave functions of the particles. Instead, we measure their probability densities. The probability density $P(x, t)$ of Eq. (4.20), it becomes

$$P(x,t) = \psi_n^*(x,t)\psi_n(x,t) = |\psi_n^*(x,t)|^2 = \frac{2}{L}\sin^2\left(\frac{\pi x}{L}\right) \tag{3.22}$$

Note that the probability density is independent of time. Hence, we can ignore it for time dependence. Even though the wave function is changing with time, the probability does not change with it. If we plot this function, it will simply be the square of the spatial part of the wave function. This is called a *stationary wave*.

In the ground state, the maximum probability of finding the electron is at the middle of the well. The electron is everywhere because its mass has a field. It has a distribution of charge as well, changing only with space and not with time. This means that the charge is not accelerating and hence not emitting any radiations. This is the last nail in the coffin of the Rutherford's model which says that an electron emits radiations even in its own orbit. Schrödinger's equation verifies the postulate of Bohr's theory that an electron can not emit any radiation in its own state.

3.4.1 A software-aided approach

The probability density function and the real and imaginary parts of the wave function are also shown here plotted with the help of a software. For the ground state of an infinite potential well it looks like the following. Now I am going to run the simulation where in the back script the Schrödinger's equation is solved in JAVA script. We want to see what happens to the plot when we change the time. We are observing that the real part is decreasing, the imaginary part is increasing in the negative direction but nothing happens to the probability density function. The real part is going towards the zero and the imaginary part is going to the negative maximum. Further increasing time the real part increases in the negative direction and the imaginary part goes up in the positive. We notice that the variation is faster in between the walls and slower towards the walls.

I am now going to run the simulation a bit faster. This is another way of showing the same results. The probability density function does not change with time. The magnitude of the wave function is not changing with time. The phase is changing with time and it is shown with different colors in the simulation.

Now lets look at the first excited state $n = 2$. There is only node in the wave function and two lobes in the probability density function. The red part of the wave function is going up and down in time, the blue part is also changing direction and it is having a phase difference of 90° with the red. However the probability density function does not change with time, you will never find the electron at the center of the well. The total area under the curve is 1 but the probability is different at different positions in the well.

Lets look at even the higher excited states. In the 5th excited state once again the probability density function does not change with time and the wave functions are oscillating. Are these wave functions oscillating at a higher frequency compared to the lower states, why is it so? Because the energy is higher so the frequency is higher. We can see five nodal points in the density function excluding the walls for $n = 6$. We can say that for the n state there are $n - 1$ nodes.

Now suppose we have an electron which is in the two energy levels at the same time or it is in a superposition. Suppose the electron is in the ground and the first excited state at the same time. The wave function for the electron will be a superposition of the two states. Now the real and imaginary parts of the wave function is changing but it is changing such that the probability density function is also changing. The density function has two humps when one goes up the other one goes down.

3.5 Superposition of states

Let's go back to the infinite potential well having states $n = 1$ and $n = 2$. We derived the general form of the wave functions as a function of time and space. In this case, the energy of the particle is fixed in any state and does not have any imprecision in it. We cannot tell where the particle *exactly* is – it is a pointless question in quantum mechanics. However, we can calculate the probability of finding the particle in a particular region. In the old quantum theory, the energy of the particle depends upon the distance from the nucleus. However, such is not the case in new quantum theory.

For the ground and first excited states, we can write the complete wave functions as

$$\begin{aligned}\Psi_1(x,t) &= \sqrt{\frac{2}{L}}\sin\left(\frac{\pi x}{L}\right)\exp\left(-\frac{iE_1t}{\hbar}\right)\\ \Psi_2(x,t) &= \sqrt{\frac{2}{L}}\sin\left(\frac{2\pi x}{L}\right)\exp\left(-\frac{\iota E_2t}{\hbar}\right)\end{aligned} \tag{3.23}$$

In the double slit experiment, we have seen that the electron can pass through the two slits at the same time. This implies that its wave function is a composition of two wave functions. In this situation, we cannot tell through which slit the electron has passed. It may make you question whether the electron can exist in a superposition of two states E_1 and E_2.

The wave function of the electron in a superposition of states will be

$$\Psi(x,t) = \sqrt{\frac{1}{2}}\Big(\Psi_1(x,t) + \Psi_2(x,t)\Big) \tag{3.24}$$

In superposition, we lose the information of the energy of the electron. Putting Eq. (4.23) in Eq. (4.24) we get

$$\Psi(x,t) = \sqrt{\frac{1}{2}}\Bigg(\psi_1(x)\exp\Bigg(-\frac{\iota E_1 t}{\hbar}\Bigg) + \psi_2(x)\exp\Bigg(-\frac{\iota E_2 t}{\hbar}\Bigg)\Bigg) \tag{3.25}$$

The probability density function is calculated as

$$P(x,t) = \Psi^*(x,t)\Psi(x,t) \tag{3.26}$$

By putting in the values

$$\begin{aligned} P(x,t) = \frac{1}{2}\Bigg(\Bigg(\psi_1^*(x)\exp\Bigg(\frac{\iota E_1 t}{\hbar}\Bigg) + \Bigg(\psi_2^*(x)\exp\Bigg(\frac{\iota E_2 t}{\hbar}\Bigg)\Bigg). \\ \Bigg(\Bigg(\psi_1(x)\exp\Bigg(-\frac{\iota E_1 t}{\hbar}\Bigg) + \Bigg(\psi_2(x)\exp\Bigg(-\frac{\iota E_2 t}{\hbar}\Bigg)\Bigg) \end{aligned} \tag{3.27}$$

after multiplication and simplification the result is

$$P(x,t) = \frac{1}{2}\Bigg(|\psi_1(x)|^2 + |\psi_2(x)|^2 + \psi_1(x)\psi_2(x)\Bigg[\exp\Bigg(\iota\frac{\Delta E t}{\hbar}\Bigg) + \exp\Bigg(-\iota\frac{\Delta E t}{\hbar}\Bigg)\Bigg],$$

where $\Delta E = E_2 - E_1$. On simplifying the equation a bit further

$$P(x,t) = \frac{1}{2}\Bigg(|\psi_1(x)|^2 + |\psi_2(x)|^2 + 2\psi_1(x)\psi_2(x)\cos\Bigg(\frac{\Delta E t}{\hbar}\Bigg)\Bigg) \tag{3.28}$$

Since $\omega = \Delta E/\hbar$, therefore

$$P(x,t) = \frac{1}{2}\Bigg(\psi_1(x)^2 + \psi_2(x)^2 + 2\psi_1(x)\psi_2(x)\cos(\omega t)\Bigg), \tag{3.29}$$

This gives us the probability density of the wave function which changes in both space and time. This is the first time in our study so far that the probability density has changed with time.

Let's now look at how it changes with time. Plot $\cos(\omega t)$ as a function of time. Starting from the point A to point B, the time $t = \pi/2\omega$ at which the function is zero. It is negative maximum at $t = \pi/\omega$. At point A, the probability is

$$\begin{aligned} P(x,0) &= \frac{1}{2}\Bigg(\psi_1(x)^2 + \psi_2(x)^2 + 2\psi_1(x)\psi_2(x)\Bigg) \\ P(x,0) &= \frac{1}{2}\Bigg(\psi_1(x) + \psi_2(x)\Bigg)^2 \end{aligned} \tag{3.30}$$

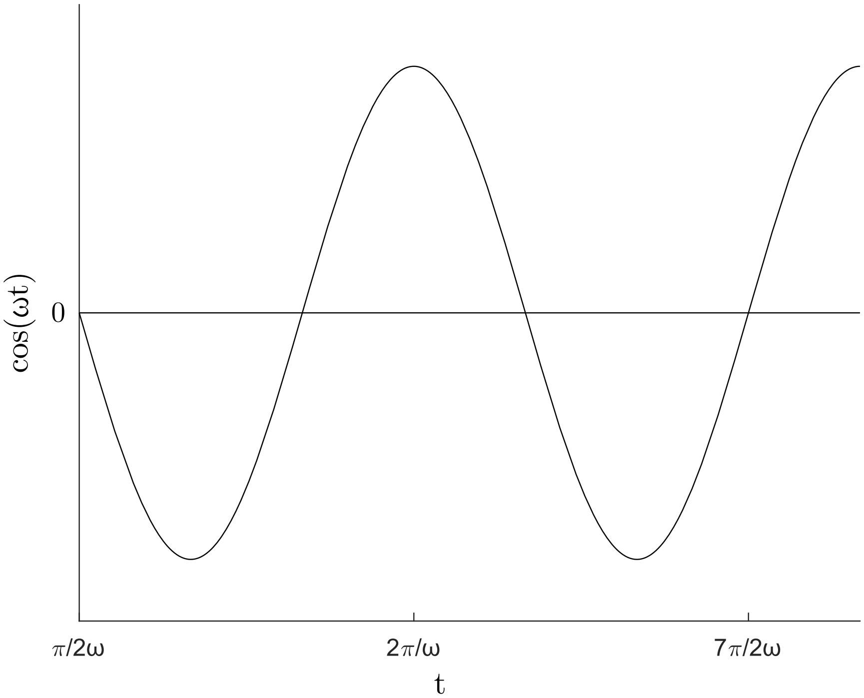


Figure 3.3: Plot $\cos(\omega t)$ as a function of time

At point B, it is

$$P\left(x, \frac{\pi}{\omega}\right) = \frac{1}{2}\left(\psi_1(x) - \psi_2(x)\right)^2, \tag{3.31}$$

Let's now have a look at the plot superposition of the states and its probability density function at points A and B. Since the cos function is sinusoidal, there is a periodic toggling in the probability density at different times with a frequency ω.

The analogy from the old quantum theory is that when an electron jumps from the upper level to the lower level, it emits a photon. However, in the new quantum theory, the probability density shows a charge that is smeared out in space and redistributing itself in time. There is equivalent to a force acting on the electron; it accelerates under that force and emits radiation. The radiation is emitted as a stream of photons. So we should stop thinking about the electron being a discrete particle in the upper level which somehow emits a photon when it decays. In the proper quantum theory, superposition gives the transition.

Let's look at a demonstration. This is a continuous simulation of what is happening. The program is solving the Schrödinger's equation. One hump

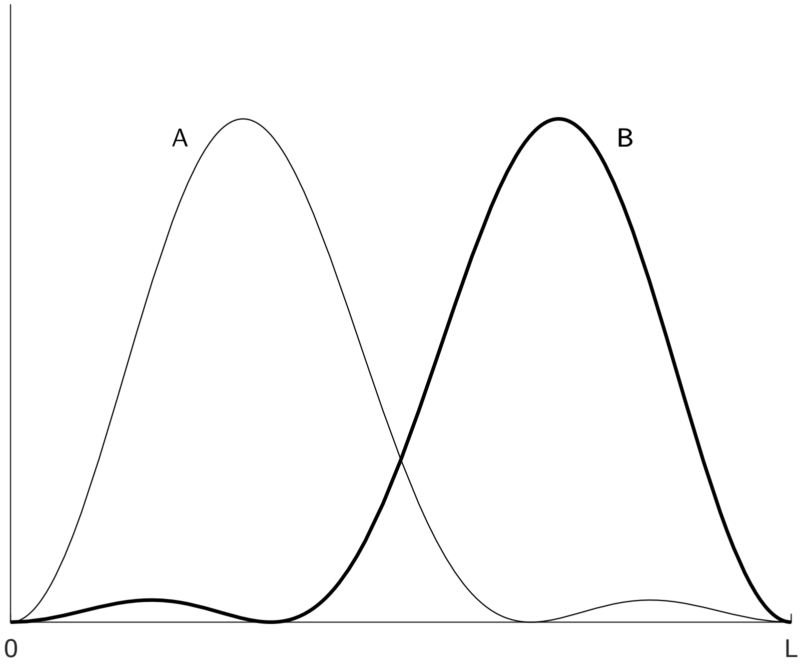


Figure 3.4: Superposition of the states and its probability density function at points A and B.

is going up and the other one going down. This toggling is at frequency ω which is also the frequency of the emitted photon. If we integrate this function from 0 to L, we will always get 1, implying that there is a 100% chance that the particle is inside the potential well. The electron is doomed to remain in this well forever.

There can be a superposition of hundreds of states, which happens in quantum computing. In these computers we can have the binary states 0 and 1 at the same time. Let's see what happens when there is a superposition of three states. The probability density function for three states in superposition toggles as well, but in this case it has two frequencies. We can also plot the superposition of higher states.

Let's look at the probability density function for the superposition of two states. It takes 0.8 s for the peak to go up and then go down. If we increase the frequency, the time taken will decrease. Let's now consider what happens if we increase the energy difference between the states. How long will the process take to repeat itself? If we increase the spacing between the two peaks, the oscillation will increase and correspondingly, the time taken will decrease. Consequently, a photon of a higher frequency is emitted.

3.6 Quantum obstacles and tunneling

Consider a step function where the step V_0 is at $x = 0$. The infinite well has two regions: region I where the potential is zero and the region II where it is V_0. The kinetic energy of the particle will be reduced when it encounters the step but it will still have enough kinetic energy to go to region II.

In region I, the wave function is

$$\psi_I(x) = Ae^{ikx} + Be^{-ikx} \tag{3.32}$$

where $k^2 = 2mE/\hbar^2$. Both terms will now be included in the wave function because of the discontinuity in the potential. The particle is encountering another medium at the step. It means there has to be some force acting on the particle for it to accelerate. It can reverse its velocity so the B in Eq. (4.32) is zero.

Let's write the solution in region II

$$\psi_{II}(x) = Ce^{ik'x} + De^{-ik'x} \tag{3.33}$$

But here, D equals zero because once the particle enters region II, it has no obstacle. Hence, we are left with

$$\psi_{II}(x) = Ce^{ik'x}\,, \tag{3.34}$$

where $k'^2 = 2m(E - V_0)/\hbar^2$.

If we plot the wavefunctions, the wavelength will be smaller in the region I but due to the decrease in k in region II, the wavelength will be greater. In this case, either the coefficient C is large or the particle is moving slowly. For this scenario, we now define the transmission probability.

3.6.1 Transmission and reflection probability

In this configuration, if a particle is injected from the left, it may either be reflected from the step or go into region II. There is a probability of reflection and transmission of a particle just like a wave. If I need to find the transmission probability I would make a plane in region I and in region II and find the probability of the particle passing this screen from left to right. When I divide the probability of particle passing the screen b by the probability of it passing the screen a, I get the transmission probability. The transmission probability is defined as *the probability that the particle crosses* *b per unit time divided by the probability that the particle crosses a per unit time.* We can also define

$$\frac{P}{t} = \frac{P}{l}\frac{l}{t} = |\psi(x)|^2 v \tag{3.35}$$

where P is probability, l is the length and t is the time. We can find the speed as

$$p = \hbar k \Rightarrow v = \frac{\hbar k}{m}$$

the transmission probability then becomes

$$T = \frac{|C|^2 \frac{\hbar k'}{m}}{|A|^2 \frac{\hbar k}{m}} \Rightarrow \frac{|C|^2 k'}{|A|^2 k} \tag{3.36}$$

this is the transmission probability. We now have to find what are A and C by boundary conditions. The probability of reflection will then be

$$R = 1 - T\,, \tag{3.37}$$

applying boundary conditions, at $x = 0$ the wave functions are equal

$$A + B = C\,, \tag{3.38}$$

the other condition is that the derivatives at $x = 0$ will also be equal

$$\begin{aligned} ik(A - B) &= ik'C \\ A - B &= \frac{k'}{k}C \end{aligned} \tag{3.39}$$

Adding Eq. (4.38) and Eq. (4.39) we get

$$A = \frac{C}{2}(1 + \beta)\,, \tag{3.40}$$

where $\beta = k'/k$. The transmission probability becomes

$$T = \frac{4|C|^2 k'}{|C|^2(1+\beta)^2} = \frac{4\beta}{(1+\beta)^2}\,, \tag{3.41}$$

This is the transmission probability of a single particle. In case there are N particles, the probability will be NT.

We define the reflectivity R as

$$R = \frac{|B|^2 k}{|A|^2 k} = \frac{|B|^2}{|A|^2} \tag{3.42}$$

here k cancels out because both incident and reflected parts are in the same medium. We also know that

$$R + T = 1. \tag{3.43}$$

Let's consider the case of negative potential again. Initially, there is no potential followed by a valley where the potential is negative. A particle with energy E enters the infinite potential well. Will this particle be reflected at the discontinuity? Yes, because if there is some discontinuity in the potential, some force will act on the particle whether the potential is positive or negative. If the potential is positive, the particle should pass without any reflection but quantum mechanically, there is reflection. Moreover, as the energy of the particle is higher than the potential, its kinetic energy should increase in region II. However, the particle can still be reflected because of the discontinuity in potential.

It's just like a pool of water – a cricket ball is thrown into the pool with some energy. Suppose the energy of the ball is 50J and the pool of water presents a potential barrier of 20 J to it. Classical mechanics says that the ball should enter the pool with its kinetic energy reduced. Inside the water, the ball experiences some viscous force and its energy drops. Once it enters the pool, there is no chance for the ball to reflect back. However, if this were a quantum cricket ball, there would have been some finite probability of its reflection from the surface (counter-intuitive, right?).

To understand it better, consider a beam of light from air reaching glass. When the beam impinges the air-glass interface, some of it will be refracted to the glass will some of it will be reflected from the interface. As the refractive index of glass is higher than that of air, the speed of light is lower inside glass.

3.7 Quantum tunneling

Suppose the energy of the incident particle is smaller than the potential barrier and the particle is injected from the left. In region I, the wave function would be same as that for $V = 0$. According to classical mechanics, the particle cannot penetrate in region II. However, quantum mechanically, this is not true. If we solve the Schrödinger's equation in region II, we can come up with a wave function

$$\psi_{II}(x) = Ce^{-\alpha x} + De^{\alpha x} \tag{3.44}$$

where $\alpha^2 = 2m(V_0 - E)/\hbar^2$. Note that there is no imaginary term in this wave function and that α is real. The second term in Eq. 4.44 blows away when $x \to \infty$. Since the wave function cannot be infinity, we neglect this term. In region I, the wave function will be oscillatory but when it touches the boundary, it decays exponentially. Even though the energy of the particle is less than the potential, it penetrates to a small length in region II which is classically forbidden region.

Suppose situation is changed such that the potential is zero in region I and III but V_0 in region II. The energy of the incident particle is less than the barrier and the length of the barrier is L. Even in this case, the particle can penetrate in region II provided that we reduce the length of the barrier. This phenomenon is called quantum-mechanical tunneling.

In region I, the wave function will be oscillatory. In region II the wave function will be decaying but it will have both the terms of Eq. (4.44) present because x does not go to ∞ here. The decay will be very fast. If the wave function has not decayed completely when we reach the other barrier – which is possible if the length is small – then some of the energy can leak out to region III. The wave function in the region III will have the same k values but the height will be very small.

Let's go back to the example of glass and air. If we increase the angle of incidence of a beam of light traveling from glass to air such that $\theta_i > \theta_c$, the beam will not be refracted anymore but will *only* be reflected back in glass. This phenomenon is called total internal reflection. This is just like a photon hitting the barrier with less energy than the barrier so that all of light is reflected. If the beam somehow leaks into the air, it will be damped very quickly. Since we are not able to see the quick damping, we place another glass tube very close the interface. The new glass is inserted such that it is in the region of the penetration depth of the leaking field. If the glass is brought extremely close to the first interface, we can observe the transmission of light beam to air. This is called optical tunneling or *frustrated total internal reflection.*

Let's see a demonstration of this effect (Fig. 4.5). Consider a green color laser light on the screen with a glass prism with a certain refractive index placed in front of it. The refractive index is such that $\theta_c = 42°$. Let's take $\theta_i = 45°$. As $\theta_i > \theta_c$, total internal reflection should take place. But, by virtue of quantum tunneling, some of the light must go into the classically forbidden region. If we place another prism close the first one, no light is transmitted. To capture the evanescent light, the prism needs to be extremely close to the first such that the dimension of the prism should match the wavelength ($\sim$ 500nm). In order to capture the light we have made another arrangement, in this arrangement two prisms are joined back to back with a fluid in between them such that the distance between them is less than a micron. In this case, we will be able to capture the evanescent light and frustrate total internal reflection.

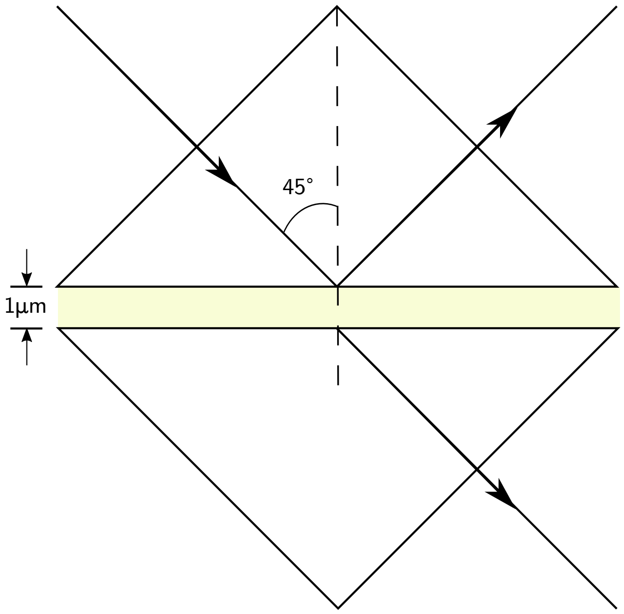


Figure 3.5: Demonstration for quantum tunneling

3.7.1 Radioactivity and nuclear stability

We are now looking for quantum devices. We discussed tunneling and we can even say that we exist because of tunneling. An alpha decay can only be explained through tunneling. One of the greatest successes of quantum mechanics is the explanation of radioactivity.

Suppose we have a mountain with a person A standing on one side of the mountain and a person B standing on the other side of it. The mountain has a height such that a person standing at the top would have a potential energy of 1000J. The person A throws a cricket ball with some energy so that it reaches the person B. B then measures the energy of the ball.

What is the minimum energy that this ball must possess? It should be greater than 1000J because on coming down all of ball's potential energy is converted into kinetic energy. Let's say that the energy is 800J – is this classically possible? No, because this would mean that the potential energy at the top of the mountain is -200J. However, since tunneling works, this ball can somehow reach B even with its energy of 800J. This effect is observed in radioactivity.

Suppose we have a ${}^{234}_{92}U$ nucleus. This (parent) nucleus decays gives rise to a daughter nucleus Thorium-232 along with α particle.

$$
{}^{234}_{92}U \rightarrow^{232}_{90} Th + \alpha particle \tag{3.45}
$$

A parent nucleus with 90 protons and large number of neutrons decays and gives a daughter thorium nucleus with 90 protons along with an α particle. The α particle moves away with some kinetic energy typically of the order of 5MeV. This process is known as α decay and it is a naturally occurring process. Both of the resulting species are positively charged particles.

Suppose I draw a plot of the distance from the nucleus r against the potential $V(r)$. What happens to the potential energy of the α particle when it is brought close to the Thorium. The potential energy increases according to the following equation:

$$
|V(r)| = \frac{1}{4\pi\epsilon_0}\frac{q_1 q_2}{r} \tag{3.46}
$$

If the parent nucleus has not decayed yet, the α particle must have a very high potential energy. Increasing the number of protons in a nucleus increases the potential energy. In that case, what keeps the nucleus stable?

In 1930, a theory was proposed to explain nuclear stability. According to this theory, when protons are brought together, the repulsive force is overcome by an attractive force. That is, there is a switch from the repulsive to the attractive force at smaller r and the potential energy of the nucleus decreases. The repulsive force is electrostatic. The attractive force, on the other hand, is not electrostatic because then the two forces would repel each other. The attractive force is called a *strong nucleus force* and it is a short range force.

In describing α decay, we say that α particle is bound to the Thorium nucleus by this strong nuclear force and it sees a potential barrier. If $r_0 = 7.5 \times 10^{-15}$m then

$$
|V(r)| = 35\text{MeV} \tag{3.47}
$$

But the α particles only have an energy of 5MeV and they are coming from inside of the nucleus. This is analogous to the example of the cricket ball. As the ball cannot go from above the barrier, it tunnels through it.

3.7.2 Tunneling probability

There is a certain probability of the tunneling of this α particle denoted by T. It depends on the height of the barrier and the initial energy of the α

particle. There are a large number of α particles hitting the barrier and there is a small probability that a single α particle can tunnel through the nucleus. The tunneling probability is given by

$$T \sim exp\left(- \frac{L}{\delta} \right) \quad \text{where} \quad \delta = \frac{\hbar}{\sqrt{2m(V_0 - E)}} \tag{3.48}$$

where E is the energy of the α particle and V_0 is the height(potential) of the barrier.

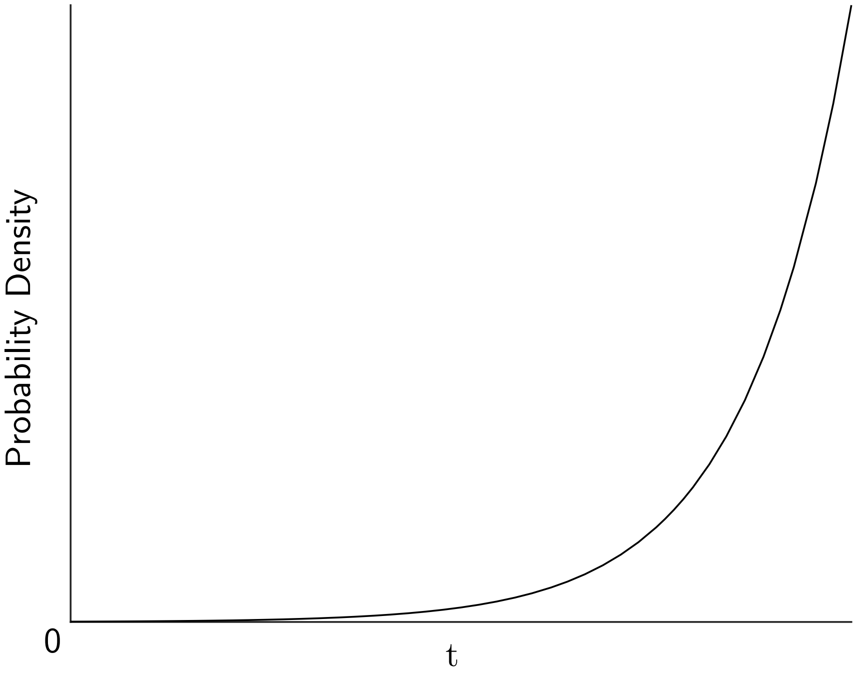


Figure 3.6: Tunneling probability density function

If $T \sim 10^{-15}$, there will still be a large number of α particles coming out. But the process is indeed probabilistic, as it depends upon the initial energy of the α particles. We can also write

$$\frac{N}{t} \simeq \frac{N_\alpha}{t} \times T \tag{3.49}$$

where N is the number of decays, t is the time and N_α is the number of α particles hitting the barrier. The mean time or the average time between two decays is given by the inverse of Eq. (4.49).

Consider α particles having energies of $4, 5, 6, 7, 8, 9$ MeV – much less than the barrier – coming out from the parent nuclei. They are escaping the nucleus because of tunneling. In other words, if $\hbar = 0$, the penetration length would be zero and the α particle would not be able to reach the classically forbidden region and no tunneling would take place. Because $\hbar$ is very low, we cannot observe tunneling in daily life. However, the mean lifetime of decay is very diverse because of the huge range of tunneling probabilities which in turn depend on the initial energy of the α particles. Suppose the length of the barrier $L \simeq 10$nm, $V_0 = 0.1eV$ and $E \simeq 0.05$eV. Now, if we increase E to 4eV, we will see a long range of probabilities and corresponding mean times of decay.

Let's have a look at a demonstration. We have some radioactive materials which are sealed. They have low activity and are safe to use. One of the sources is placed in a holder with a lead sheet. This is continuously emitting radiations which are picked up by the Geiger-Muller tube, a device which detects the number of counts per unit time. When an α or a β particles hits the tube, it produces a flash which can be heard. We can observe the decays on a computer screen as well as the counts per second. As we are not receiving a constant count, we can assess that the process is probabilistic. Let's plot the number of occurrence versus the number of counts after running the experiment for some time (Fig 4.7). The plot is in

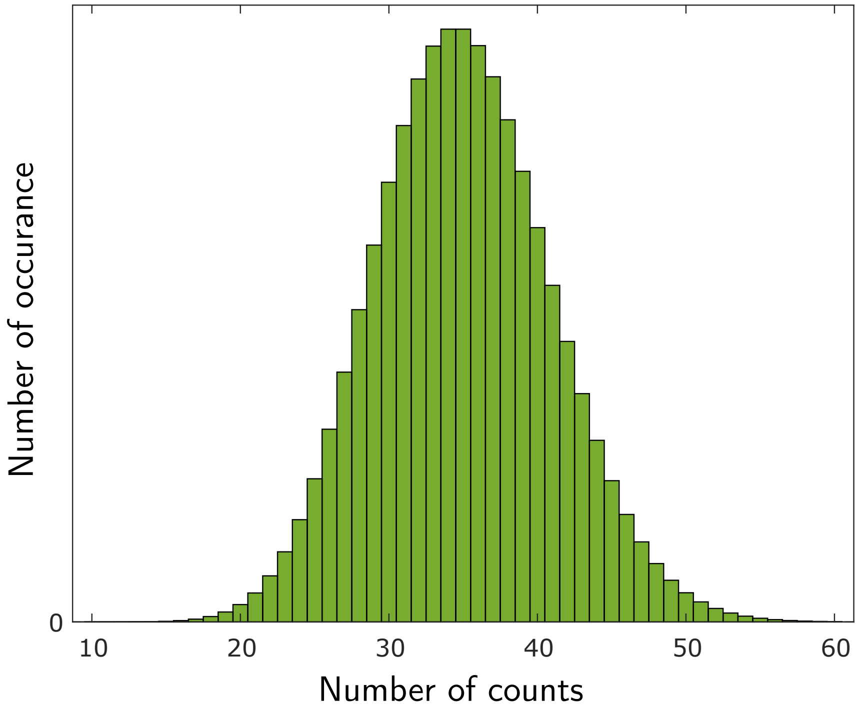


Figure 3.7: Histogram of number of occurrence versus the number of counts

the form of a histogram. This histogram peaks at 35 but it also has lower regions. This is due to the probabilistic nature of the radioactivity. Such a process is called the Poisson distribution.

3.7.3 Scanning tunneling microscope

Imagine a solid surface made up of atoms which are bumping out. There are free electrons inside the metal which cannot come out of the metal because of the work function. The electrons initially experience zero potential followed by a potential step of work function ϕ. Therefore, they cannot escape the metal.

Let's change the situation a little bit. Suppose we have two pieces of metal and one of them has a tip. We apply a voltage to the metals such that the tip is a bit positive with respect to the other metal. The distance between the tip and metal is L. If we bring the tip closer and closer to the metal, the metal tip will eventually 'scan' the surface of the metal. Whenever there is an atom, there is a bulge in the surface. The atoms are spaced apart at a certain distance. When the tip is directly above the atom, the effective distance L decreases. As the tip is positively charged with respect to the metal surface, the electrons will favorably tunnel from the tip to the metal. If we connect an ammeter, we can actually measure the current flowing in the metal plates. However, if the distance L is increased, the current will decrease. This is called tunneling current. The current is then converted to an audio signal which indicates the detection of the atom.

If the tip is brought extremely close to the atoms, a chemical bond can be formed between the atoms. Now, if the metal tip is carried to a different location, the atom will be carried off along with it. This allows us to pick individual atoms from the metals and place them wherever we like. In this way, scanning tunneling microscope can manipulate metals.

Scanning tunneling microscope has revolutionized the physics. This microscope was invented in the late 1980s and the inventors, Gerd Binnig and Heinrich Rohrer, were awarded the Nobel Prize in 1994 for designing it.

3.7.4 Our sense of smell

You might have studied about cell signaling in biology courses. Our cells have a cell membrane and the atoms present outside act like messengers (e.g. adrenaline is a molecule which gets excited and transmits signals to the cell

membrane). This is possible due to the fact that the cell membrane has receptors embedded in it. The messenger molecules bind to the receptors and trigger the activation of some proteins known as G-proteins. Once activated, these G-proteins act like a molecular switch which is normally in off state. When it is turned on, other channels in the cell membrane are opened which leads to the release of sodium and potassium ions. If their concentration surpasses the threshold, a neuron will be triggered and the electric impulse will travel to the brain. This is how we breath or smell.
There are cilliary molecules and olfactory receptor cells on the nasal epithelium. The receptor cells are triggered by molecules and a signal is transmitted to a dedicated region inside the brain. As another example of G-proteins, consider a small adrenaline molecule traveling in the extra cellular environment. This molecule binds with a receptor. The receptor gets activated and changes its geometry. Guanosine triphosphate (GDP), a component of the G-protein, acts like a switch; teh G-protein is in OFF state as long as GDP is attached to it. As soon as adrenaline binds to the receptor, the G-protein is activated. The resulting change in geometry causes the G-protein to exchange its GDP for a GTP molecule. This GTP moves along the cell membrane and results in the production of cyclic adenosine monophosphate (cAMP) molecules. As the cAMP molecules are produced by a single adrenaline molecule, this whole mechanism is acting like an amplifier. These molecules result in the opening of ions and if this is a neuronal channel, neurons are fired. In this way, the signals are transmitted to the brain.

To understand our olfactory senses better, we need an odorant molecule. Let's choose an ester with the chemical formula CH_3CH_2COOH to serve the purpose. The geometry of this molecule is such that it fits the receptor. An electron from one side of the receptor has to cross the barrier after the G-proteins are linked up with the receptor molecules. We have to activate this molecule via switch. The G-protein has to gain an electron so that the link is broken. The other side of the receptor molecule has some excess electrons and the region is called the donor region. The electron has to cross the barrier to go to the acceptor region. The barrier ester does not conduct and is an insulator. The electron crosses the barrier by tunneling. Both inside the donor and acceptor regions the electrons are at some energy which is different from each other. A free electron enters the donor region which comes from some other molecule present nearby which is called a reducing agent. One such reducing agent is NADH which can supply electrons. The electron has energy E_1 in the donor region. The electrons through the odorant molecules tunnels towards the acceptor region having energy E_2. Why do we need an odorant molecule? The electron loses its energy while tunneling which has to be depleted somewhere and that excess energy goes to the odorant molecules. This molecule shows vibrations which are called phonons. This is an example of “Phonon assisted tunneling”.

Phonon is a quantum of vibrational energy. The energy levels for the vibration of molecules are quantized. The lower energy level represents the ground state of the molecule having $n = 0$. If the molecule is in the ground state, its energy will be

$$E_0 = \frac{1}{2}\hbar\omega_0 \tag{3.50}$$

where ω_0 is the fundamental frequency. The energy of the first excited state having $n = 1$ is

$$E_1 = \frac{3}{2}\hbar\omega_0\,, \tag{3.51}$$

If the molecule is in the ground state, we can excite it, for instance, by an incoming photon. It absorbs that photon and jumps to a higher level. Another way of exciting the molecule can be by increasing the temperature. The difference between the states is given by $\hbar\omega_0$. We define this energy gap by phonons. In the nth state, the oscillator has energy

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega_0 \tag{3.52}$$

In this state, we can say that the oscillator has $n + 1/2$ phonons. If n is very large, $n+1 \sim n$ i.e. we can say that the oscillator has n phonons, each having energy of $\hbar\omega_0$. A phonon can be treated as a particle just like a photon. If one of the electrons is excited from its energy level, it leaves a hole behind. This hole is not a real particle – it is a quasi-particle but it can be treated like a real particle. Similarly, a phonon is not a particle but it can be treated like one. If the electron is de-excited from E_1 to E_2, it will emit a phonon of the corresponding energy. This is what happens in tunneling. Hence, the process of smelling involves the process of phonon-assisted tunneling.

A photon is a packet of energy which cannot be destroyed but a photon itself can be. Phonons are not conserved – they can be created or destroyed through temperature. If the atom is excited to higher levels, new phonons are created all by themselves.

3.8 Bohr’s correspondence principle

Consider a teardrop on the tip of an eyelash. Suppose the tear drop has a mass of 1 μg and is being pulled down by gravity. This scenario can be considered analogous to a harmonic oscillator. If the spring constant k is 0.1N/m, how many phonons are excited in this process? The total downward displacement due to gravity is given by

$$x_0 = \frac{mg}{k} \tag{3.53}$$

Therefore, the energy will be

$$E = \frac{1}{1}kx_0^2 = \left(n + \frac{1}{2}\right)\hbar\omega_0 \tag{3.54}$$

We can calculate the value of n from this equation, which comes out to be $n \approx 3 \times 10^{22}$. It is such a large number that the wave-nature of quantum mechanics is washed away. This is called *Bohr's correspondence principle.*

3.8.1 Capacitors

Consider a circuit with a capacitor C and a resistor R. The capacitor is initially uncharged. If we close the circuit at $t = 0$, the current will flow through the circuit but it will be resisted by the positive end of the charging capacitor. When the voltage across the capacitor equals the voltage across the battery, the current stops flowing.

We are interested in observing the change in the capacitor charge with time. Applying Kirchoff's voltage law to the circuit

$$V_0 = \frac{Q}{C} + R\frac{dQ}{dt}$$

rearranging this a little

$$\frac{dQ}{dt} = \frac{1}{R}\left(V_0 - \frac{Q}{C}\right)$$

Multiplying both sides by an integrating factor we get

$$\frac{d}{dt}\left(Qe^{t/RC}\right) = \frac{V_0}{R}e^{t/RC}$$

integrating this equation we get

$$Q = V_0C + Ae^{-t/RC}$$

At $t = 0$, $Q(t = 0) = 0$ so $A = -CV_0$. Putting it back, we get

$$Q(t) = CV_0(1 - e^{-t/RC}) \tag{3.55}$$

This is the maximum charge the capacitor holds at a particular time and is shown in Fig. 4.8. It shows that the capacitor does not immediately charge to CV_0 and requires time. At $t = RC$ the capacitor is charged to 63% of its maximum value. This RC is called the time constant.

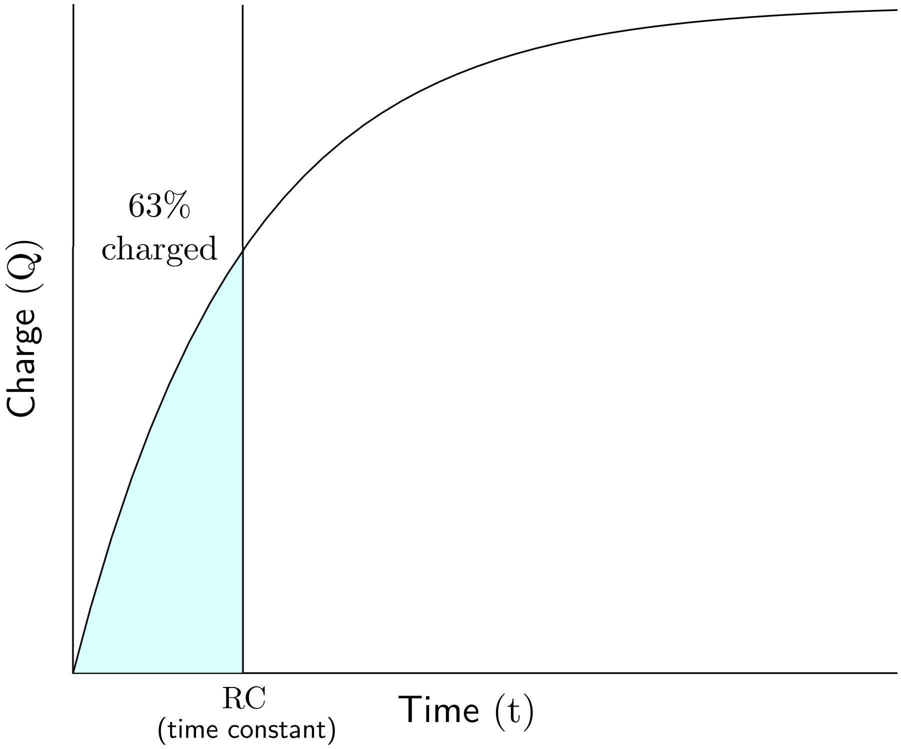


Figure 3.8: Plot of charge against time

Now, let's open the switch and see what happens. The equation in this discharging condition is

$$\frac{Q}{C} + \frac{dQ}{dt}R = 0 \tag{3.56}$$

By solving this differential equation, we get

$$Q = Q_0 e^{-t/RC} \tag{3.57}$$

This is an exponentially decreasing curve.

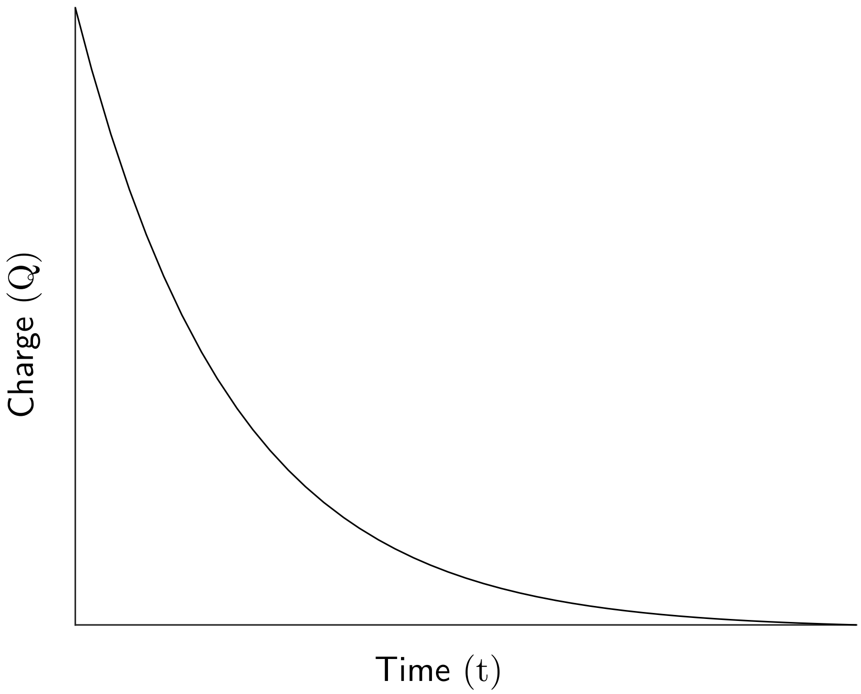


Figure 3.9: Exponentially decreasing curve

3.9 The quantum revolution in electronics

Transistors were invented in 1947 by Shockley and Bardeen and can act like switches and amplifiers. There are different kinds of transistors. We will briefly discuss a specific type known as field-effect transistors (FETs).

In a metal-oxide-semiconductor field-effect transistor (MOSFET), we have a slab of silicon, which is a group 4th semiconductor. We know that energy levels are quantized. If we have free electrons and positive ions in the solid, the energy levels are bunched up to make an energy band. One bunch of levels is separated from the other by some gap called a forbidden or energy band. The electrons are filled according to the Pauli exclusion principle. The lower band is called the valence band and the upper band is called conduction band. If the valence band is completely filled and the conduction band is empty, then such a material is called an insulator. If we have electrons at the top of valence band and we provide it with energy by an incoming photon, they can go to the conduction band. Then, an electron-hole pair is created and the material becomes a semiconductor. Even room temperature is sufficient to start conduction. If we increase the temperature of a semiconductor, the conductivity goes up whereas for the insulator it goes down. When the excited photon comes back a photon is emitted which is the working principle of a light-emitting diode (LED). Converse is the working principle of a solar cell.

A MOSFET is made of a silicon chip which is a very abundant and pure element. We can purify the silicon to such an extent that the impurity is less than 1 atom in 10^8 silicon atoms . This transistor has two metal electrodes made of gold. These electrodes have wires known as interconnects. One of the electrodes is called a source and the other is called a drain. On top of the silicon we have an insulator which is a metal oxide such as silica. On top of it we have another electrode known as the gate. The thickness of the gate defines the kind of technology under use. Classically, no electrons can go from the gate to the silicon and vice versa. This device is kept at room temperature. The electrons are spread out in the silicon – the higher the temperature, the higher will be the density of these electrons. Now we apply a voltage to the source. Suppose it is a negative voltage and to the drain we apply a positive voltage. Electrons should flow from the source to the drain but can they flow through the silicon to make up the current? They cannot, because no conduction can take place through it. This transistor is now in the OFF state. If we want to make it go to the ON state, we would need to apply a voltage of 10 volts to the gate. We apply this potential difference to the source and drain. The electrons in the silicon will be attracted to the vicinity of the gate. Free electrons are already available in the source and now, due to repulsion, the conduction channels open up between source and drain. By switching the gate voltage ON and OFF we can turn the current ON and OFF.

If the number of electrons controlling the transistor is very high, power dissipation will also increase accordingly. In 1970 when these ICs were made, there were 10 million electrons which made the transistor switch. In 2013, only $5 - 10000$ electrons were able to switch it. This means that the number of electrons controlling the transistor are reducing day by day, resulting in less power dissipation and faster operation. Currently a microchip of the dimensions of 1in^2 squared has 100 million transistors and each one is switching in 100 millions per second. This, however, is not desirable and we want to decrease it. This is where quantum mechanics comes to the rescue.

Flash memory cell

A flash memory cell is similar to MOSFET. It consists of a source and a drain and there are two electrodes on a metal oxide wafer. The top electrode is the gate while the lower electrode is the floating electrode. To study the working of the flash memory cell, let's apply a voltage to its source and drain. If a voltage is applied to the gate, a conduction channel known as an N-channel opens up. When the number of conduction electrons is high, they can tunnel through the metal oxide to the gate because the channel width is small ($l = 1$mm). Now, if the gate voltage is turned OFF, the conduction electrons fade away. However, the tunneling electrons remain there.

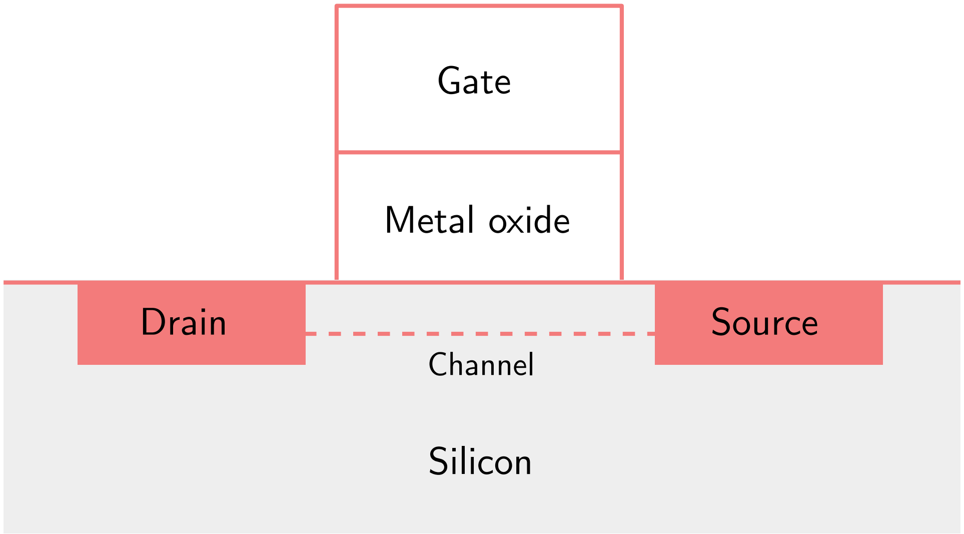


Figure 3.10: Flash memory cell

Let's write something onto the memory cell. This process is carried out through quantum mechanical tunneling. These cells are non-volatile devices because turning OFF the power does not change the state of this cell. We can also erase the information stored in these cells if we apply a positive voltage to the drain and a ground voltage to the source and gate. In this case, the electrons will now tunnel through the metal oxide to the drain.

3.9.1 Single electron transistor

Once again, we have the three electrodes: a source, drain and gate, all of them having voltages. Instead of electrons, we now have a quantum dot inside the chip. We would like to put one electron in the quantum dot or quantize the current because charge is quantized. We would like to have tunneling on purpose for this to happen. If the quantum dot is neutral, its electrostatic energy is zero. If we put one electron in the quantum dot and then another, is it easy for the second electron to go in? The first electron will repel the incoming electron, so adding the electron to the quantum dot increases its energy. If this quantum dot acts as a capacitor, its electrostatic energy would look like

$$E_C = \frac{e^2}{2C_{dot}} \tag{3.58}$$

This increase in energy acts like a potential barrier. This effect is known as Coulomb blockade. If we have large temperatures electrons from the source and gate can tunnel to the quantum dot. We would like to make the temperature smaller so that the electrons cannot overcome the Coulomb barrier or we can also write this condition as

$$10k_BT < \frac{e^2}{2C_{dot}}$$

At this condition we can control the tunneling. The capacitance of a quantum dot is given by

$$C_{dot} = 2\pi\epsilon d\,, \tag{3.59}$$

where ϵ is the conductivity and d is the diameter of the dot.
The second condition to control tunneling is given as

$$\Delta E < E_C\,, \tag{3.60}$$

where ΔE is the uncertainty in energy. If it is so large that it is in the order of the charging energy E_C, then the charging energy is as good as nothing. If this condition is not met, then electrons will tunnel to and from the dot. From the uncertainty principle we know that

$$\Delta E \sim \frac{h}{2\Delta t} = \frac{h}{2R_tC_{dot}}\,, \tag{3.61}$$

where R_t is the tunneling resistor, and if Δt is large then the ΔE should be small. Writing the second condition again by putting these values, we get

$$\Delta E = \frac{h}{2R_tC_{dot}} < \frac{e^2}{2C_{dot}} \Rightarrow R_t \simeq 26k\,\Omega\,. \tag{3.62}$$

Electrons in the source and the drain are filled upto the Fermi energy. In between these, we have the quantum dot and the barrier is the Coulomb blockade. Suppose there is an energy level inside the dot which is coincident with the Fermi level. If an electron is in the source can it tunnel to the dot? It can, but when it happens, the energy of the quantum dot increases which is against the conservation of energy and is classically forbidden. So the electron has to pay the price to enter the quantum dot. The drain is moved downwards because of the positive voltage applied to it. If a positive voltage is now applied, the the energy levels in the quantum dot will decrease. Now, the electron can tunnel through the barrier by paying the price which is met at the condition

$$V_g e = \frac{e^2}{2C_{dot}}\,, \tag{3.63}$$

At this condition, the quantum dot increases its number of electrons by 1 only.

Chapter 4

Double slit experiment

Richard Feynman stated that double slit experiment captures all of the mysteries of quantum mechanics. The experimental setup is as follows.

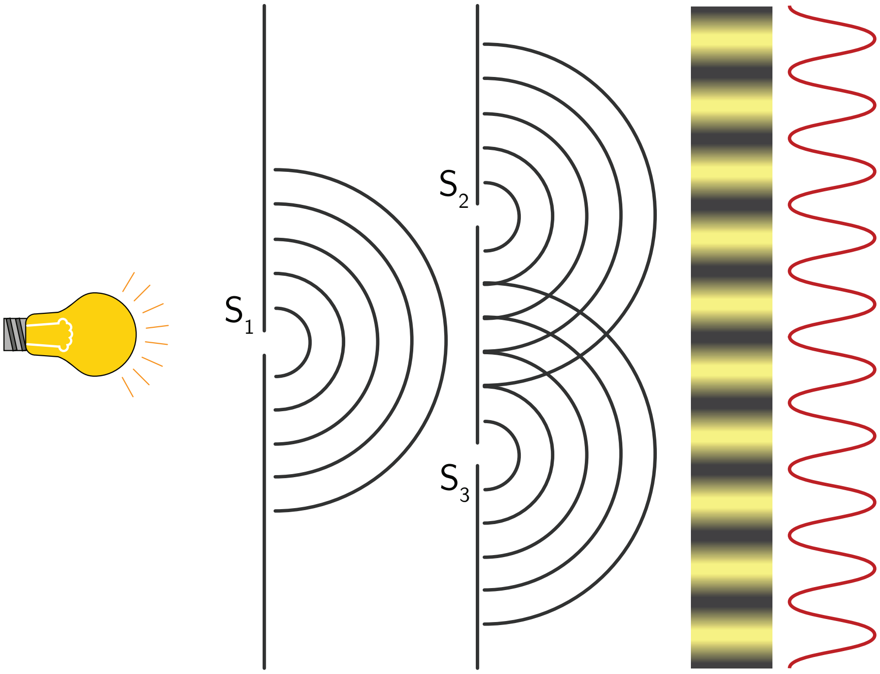


Figure 4.1: Young's doulbe-slit experiment

We have a coherent source of light. Light falls from this source onto a screen that has one slit in it, this is followed by another screen that has two slits. After the screen with two slits, we have another screen that we use to observe the pattern that the light makes after passing through the two slits.

Light is an electromagnetic wave, and waves interact with each other. Light waves passing through the two slits interact and we observe fringe formation on the screen. The graph of light intensity against the position on the screen is called the interference pattern.

If instead of waves we were sending particle through the slits (bullets let's say), we would not get an interference pattern, but just two humps. Closing one slit will give us one hump and closing the other will give us another hump at a different position on screen. Opening both the slits simultaneously will give us a sum of these two humps but not an interference pattern. This is because particles do not interfere.

4.1 Quantum realm

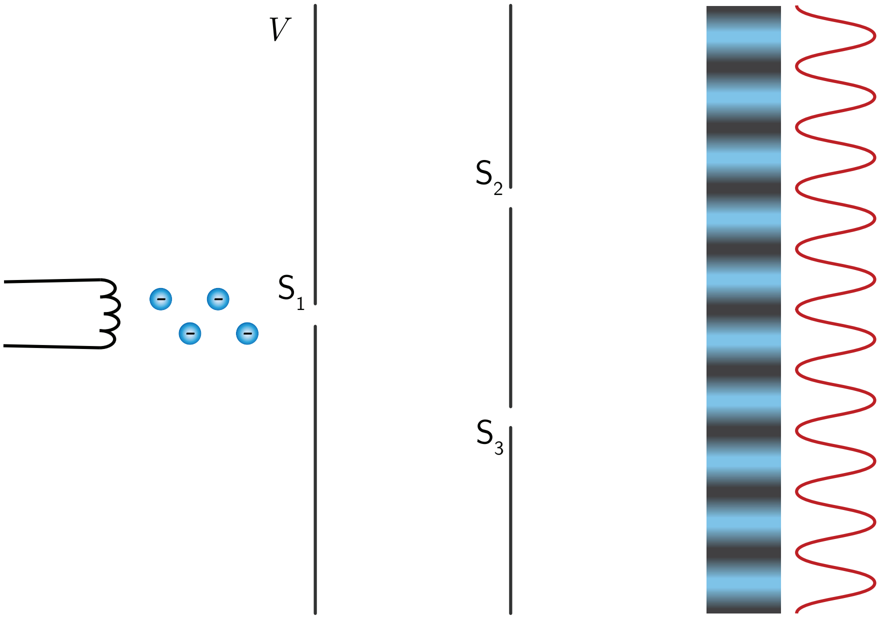


Figure 4.2: Young's doulbe-slit experiment with electrons instead of light

We alter the previous experiment a bit and instead of a light source we put a source that produces electrons using thermionic emission. It consists of a filament that heats up when current passes through it. When the temperature gets large enough, the electrons come out from inside the metal to the surface and evaporate out of it.

We accelerate these electrons by providing a potential difference between the filament and the screen containing slit. Kinetic energy gained by the electrons will be

$$E = Ve$$

Where V represents potential, and e represents elementary charge.

When the electrons pass through the screen with two slits and fall on the detection screen coated with Zinc Sulfide, we see the interference pattern instead of two distinct humps. This reveals a strange phenomenon that electrons that we are used to consider as particles are behaving as waves in this experiment. This idea was first proposed by deBroglie when he said that there is a wave associated with each and every particle, big or small.

Our general description of this experiment is that we have millions of electrons passing through the slits at a given time. Some of the electrons pass through S_2 while the others pass through S_3. The electrons interfere with one another and give us the interference pattern we observe. Let's change this experiment a bit more. We reduce the intensity of electrons produced to a level that a new electron is released only when the previously released electron has already reached the final screen. This means that at a time only one electron passes through the slits. One can intuitively think that since it has no other electron to interact with, it will simply pass though one of the two slits, fall on some point on the screen and we should not see a pattern. But, contrary to our intuition, the experimental outcome came out to be an interference pattern. This means that the electron passed through the two slits simultaneously, interfered with itself like a wave, and gave us the interference pattern.

But how is it possible for an electron to split? The electron has to pass through either of the two slits to reach the final screen. So is it passing through S_2 or S_3? To find this out we introduce an observation device in the setup: a light bulb. Light from the bulb bounces off the electron and it shines. A shine near S_2 means that the electron chose that slit and vice versa for S_3. By sending one electron at a time, we find that the electron actually behaves like a particle and passes through either of the two slits but not both.

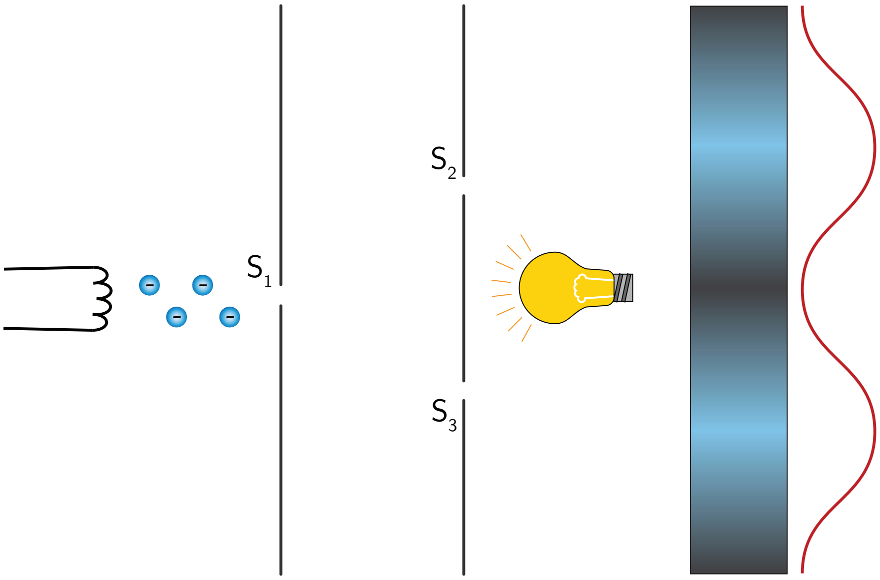


Figure 4.3: Observation device introduced in experimental setup

However, after installing this bulb in the setup, we do not get the interference pattern anymore. Instead, we see only two humps on the intensity plot. This means that the electrons started acting as particles the moment they were observed even though the observer did not interact with the electrons.

This problem occurs because introducing an observer (bulb) alters the nature of the experiment. The overall experiment has changed because of this measurement device. Even if we were focusing on a single slit, and from a very large distance (through a telescope let's say), we will still see this change of nature.

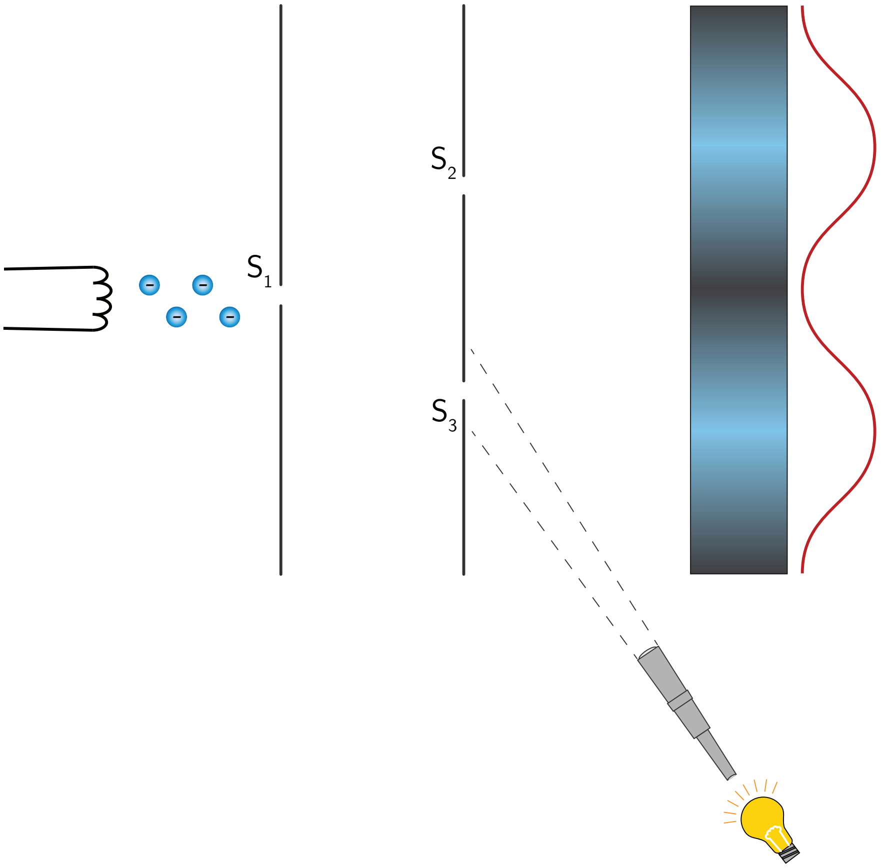


Figure 4.4: Telescope as observation device

Nature or properties of something should not change because of our choice to observe it or not. At least that's what we thought. Quantum mechanics negates that perception of ours.

4.2 Particels or waves?

In the slit region the electrons are acting as waves and energy is spread out in the whole setup; but as soon as the electron reaches the final screen, the wave collapses into a particle, and all of that energy confines into a single point giving us a dot on the screen. Before detection, there was a realm of possibilities for position on screen; Measurement confined that to a single point.

4.2.1 Copenhagen interpretation

This problem was resolved by quantum mechanics. According to Copenhagen interpretation, we do not consider an electron as either a particle or a wave. Instead, we say that it is a quantum field, represented by the Greek letter *psi* ψ.

We say that the field associated with the path of slit S_2 is ψ_1 and with S_3 the associated field is ψ_2. The overall field ψ of the electron that passes through S_1 is a superposition of these two fields.

$$\psi = \psi_1 + \psi_2 \tag{4.1}$$

A superposition is a field state that can exist in quantum interpretation of world, but not in the world how we perceive it. So our act of trying to find the electron destroys this superposition state and it collapses to either of it's two components states.

There is another interpretation of this changing nature called the *many world interpretation*. It says that another universe is created when we make a measurement. In some other universe with the exact same conditions, the observer will not see a flash.

4.3 Heisenberg's uncertainty principle

A bit of mathematics is required before this concept can be explained.

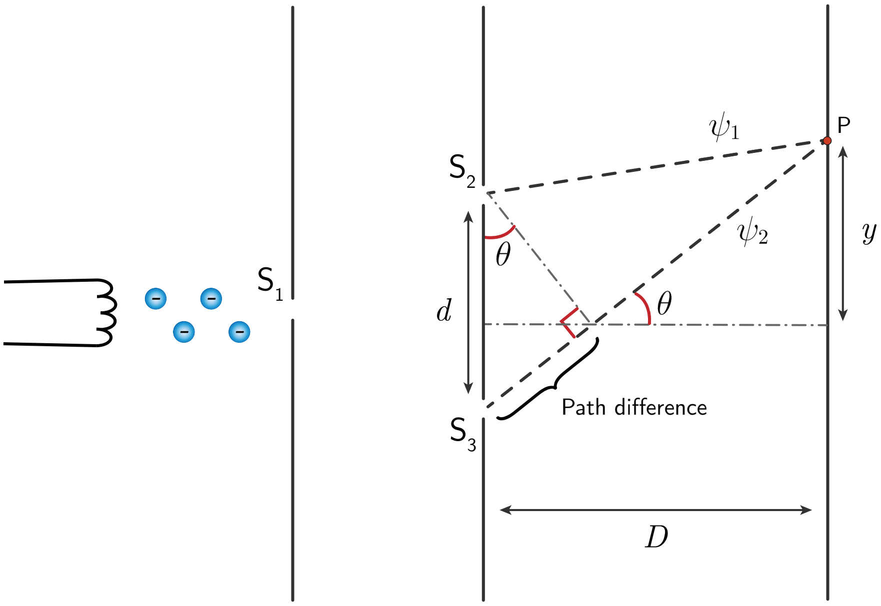


Figure 4.5: Path difference in double-slit experiment

Equation of a wave can be written as

$$e^{i(\phi)}$$

Without considering its propagation with respect to time. The factor of phi in the exponential is the phase difference in radians between the two waves that interfere given by $\phi = kx$. x is the path difference between the two waves and k is the wave number given by

$$k = \frac{2\pi}{\lambda}$$

Path difference can be calculated from the figure using simple trigonometry and turns out to be $dsin(\theta)$. Using this path difference, the phase difference becomes

$$\phi = kdsin(\theta)$$

Substituting expression of k,

$$\phi = \frac{2\pi}{\lambda}dsin(\theta)$$

We know $\lambda = \frac{h}{\rho}$. Substituting this, we get

$$\phi = \frac{2\pi}{h}\rho dsin(\theta)$$

The point at which we want to find the intensity will be dictated by the angle θ as the rest is constant.

Intensity at any point is given by amplitude squared.

$$I = A^2$$

In our experiment we have wave functions which can also be thought of as probability amplitudes like we did in the previous chapter. So, intensity at any point P on the screen is given by

$$^{1}I_P = |\psi|^2 I_P = \psi\psi^*$$

Using eq(5.1) and the fact that ψ_1 and ψ_2 are same except for a phase difference of ϕ, the overall wave function can be written as

$$\psi = \psi_1 + \psi_1 e^{i\phi}$$

and the conjugate is given by

$$\psi^* = \psi_1^* + \psi_1^* e^{-i\phi}$$

Intensity can then be written as

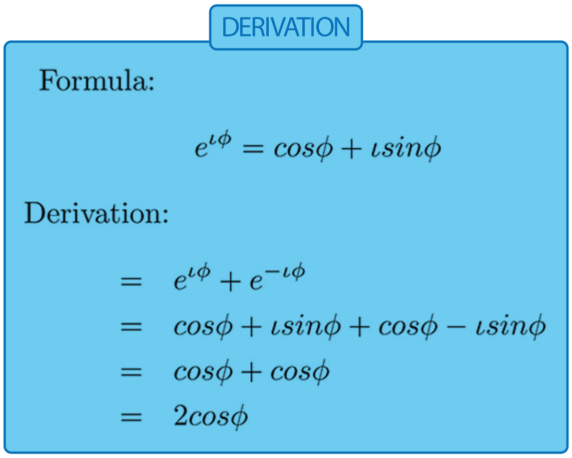
$$\begin{aligned} I_P &= (\psi_1 + \psi_1 e^{i\phi})(\psi_1^* + \psi_1^* e^{-i\phi}) \\ &= \psi_1\psi_1^*(1 + e^{i\phi})(1 + e^{-i\phi}) \end{aligned}$$

Let $\psi_1\psi_1^* = c$ as it is a constant value.

$$\begin{aligned} I_P &= c(1 + 1 + e^{i\phi} + e^{-i\phi}) \\ &= c(2 + 2cos(\phi)) \\ &= 2c(1 + cos\phi) \\ ^{1}I_P &= 4c\left(cos^2\frac{\phi}{2}\right) \end{aligned}$$

---

[1] Modulus sign is used because ψ is, in general, a complex number



DERIVATION

Formula:

$$e^{\iota\phi} = cos\phi + \iota sin\phi$$

Derivation:

$$\begin{aligned} &= e^{\iota\phi} + e^{-\iota\phi} \\ &= cos\phi + \iota sin\phi + cos\phi - \iota sin\phi \\ &= cos\phi + cos\phi \\ &= 2cos\phi \end{aligned}$$

Figure 4.6: Derivation of possibly confusing step

Plotting this equation will give us the intensity pattern that we observe on the screen.

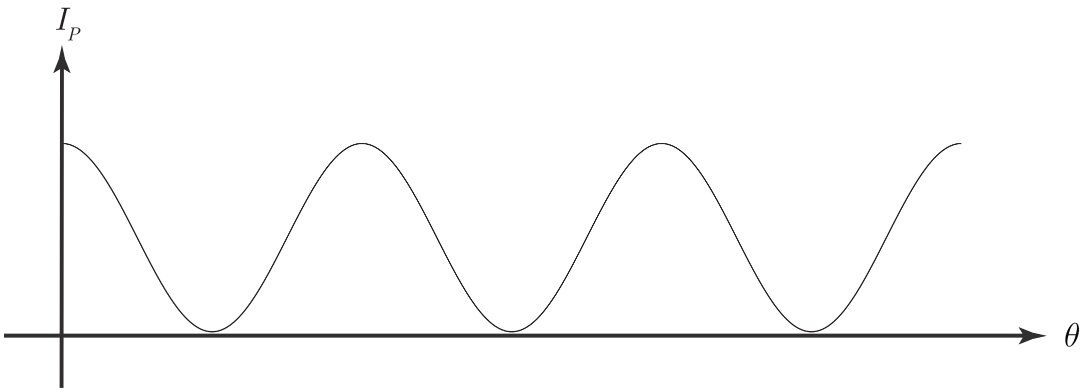


Figure 4.7: Intensity against angle θ

If one wants to plot the graph of intensity against the displacement from mean position y, it can be achieved using the following calculations.

$$\phi = kdsin\theta$$

For small $\theta, sin\theta \approxeq \theta \approxeq tan\theta$.

$$tan\theta = \frac{y}{D}$$

So,

$$\phi = \frac{kdy}{D}$$

Substituting in eq5.2

$$I_P = 4c\left(cos^2\left(\frac{\pi dy}{\lambda D}\right)\right)$$

Here we get a minimum when the angle is a multiple of $\frac{\pi}{2}$, i.e. $y = \frac{n\lambda D}{d}$ where n is an integer.

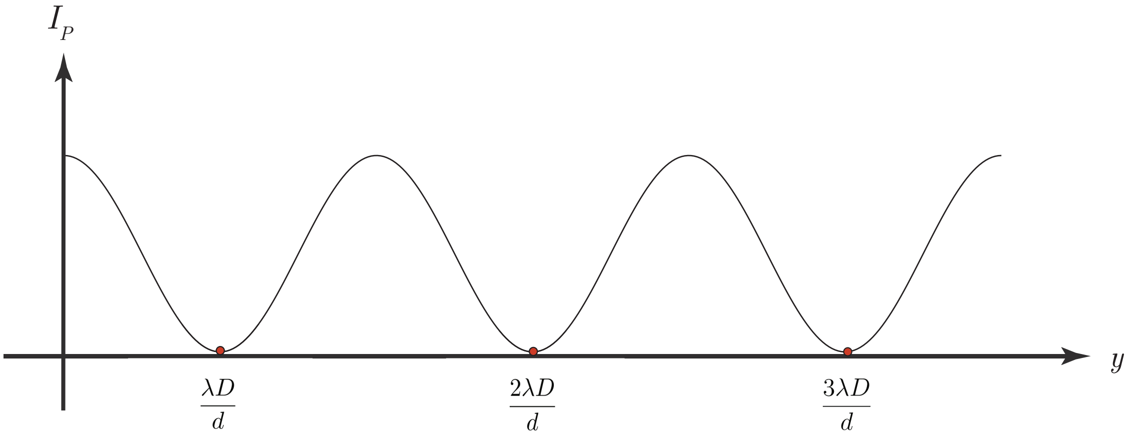


Figure 4.8: Intensity against position y

The uncertainty Δy in the path that the electron chooses depends on the distance between the slits. Farther apart the slits are, greater will be the uncertainty in path.

$$\Delta y \approx a$$

Let's reduce the distance d between the slits and see how it effects our graph; considering only the central maxima for now.

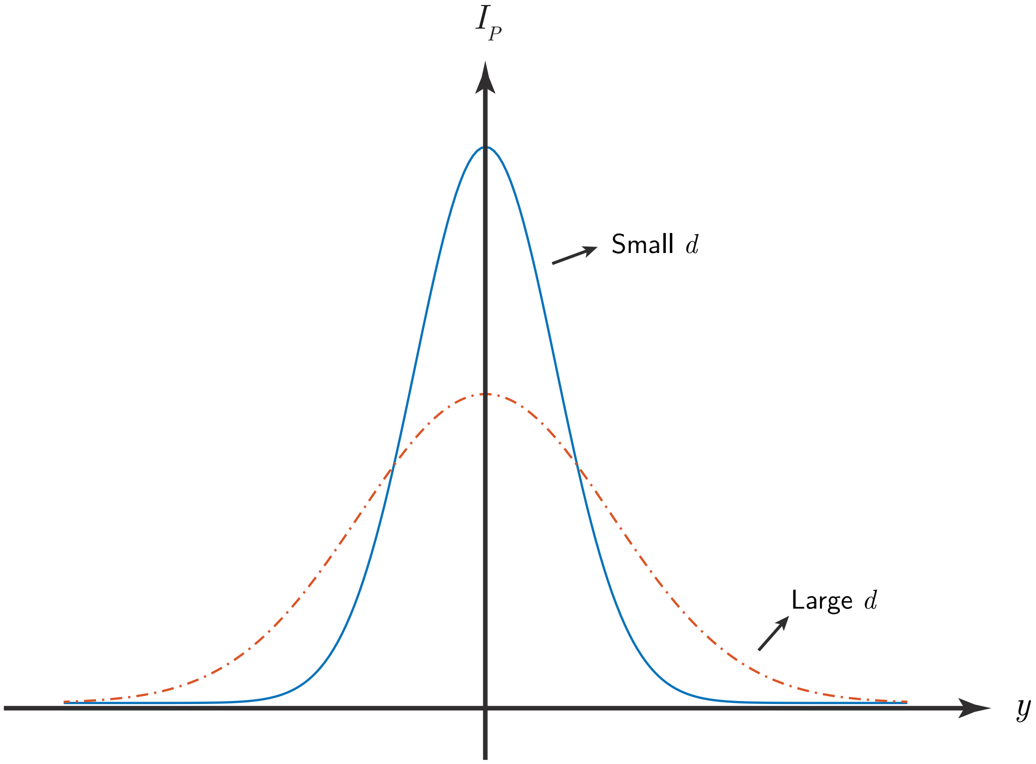


Figure 4.9: Intensity against displacement y from mean position

Decreasing the slit separation d, increases the distance where first minima appears. Intensity must decrease if that happens because the area under the curve corresponds to the number of electrons and it should stay constant.

The spread in the intensity shows that there is some uncertainty associated with what direction the electron moves in after passing through the slits. Motion of a particle is related to momentum. So, there is a whole range of momentum vectors that can be attained by the electron.

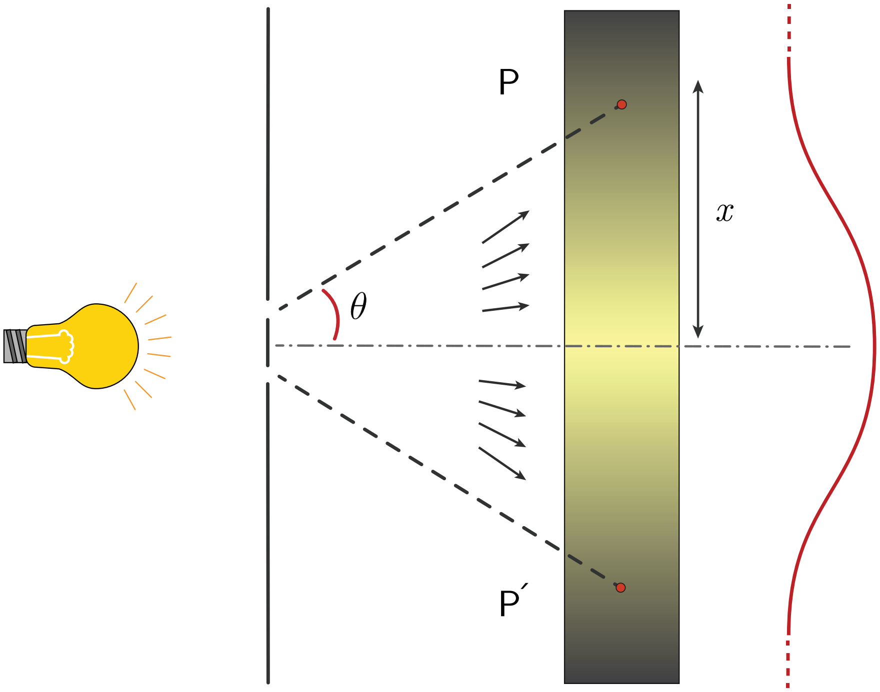


Figure 4.10: Uncertainty in momentum

The uncertainty in momentum $\Delta \rho_y$ depends on the spread of the peak i.e. $2\rho sin\theta$.

$$\Delta \rho_y \approx 2\rho sin\theta$$

Product of the two uncertainties turns out to be

$$\Delta \rho_y \Delta y \approx 2d\rho sin\theta$$

We know that $\rho = \frac{h}{\lambda}$, so

$$\Delta \rho_y \approx \frac{2dh}{\lambda} sin\theta$$

This can be simplified a bit more. We see that a single hump like this central maxima of ours can be seen even if we have a single slit. This can be explained if we consider our single slit as a combination of two smaller slits with a slit separation of $\frac{a}{2}$.

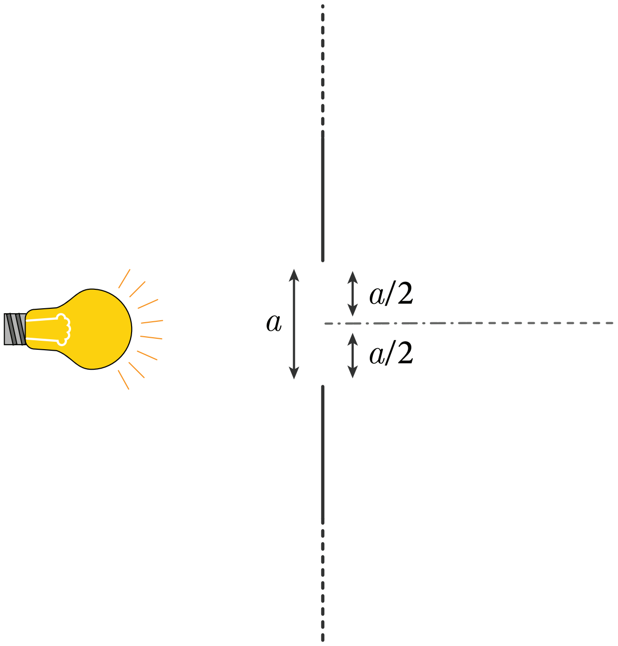


Figure 4.11: Magnification on a single slit

At the ends of this peak we see that the brightness is low, so destructive interference must have taken place. Condition for which is that the phase difference must be $\frac{\pi}{2}$. So,

$$\begin{aligned}\phi &= kdsin\theta \\ \frac{\pi}{2} &= \frac{2\pi}{\lambda}\frac{a}{2}sin\theta \\ \frac{1}{2} &= \frac{asin\theta}{\lambda}\end{aligned}$$

Substituting this result in eq5.3, we get

$$\Delta x \Delta \rho_x \approx h$$

This is where Heisenberg's uncertainty principle comes from.

Heisenberg's uncertainty principle says that the product of uncertainty in position and uncertainty in momentum has a minimum value it can not go below no matter how precise our measuring instrument is. This minimum value is given by $\frac{\hbar}{2}$.

$$\Delta x \Delta \rho_x = \frac{\hbar}{2}$$

Where $\hbar = \frac{h}{2\pi}$.

It should be kept in mind that these uncertainties are quantum uncertainties. When we flip a fair coin, we say that it will land on either heads or tails with equal probability; but we can improve this prediction if we do some extra measurements and calculations like using the force of our initial flick, the strength of wind flow, the angle at which the coin leaves the hand and so on. There is no minimum value of uncertainty that restricts us from making our prediction better. Such is not the case with quantum uncertainties.

Suppose that we have an electron in our room moving from one wall to the other in x-direction. Complete information of this electron is given by

$$\psi = \psi(x,t) = \mathrm{A}e^{i(kx-\omega t)}$$

Real part of this wave function will be $\mathrm{A}cos(kx - \omega t)$

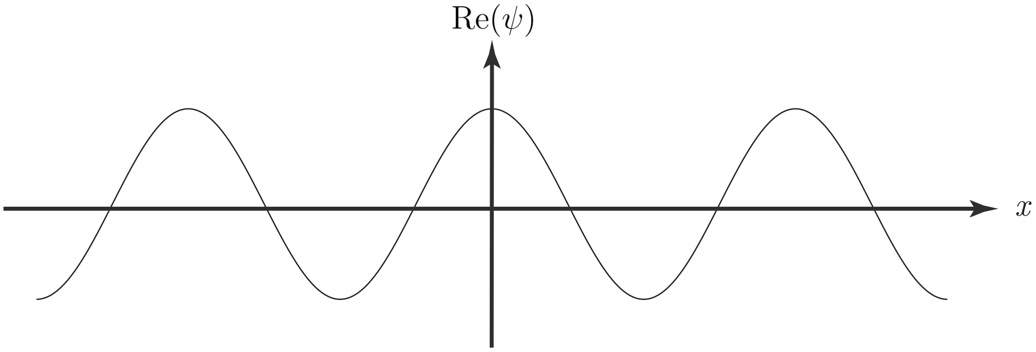


Figure 4.12: Real part of wave function

Note that this time the wave is not stationary in time and at some other time it will have displaced.

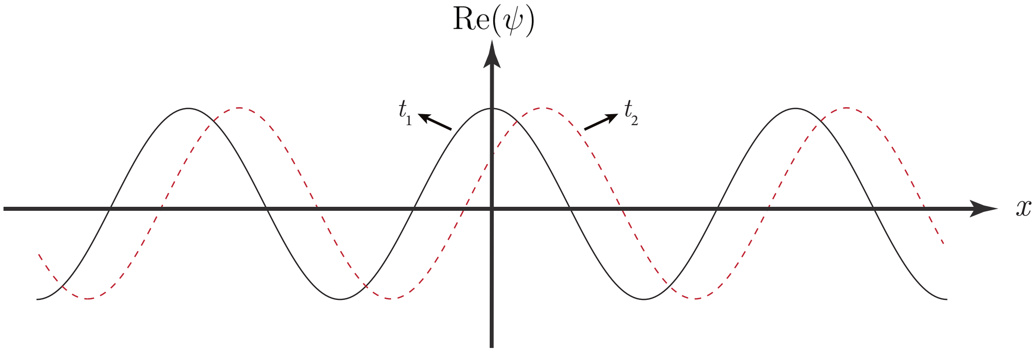


Figure 4.13: Real part time-varying

Now, we are posed with the question *"Where is this electron?"*. It might seem a bit too simple considering the problems you have already solved in mechanics. We had objects attached with pulleys, sliding on planes etc. and using Newton's laws we could calculate where our object of interest was at any instant of time.

However, simple it may seem at first glance, we can not determine the exact position of this freely moving electron in our room. Rather, we can only propose a realm of possible locations where it might turn out to be if a measurement is made. The act of measurement collapses the wave function to a single point out of those realm of possible points and gives it a new state given by $\delta(x - x_i)$ [2]. This means if we perform subsequent measurements of position on this very electron again and again, we will always find it at x_i. However, if we take another electron, keep all the conditions same as were for the first one, and perform a measurement, we will find the electron at a different position than what was determined in the previous try. Every new electron whose state we have not collapsed yet, will land at a different position even though they face the exact same conditions.

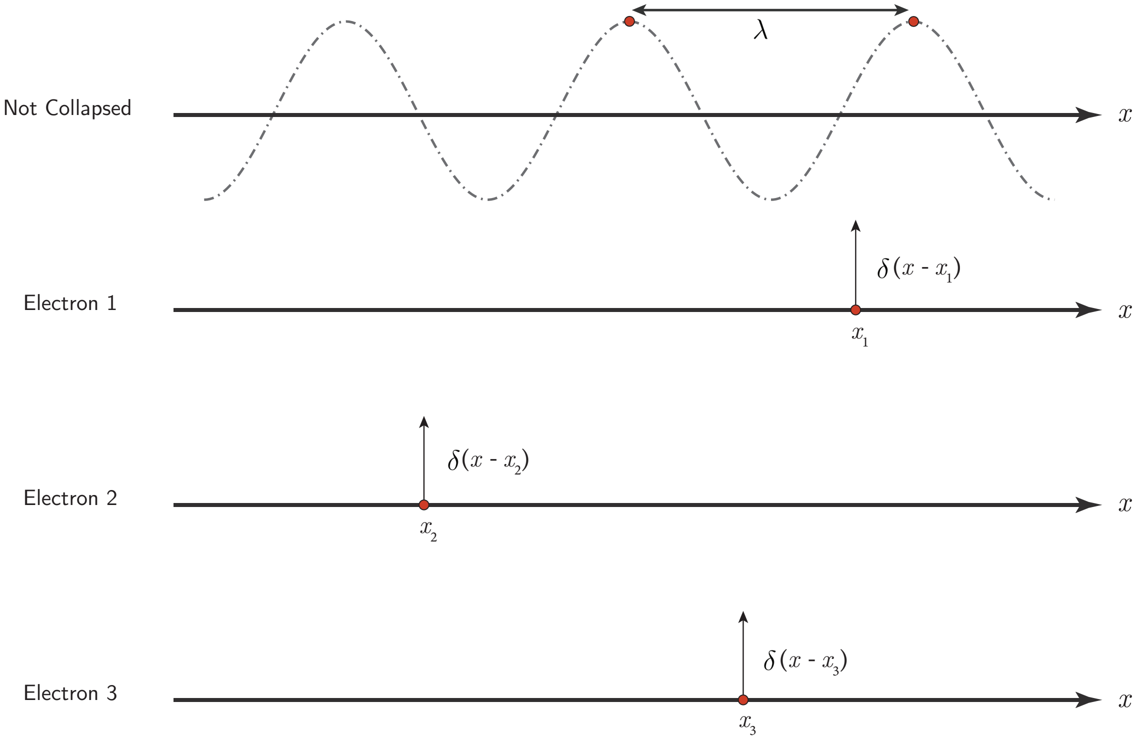


Figure 4.14: Similar electrons collapsing at different positions

Before we made the measurement, we knew what was the wavelength of our function and using it we could find the momentum, $\rho = \frac{h}{\lambda}$. But once we

---

[2] This is a special kind of function called a *Dirac Delta* function. Its value is non-zero only when the input parameter is zero itself. In this case we will only get a value when $x = x_i$

have made the measurement, we know exactly where we have confined our electron and lose all the information about the wavelength of our function. Hence, we become completely clueless about the momentum of our electron. Getting information about one of the two variables destroys information about the other. There is a trade-off between position and momentum.
If $\Delta x \to 0, \Delta \rho_x \to \infty$.
If $\Delta \rho_x \to 0, \Delta x \to \infty$.

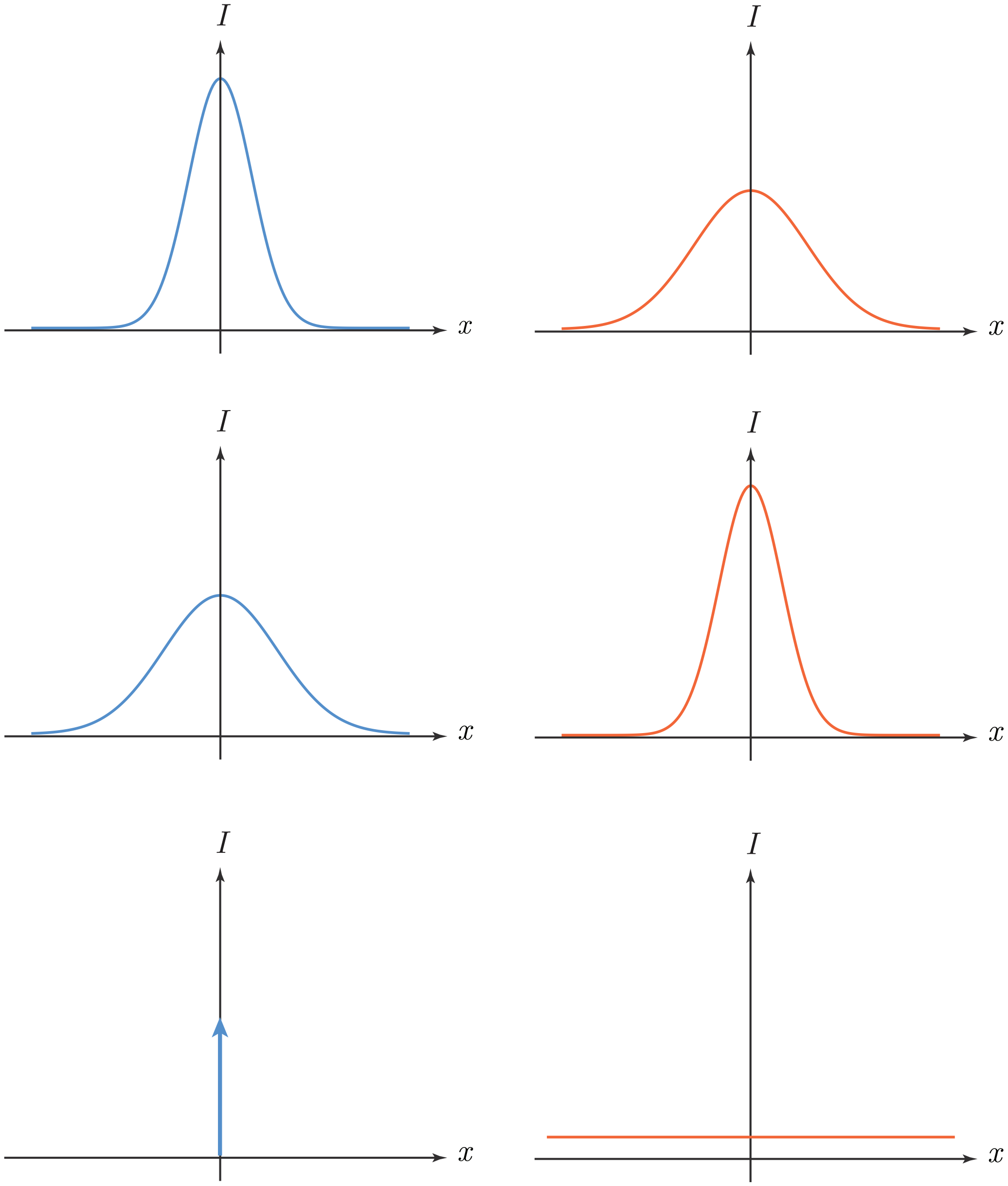


Figure 4.15: Trade-off between position and momentum

4.3.1 Uncertainty Diagram

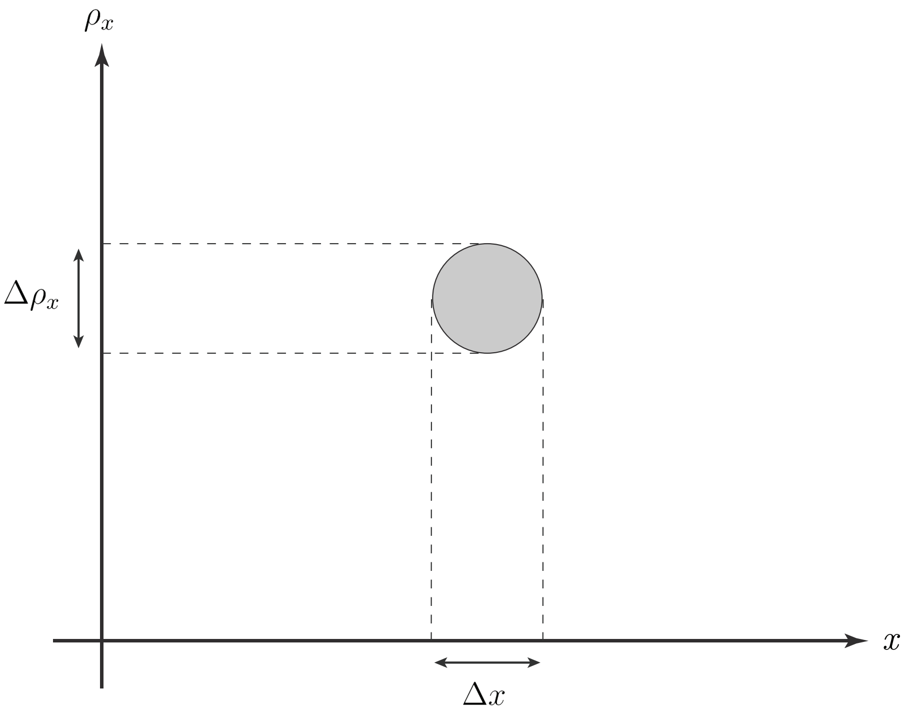


Figure 4.16: Uncertainty diagram

For uncertainty diagram, we get take the values of Δx and $\Delta\rho_x$ using full width at half the maximum height of Gaussian distribution of the uncertainty.

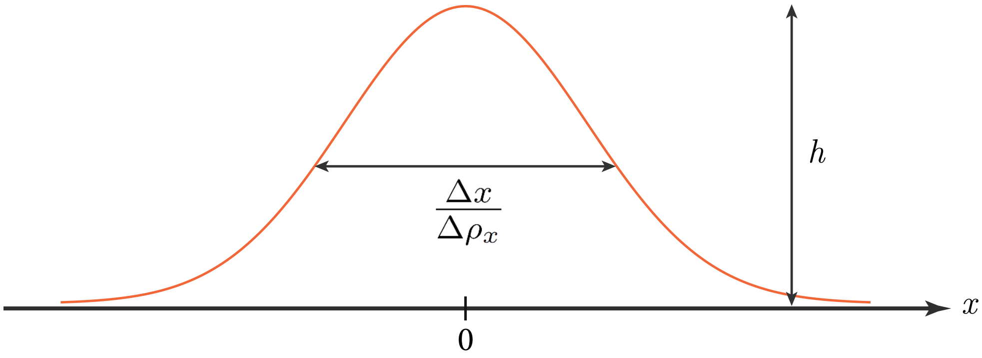


Figure 4.17: Gaussian distribution

We now know that the act of measurement collapses the wave function and uncertainty in position should be zero, turning it into a Dirac delta function (figure 3a). However, due to the lack of precision of our equipment (classical uncertainty), we see a little spread in uncertainty diagram instead of a perfect straight line, giving us Fig. (5.18(a)) instead of Fig. (5.18(b)).

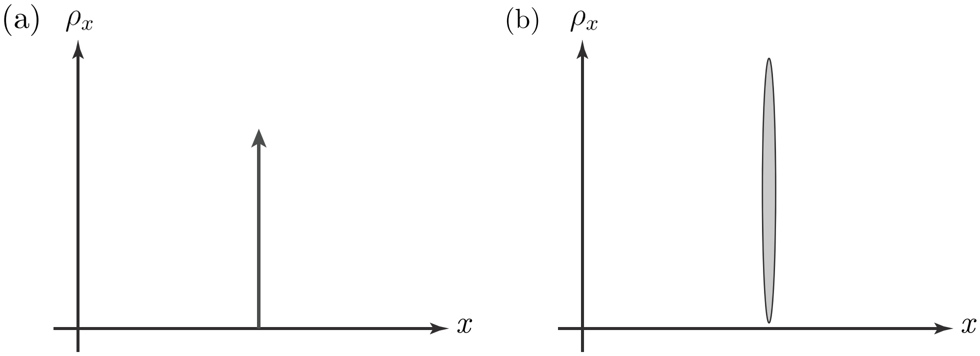


Figure 4.18: Uncertainty theory diagram vs experimental result

4.3.2 Another form of uncertainty principle

So far, the form of uncertainty principle that we have seen is the relationship between position and momentum

$$\Delta x \Delta \rho_x \geq \frac{\hbar}{2}$$

This equation can be manipulated to describe the uncertainty relationship of some other parameters as well e.g. energy-time uncertainty principle. Let's consider an electron that has an uncertainty in momentum $\Delta\rho_x$. Mass of an electron is definite and we can write this uncertainty as

$$\Delta \rho_x = m \Delta v_x$$

If we have to observe this electron, we shine light on it. The photon that comes to interact with the electron and bounce off of it, has a time window of Δt to do so

$$\Delta x = v_x \Delta t$$

Energy of an electron can be written as (considering kinetic energy only because electrons are very light)

$$E = \frac{1}{2} m {v_x}^2$$

Written in terms of momentum, this becomes

$$E = \frac{\rho_x^2}{2m}$$

To find the uncertainty in energy, we take partial derivative on both sides.

$$\begin{aligned}
\Delta E &= \frac{\rho_x}{m}\Delta\rho_x \\
\Delta\rho_x &= \Delta E\frac{m}{\rho_x} \\
\Delta\rho_x &= \frac{\Delta E}{v}
\end{aligned}$$

Plugging these in the uncertainty relationship, we get

$$\begin{aligned}
\Delta x\Delta\rho_x &= v\Delta t\frac{\Delta E}{v} \\
&= \Delta E\Delta t \geq \frac{\hbar}{2}
\end{aligned}$$

Intuitively, this can be thought of like this: an object with higher energy is moving faster and more time is required to observe what it is doing, while it is easier to observe a slow moving object and requires less time. This new form is just as important as the previous one. It's just a matter of application that we choose one or the other.

4.3.3 Applications of Uncertainty Principle

1. Measuring the energy of a photon (Light)

For this application we use monochromatic source of light, a Laser. We send a pulse of light at a spectrometer that gives us the frequency of the light that it received. The time period for which the laser is sent to the spectrometer is varied using a shutter placed in the laser path.

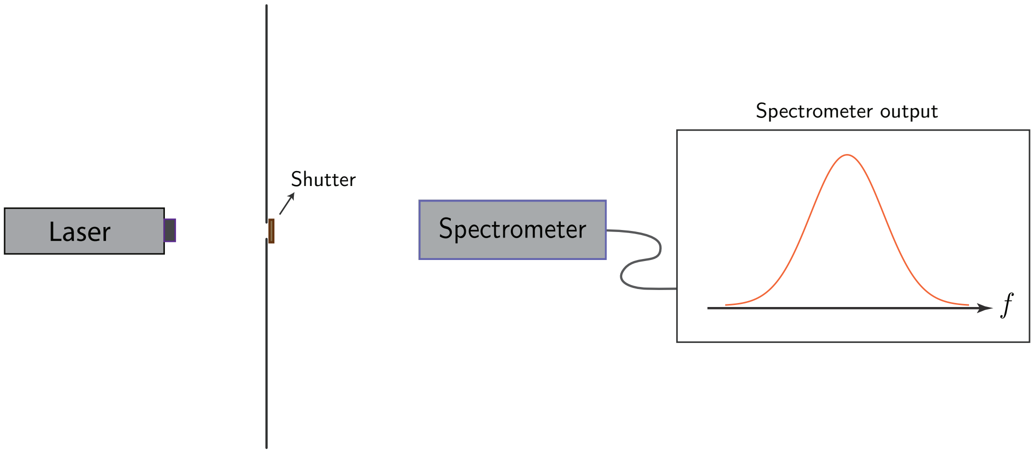


Figure 4.19: Laser observed by spectrometer

If we keep the shutter open time small and repeat the experiment with several pulses, we will observe that the spread in the frequency Δf displayed by the spectrometer is larger than what we get if the shutter is opened for a larger interval of time.

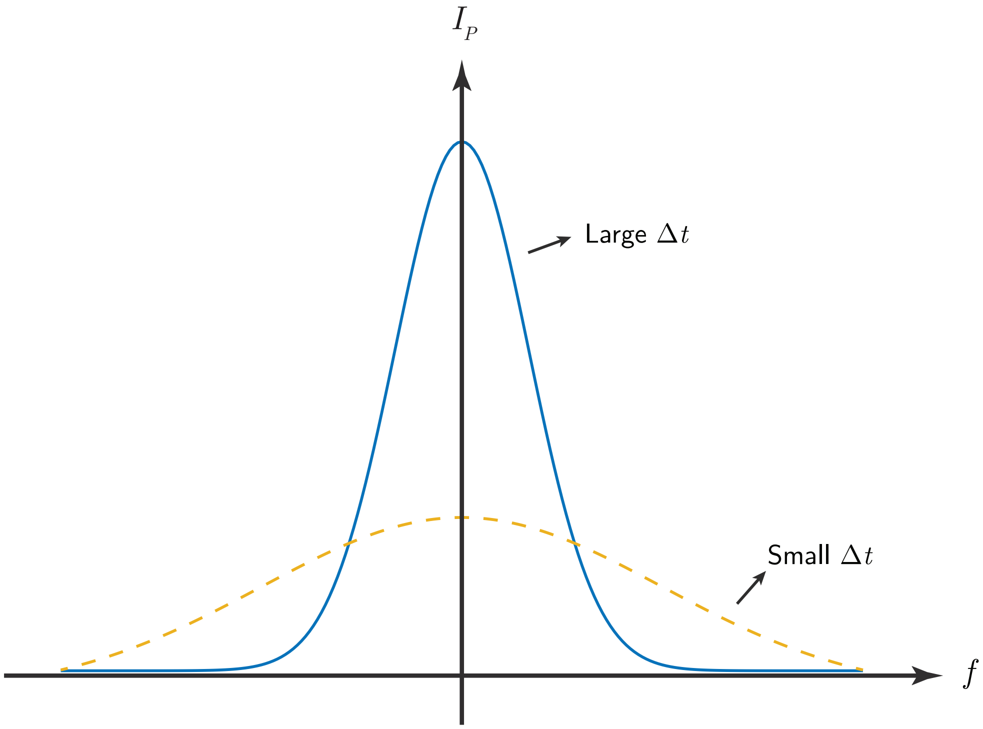


Figure 4.20: Relation of Δt and Δf

This spread or uncertainty in frequency is directly related to energy

$$\Delta E = \lambda \Delta f$$

But why do we get this spread in frequency when we send in pulses to the spectrometer? This happens because when we increase the sample time, we get more information about the wave nature and can easily determine its wavelength. Lowering this sample time creates some blank portions in our wave plot/information about the wave. These blank portions are what increase our uncertainty about the wavelength of sample received.

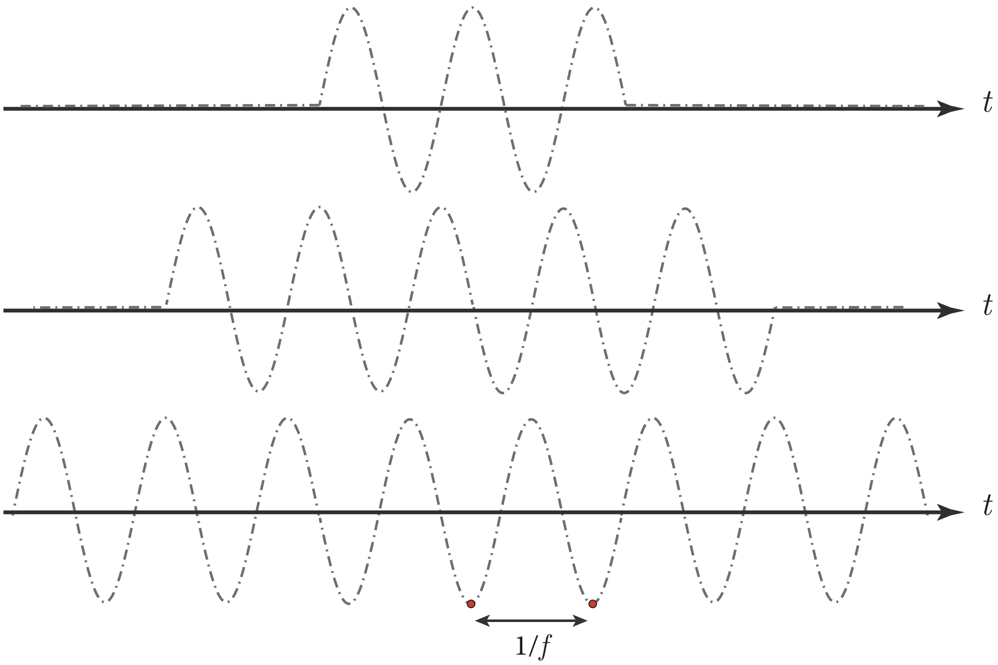


Figure 4.21: Different samples of light

Even if we remove this shutter, we do not get perfect information about energy. This is because we have an in-built shutter in the electrons themselves. We saw in chapter 2 lasers' section that the electrons stay in meta-stable state for some time known as the lifetime of that excited level. Pulses are generated after every one lifetime when the electron transitions to the lower level. More the time an electron gets to make transition (higher Δt), the more precise frequency it will release (lower Δf).

2. Bohr's orbitals

In 1913, Bohr proposed his model of a hydrogen atom. In this model he said that the electrons move around the nucleus in quantized circular orbitals, and each orbital has a quantum number n associated with it. Distance of each orbital from the nucleus is given by r_n, and the electrons are bound to stay in these quantized orbitals.

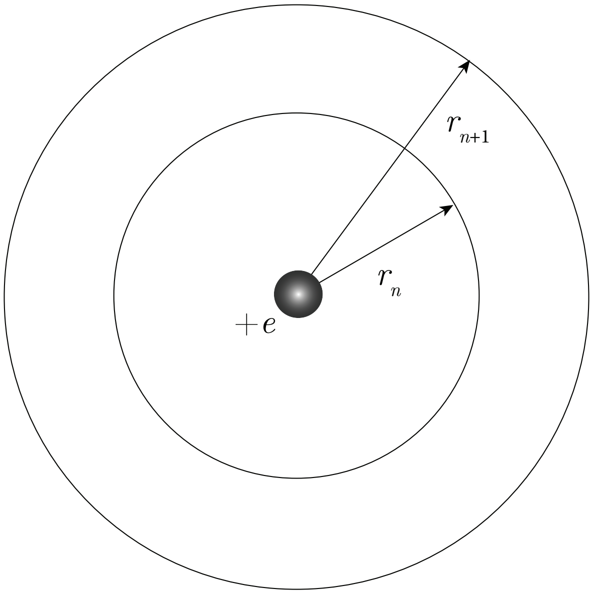


Figure 4.22: Atomic orbitals

But what if we somehow try to confine the electron in the region between two orbitals? What stops us from doing that? Let's see. If we try to confine the electron in between the orbitals r_n and r_{n+1}, the uncertainty in the position will be of the order of the difference between the two orbitals.

$$\Delta r = r_{n+1} - r_n$$

There will also be an uncertainty in momentum according to the Heisenberg's uncertainty principle

$$\Delta\rho_r \approx \frac{h}{\Delta r}$$

This will then result in uncertainty in energy given by

$$\Delta E = \frac{\rho_r}{m}\Delta\rho_r$$

Plugging in the value of $\Delta\rho_r$ we get

$$\Delta E\Delta r\ \frac{\rho_r h}{m}$$

This shows that when we try to confine the electron in the region between two orbitals (reducing Δr), the uncertainty in energy ΔE will become large. This means that the minimum value of energy E_{min}, which was previously zero, is also uncertain by an amount ΔE now. This makes the overall energy to go up and enough for the electron to get rid of the hold of nuclear force.

This is the reason for why the states between orbitals are highly unstable and difficult to occupy.

Figure 4.23: Increase in minimum energy

Let's quantify these uncertainties. Centripetal force on the electron moving in an orbital is given by $\frac{m{v_n}^2}{r_n}$ and this is provided by the nuclear pull on the electron given by Coulomb's law of attraction

$$\frac{m{v_n}^2}{r_n} = \frac{e^2}{4\pi\epsilon_0 {r_n}^2} \tag{4.2}$$

With every electron there is an associated wave. The orbital where the electron can exist has to be such that this wave can sustain in it, it only happens when the size of orbital is appropriate for the wave to constructively interfere when it loops around the circular path.

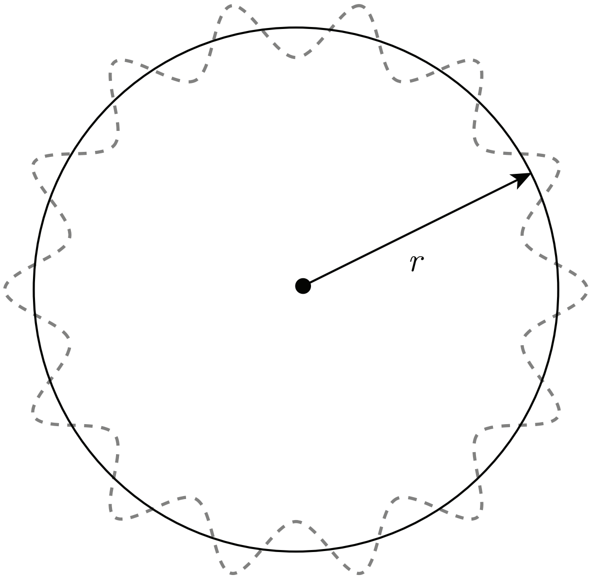


Figure 4.24: DeBroglie wave sustained in an orbital

This requires an integer number of wavelengths to be accommodated in the path

$$2\pi r_n = n\lambda$$

usnig $\rho = \frac{h}{\lambda}$

$$2\pi r_n \rho_n = nh$$

substituting $\rho = mv$

$$2\pi r_n m v_n = nh$$

We know that angular momentum $L = mvr$. Using this and rearranging the equation a bit, we get

$$m v_n r_n = L_n = n\hbar$$

We can use this expression to get rid on v_n in eq5.2

$$\begin{aligned} v_n &= \frac{n\hbar}{m r_n} \\ {v_n}^2 &= \frac{n^2\hbar^2}{m^2 r_n^2} \end{aligned}$$

Substituting in 5.2,

$$\begin{aligned} \frac{m}{r_n}\frac{n^2\hbar^2}{m^2 r_n^2} &= \frac{e^2}{4\pi\epsilon_0 r_n^2} \\ \frac{r_n m}{n^2\hbar^2} &= \frac{4\pi\epsilon_0}{e^2} \\ r_n &= \frac{4\pi\epsilon_0\hbar^2}{me^2} n^2 \end{aligned}$$

All of the terms in fraction are constants and can be combined as a_0, the value of which is 0.529Å.[1]

$$r_n = a_0 n^2$$

Using this expression, the uncertainty Δr becomes

$$\begin{aligned} \Delta r &= r_{n+1} - r_n \\ &= a_0[(n+1)^2 - n^2] \\ &= a_0[2n+1] \end{aligned}$$

If n is large as compared to 1, this expression becomes

$$\Delta r \approxeq 2a_0 n$$

Similarly, we can use $\rho_r = \frac{L_n}{r_n} = \frac{n\hbar}{r_n}$ in the expression for ΔE,

$$\Delta E = \frac{\rho_r}{m}\Delta\rho_r = \frac{n\hbar^2}{m r_n \Delta r}$$

---

[1] Å is a Greek symbol called *angstrom*. It is equal to 0.1nm.

Substituting the value for Δr, we get

$$\Delta E = \frac{n\hbar^2}{ma_0 n^2 2a_0 n} = \frac{\hbar^2}{2a_0 m}\frac{1}{n^2}$$

The value of this uncertainty turns out to be around $140eV$, much larger than the energy of an electron in ground state of the hydrogen atom i.e. $13.6eV$. This shows why it is not possible to confine an electron between the orbitals. This energy is so high, that the electron will leave the atom before it attains that much energy. For this very same reason, we do not find electrons in the nucleus. Energy of electrons is of the order of tens of eV but for nucleons, it is in MeV. Radiation emitted from electron transitions give us light, while radiations originated from nucleus contain much higher energy and give x-rays and γ-rays.

3. Zero point energy

Classically, it is possible to reduce the temperature of an object to 0K by reducing it's motion by cooling it to a point when all the molecules stop completely. Temperature is just a manifestation of the total kinetic energy of the molecules $\frac{1}{2}mv^2$. Let's consider Einstein's model of solid in which all the atoms are attached with springs representing interactive forces. Each atom has its kinetic energy and the average energy is given by $\left\langle \frac{1}{2}mv^2 \right\rangle$[1]. This energy can also be written as

$$E = \left\langle \frac{1}{2}mv^2 \right\rangle = 3 \times \frac{1}{2}k_B T$$

Where $\frac{1}{2}k_B T$ represents the energy associated with each degree of freedom, and the multiplicative factor of 3 is there because every atom has three degrees of freedom $x, y,$and$z$. If we cool our solid using liquid nitrogen, we can reduce the temperature to 77K, using liquid Helium can lower it down further to 4.2K, using dilution mechanism we can take it to the mK (milliKelvin) regime, and finally the laser cooling can take the temperature to pK (picoKelvin) regime. Seemingly so, we can take it to even 0K if we try hard enough. However, this is not possible according to Heisenberg's uncertainty principle, and there is a minimum value of energy that can not be receded.

$$\Delta E = \frac{\rho}{m}\Delta\rho$$

Minimum value of momentum is also bounded by the uncertainty $\Delta\rho$, so

$$E_{min} \approx \frac{\Delta\rho^2}{m}$$

---

[1] $\langle\rangle$ is used to represent average.

For simplicity, we consider our solid to be one dimensional and that dimension has a length of a units. This turns the uncertainty in position equal to a. Using Heisenberg's uncertainty principle again

$$\begin{aligned} E_{min} &\approx \left(\frac{\hbar}{2a}\right)^2 \frac{1}{m} \\ &\approx \frac{\hbar^2}{4a^2 m} \end{aligned}$$

No energy can go below this value even if the whole universe was to be frozen still.

4. Optical Imaging

The mechanism of camera is just like a single slit experiment. The light passes through a slit, then some optical components, and falls on the image plane where it is observed. As we already know, the light would not fall on the screen at an exact spot, rather, there will be a spread on the screen because of the uncertainty in momentum. This uncertainty is inversely proportional to the slit size i.e. uncertainty in position.

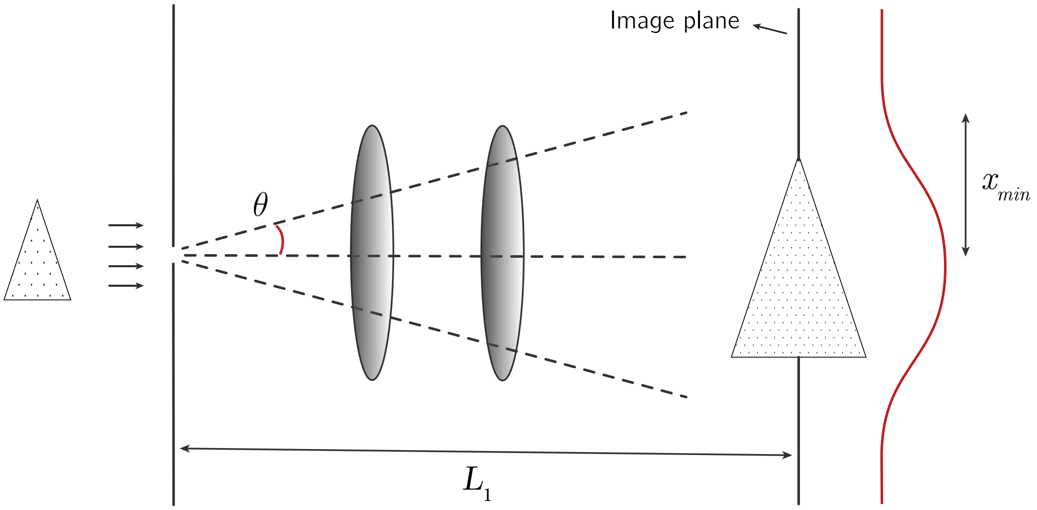


Figure 4.25: Image formation on a screen

For small θ, the value of x_{min} turns out to be as discussed in section 4.3.

$$x_{min} = \frac{\lambda L_1}{a}$$

If we want the image formed to be sharper (have higher resolution), we need to make the spread of light on the screen smaller i.e. making x_{min} smaller.

One way to do so is to increase the slit size a, and the other way is to use waves of smaller wavelength. We could use electrons as an example. Electrons have a much smaller wavelength than light, given by

$$\begin{aligned}\lambda &= \frac{h}{\sqrt{2mK}} \\ \lambda &= \frac{h}{\sqrt{2meV}}\end{aligned}$$

Where K is the kinetic energy, e is the elementary charge, and V is the potential difference. Higher the energy of electrons, shorter will be the wavelength, and sharper will be the image formed. Short wavelength of the electrons is the reason why electron microscopes give much sharper images than optical microscopes.

Resolution of an imaging device is determined using *Rayleigh's Criterion*. It states that two distinct objects will be distinguishable by our imaging device if the peak formed by one object on the screen goes to zero before the peak of the other starts to rise. For this purpose, the separation of the two maxima should be at least $2x_{min}$.

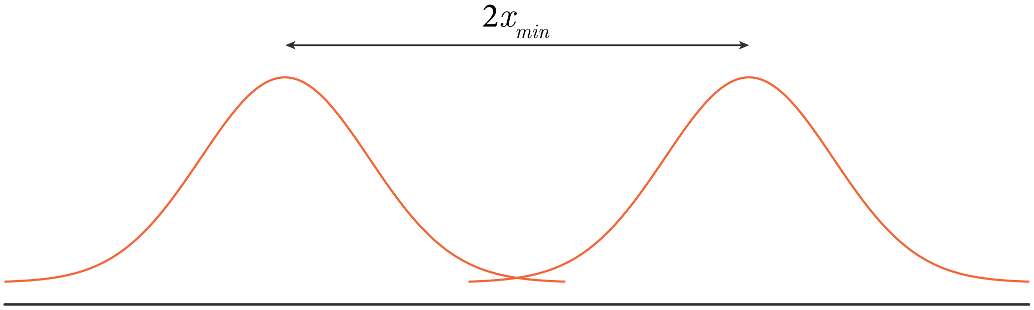


Figure 4.26: Rayleigh's criterion

Let's examine a telescope we are using to observe the surface of the Moon. We are interested in finding out the minimum distance d that should be there between two distinct objects on the surface of the moon for our telescope to be able to resolve them as distinct.

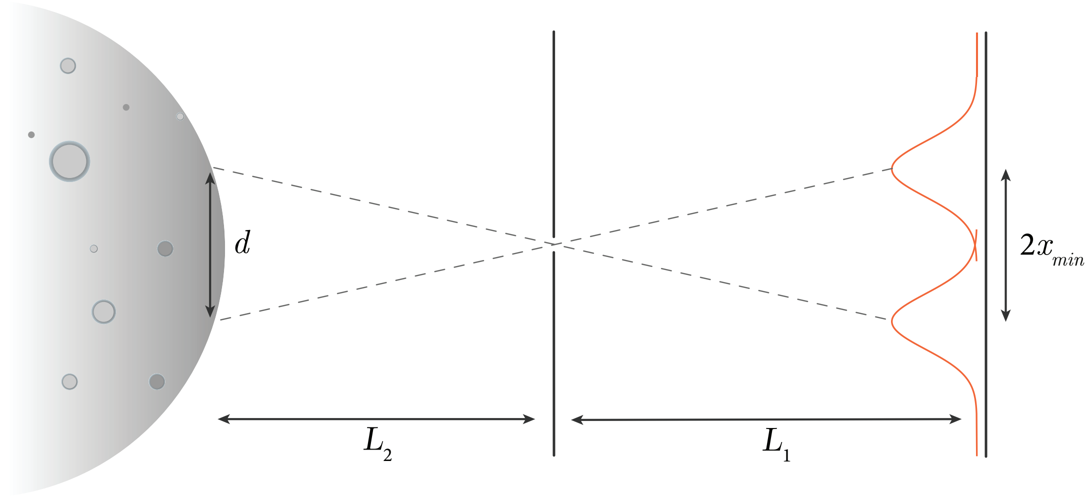


Figure 4.27: "Note: The figure is not drawn to scale and L_2 is much larger than L_1"

Comparing ratios of one right-angel triangle from each side,

$$\frac{x_{min}}{L_1} = \frac{d}{2L_2} \tag{4.3}$$

$$d = \frac{2L_2}{L_1}x_{min} \tag{4.4}$$

Substituting $x_{min} = \frac{\lambda L_1}{a}$,

$$d = \frac{2L_2}{L_1}\frac{\lambda L_1}{a} \tag{4.5}$$

$$= \frac{2L_2\lambda}{a} \tag{4.6}$$

There are two ways to resolve smaller and smaller distances. One is to increase the slit size a i.e. using a larger telescope, and the other is to use shorter wavelengths. So, instead of using visible light, we could do γ-ray or x-ray imaging for better resolutions.

EXAMPLE

If we have a telescope of 1 m opening and we observe 500 nm waves coming from Moon's surface, what should be distance between two objects on Moon for them to be resolved by our telescope?

Data: $\lambda = 500$ nm, a $= 1$ m, $L_2 \approx 10^8$ m

Solution:

$$\begin{aligned} d &= \frac{2L_2\lambda}{a} \\ d &\approx \frac{2 \times 10^8 \times 500 \times 10^{-9}}{1} \\ &\approx 100m \end{aligned}$$

All of this discussion has been purely classic. So far, making better instruments can give us better resolutions, however, there must be some limit where uncertainty principle kicks in. After all, we are dealing with waves and there must be a limit for how small Δx can get.

Chapter 5

Quantum Computing

Detailed study of natural systems reveals that they are inherently quantum-mechanical and, hence, cannot be modeled by classical means. Even a hydrogen atom, one of the simplest quantum-mechanical systems, cannot be modeled in a classical computer without some assumptions. Even large-scale supercomputers have computational limitations. As an example, calculating the prime factors of millionth-order number ($\approx 10^6$) is a computationally intensive task and will take classical computers more time than the age of the universe to enumerate those factors (can't wait for that long, can we?).

That's where the quantum revolution comes into play.

The quantum-mechanical revolution came in the 20th century and changed the very foundations of natural sciences. On the other hand, the revolution in computer sciences came about later in the century, when Claude Shannon found the fundamental limits of signal processing by studying the quantification of information. This was followed by the invention of transistors and the creation of neural networks and the birth of artificial intelligence, machine learning, and deep learning. The merger of these two fields gave rise to a novel field: quantum information processing. In the next few years, quantum computers are likely to become commercial.

By this point in the book, you have all the concepts necessary for understanding quantum computing and quantum information processing. So, let's get to it!

5.1 Quantum Bit

Generally, computers work on Boolean logic, having two distinct states: 0 and 1. These states are classically achievable. For example, if a capacitor is charged, we can consider it state 1, and if the capacitor is uncharged, we can consider it state 0. At a certain time, the capacitor can either be charged (1) or discharged (0) but not both. Such a system with a mutual exclusivity between the states 0 and 1, is known as a bit.

5.1.1 Quantum Interference

Quantum mechanics, however, is probabilistic. We have a pretty good idea of this fact, thanks to the double-slit experiment. As we have seen before, it comprises of two slits followed by a screen. Thus, there are two paths available to the photons entering the apparatus. Electron passing through slit 1 was said to be in state ψ_1, and electron passing through slit 2 was said to be in state ψ_2.

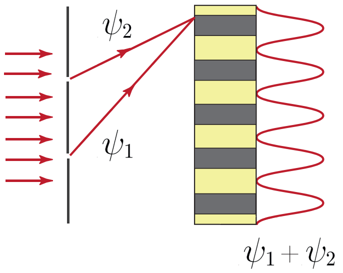


Figure 5.1: Double slit experiment.

If it were to be a classical system, we would have obtained two distinct dots on the screen. But what we observed instead, was an interference pattern on the screen. This means that in between the slits and screen, there exists a superposition state of two wave functions ψ_1 and ψ_2. This brings quantum uncertainty within the quantum channels.

Let's make a system that is quantum in nature and still gives us two distinct values; a quantum bit. For this purpose, we will need a Beam Splitter (BS). A BS is a device through which a photon of light can either get transmitted or reflected. A $50:50 BS$ is the one that transmits the photon with a 50% probability, regardless of what polarization state the incoming photon was in. If the intensity of incoming light is I_o, half of the light ($I_o/2$) will be transmitted and half of it will be reflected, and our detectors D_1 and D_2 for transmitted and reflected beam respectively, will detect an equal amount of light. This experiment can be performed in a laboratory.

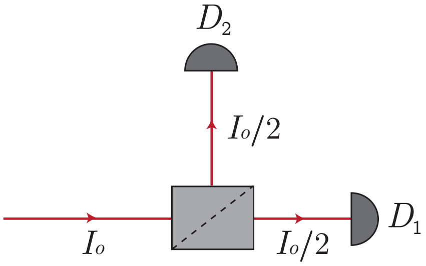


Figure 5.2: Transmission and reflection of light beam by a beam splitter

Let's repeat this experiment with a single photon now. If a single photon enters the slit, only one of the detectors D_1 or D_2 will detect the photon at a particular instant. The path undertaken by the photon defines the state of the system to be either 0 or 1. Consider transmitted photon to be in state 0 and reflected photon to be in state 1. A click on D_1 means that the beam splitter transmitted the photon (turned to state 0 by the BS) while a click on D_2 means that BS reflected the photon (turned the photon to state 1). Note that 0 and 1 are orthogonal states, which means that the state of the system must be either 0 or 1 at a single instant, but not both.

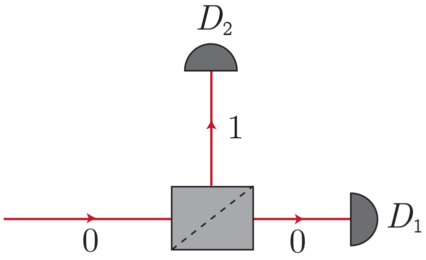


Figure 5.3: Discrete paths for a photon.

The photon also possesses other properties, but we will confine our discussion to its path for now. The path of the photon is a quantum two-level system that will serve as the basic unit of information for our quantum computer. Such a two-level system is called a quantum bit (aptly called qubit or qbit).

In quantum mechanics, we will use a useful notation called 'bracket'. This notation was introduced by Dirac[1]. According to this notation, the state 0 of the system is represented by $|0\rangle$ and the state 1 of the system is represented by $|1\rangle$. The symbol $|\rangle$ is called 'ket' and the symbol $\langle|$ is called 'bra'.

5.1.2 A Simple Quantum Computer

Now that we have a qubit, let's try to build a simple quantum computer. This computer consists of two 50 : 50 beam splitters BS_1 and BS_2. Two mirrors M_1 and M_2 are placed to direct the photons coming from BS_1 to BS_2. The mirrors are perfectly reflecting i.e. any photon that falls on this mirror will reflect with a 100% probability. Two detectors D_1 and D_2 are placed to detect the photons coming from BS_2. The state of transmission is defined as $|0\rangle$ and the state of reflection as $|1\rangle$ like in the previous experiment.

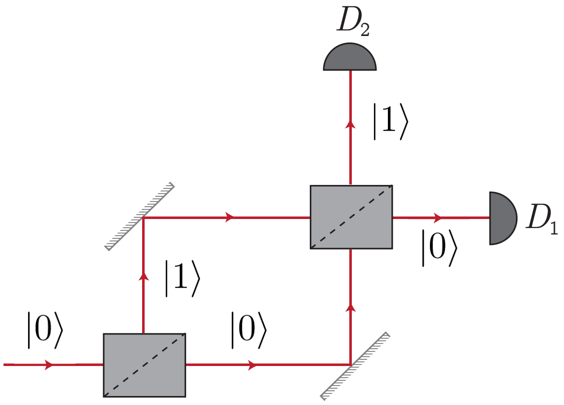


Figure 5.4: Single photon interferometer.

The incoming photon has two options of path to enter BS_1: $|0\rangle$ and $|1\rangle$. If the photon enters with state $|0\rangle$, the beam splitters create a superposition of states $|0\rangle$ and $|1\rangle$, i.e. the coherent sum of $|0\rangle$ and $|1\rangle$.

$$|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

If the incoming photon is in state $|1\rangle$, it creates a superposition state again. However, the relative signs are changed this time.

$$|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

---

[1] Named after the English theoretical physicist Paul Adrien Maurice Dirac

Wondering where this $\sqrt{2}$ factor came from and how the signs changed? Don't worry, we will get back to it later in the chapter.

Notice that the beam splitters are acting like logical gates (AND, OR and NOT etc.) i.e. they are transforming the state of the photon as it passes through them.

The beam splitters take the quantum states and create a superposition. This superposition cannot exist classically. If you have a classical bit, you cannot have a 0 and 1 at the same time. You cannot have a capacitor that charges and discharges simultaneously. But quantum mechanics allows you to have a superposition of orthogonal states. This is what happens in the double-slit interference experiment: both paths interfere with one another to produce a final outcome that is a superposition of the two paths.

Now that we have a superposition of states, let's see what happens.

As shown in the figure, the incoming photon is in state $|0\rangle$. This photon will be either transmitted or reflected by BS_1. After getting reflected by mirror M_1 or M_2, it will enter BS_2 where it will either be reflected or transmitted again. The detector D_1 detects state $|0\rangle$ and the detector D_2 detects the state $|1\rangle$. The complete mathematical description is as follows

$$\begin{aligned} |0\rangle &\xrightarrow{BS_1} \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) \\ &\xrightarrow{BS_2} \frac{1}{\sqrt{2}}\left(\frac{|0\rangle + |1\rangle}{\sqrt{2}} + \frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) = |0\rangle \end{aligned}$$

Our final state is just $|0\rangle$. Hence, D_1 clicks with a 100% probability while D_2 detects nothing.

Quantum Superposition of Qubits

For a system with single beam splitter and two detectors, if the beam splitter BS is a perfect $50:50$ beam splitter, both the detectors click with a 50% probability. Thus, this beam theory is like a random coin toss: you toss the coin and there is 50% probability that a head will turn up and a 50% probability that a tail shows up. It is a completely random device in this sense. However, when we concatenate two of these devices together, instead of a 50% probability of clicking of either of the detectors, we see that only one of these detectors clicks, and this is something counter-intuitive.

This counter-intuitive result can be explained by the fact that inside the apparatus, the quantum state is actually a superposition of two states: $|0\rangle$

and $|1\rangle$ i.e. $|\psi\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$. We have chosen these particular labels to show the correspondence of qubits to the classical bits.

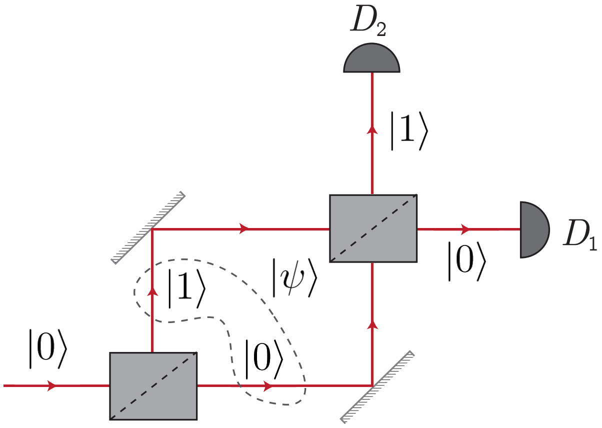


Figure 5.5: Single photon with the super position state in interferometer.

Let a single photon entering the apparatus take its course without you observing it (if you take a peek, you will spoil both the quantum superposition and the surprise!). You might think that it is taking both paths simultaneously, but we cannot make such a claim as we are not observing the path. Now, these two states interfere on the second beam splitter, and an interference pattern emerges such that one of these paths has the two destructively interfering states and the other one has the constructively interfering states. Therefore, only one of the detectors clicks. This is something strange: we have a random device (a single beam splitter) that is giving us a random outcome, but when we concatenate two of these random devices together, our outcome is no longer random; it becomes definite as only one of the detectors clicks.

Measuring the superposition

What if instead of being patient, we take a peek inside? What if we measure that state? The answer is that there will be some probability for $|1\rangle$ occurring and some probability for $|0\rangle$. But how do we calculate the probabilities of these states?

It's quite easy. If we have the state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

The coefficients that appear with $|0\rangle$ and $|1\rangle$ i.e. both $\frac{1}{\sqrt{2}}$ in this case, are called probability amplitudes. And the probability is the modulus square of these amplitudes.

$$\text{Probability that detector } D_1 \text{ clicks} = P(D_1) = \left|\frac{1}{\sqrt{2}}\right|^2 = \frac{1}{2} \tag{5.1}$$

$$\text{Probability that detector } D_2 \text{ clicks} = P(D_2) = \left|\frac{1}{\sqrt{2}}\right|^2 = \frac{1}{2} \tag{5.2}$$

Let's take a special kind of beam splitter (i.e. not the normal 50 : 50 beam splitter), which creates a state

$$|\psi\rangle = \frac{1}{\sqrt{3}}|0\rangle + \sqrt{\frac{2}{3}}|1\rangle \tag{5.3}$$

Now, the probability that D_1 clicks is $\frac{1}{3}$ and the probability that D_2 clicks is $\frac{2}{3}$. The total probability is 1, as it should be.

Suppose we have another beam splitter that gives the state

$$|\psi\rangle = \frac{1}{\sqrt{3}}|0\rangle + \iota\sqrt{\frac{2}{3}}|1\rangle$$

The quantum state is still correct. Detectors D_1 and D_2 will click with the same probability. i.e

$$P(D_1) = \left|\frac{1}{\sqrt{3}}\right|^2 = \frac{1}{3}$$

and

$$P(D_2) = \left|\iota\frac{2}{\sqrt{3}}\right|^2 = \frac{2}{3}$$

Now, suppose we have a qubit which is in a general state,

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \tag{5.4}$$

α and β both are complex numbers in general. There is, a constraint on this state, $|\alpha|^2 + |\beta|^2 = 1$ i.e. sum of probabilities should be equal to 1. This constraint is called the normalization condition. Two vectors $\vec{A}$ and $\vec{B}$ are said to be orthogonal if $\vec{A}.\vec{B} = 0$ i.e. there is no overlap between the two vectors. If we have two states $|0\rangle$ and $|1\rangle$ then $\langle 0||1\rangle = 0$. This is how we show the projection or overlap of the state $|1\rangle$ on state $|0\rangle$. Hence, $|0\rangle$ and $|1\rangle$ are both orthogonal. This product is called the inner product of two states $|0\rangle$ and $|1\rangle$ in Dirac notation. If the two states are same, they will completely overlap and their inner product can be written as $\langle 0||0\rangle = \langle 1||1\rangle = 1$.

Now consider a beam splitter i.e.

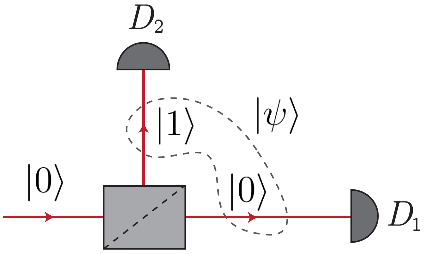


Figure 5.6: Probability of the photon

We would like to find out the probability $P(D_1)$ that the detector D_1 clicks. Note that if we take the inner product of a general state qubit and take its inner product with one of the orthogonal states, we will only be left with the coefficient of that orthogonal state in the general state. Modulus square of that coefficient, as we already know, is the probability of measurement of that state. So, the probability that D_1 clicks, can be written as

$$\begin{aligned} P(D_1) &= |\langle 0|\psi\rangle|^2 \\ &= |\langle 0|(\alpha|0\rangle + \beta|1\rangle)|^2 \\ &= |\alpha\langle 0|0\rangle + \beta\langle 0|1\rangle|^2 \\ &= |\alpha \times 1 + \beta \times 0|^2 \\ &= |\alpha|^2 \end{aligned}$$

Similarly, the probability that D_2 clicks is

$$\begin{aligned} P(D_2) &= \langle 1|\psi\rangle^2 \\ &= |\langle 1|(\alpha|0\rangle + \beta|1\rangle)|^2 \\ &= |\alpha\langle 1|0\rangle + \beta\langle 1|1\rangle|^2 \\ &= |\alpha \times 0 + \beta \times 1|^2 \\ &= |\beta|^2 \end{aligned}$$

The quantum state Eq. (6.4) is a qubit in the superposition state. Note that there is no superposition in the classical bits 0 or 1, i.e. we cannot have superposition in the classical world. The possibility of superposition is something special to the quantum world. And this is what gives an edge to quantum computing over classical computing.

5.2 Quantum Logic Gates

NOT gate is a logic gate that inverts the input. If we have two classical bits 0 and 1, it inverts 0 to 1 and 1 to 0.

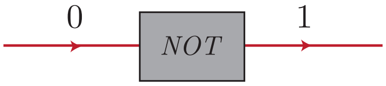


Figure 5.7: NOT gate

Hence, NOT gate performs a transformation of information or, in other words, processes the information.

input	output
0	1
1	0

Table 5.1: Truth table of NOT gate

In a controlled-NOT (CNOT) gate, we have two inputs. It is defined such that *the second bit is inverted if the first bit is* 1.

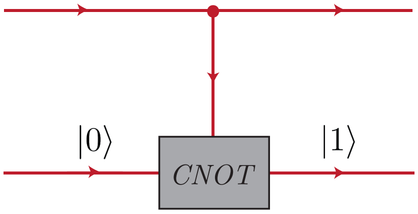


Figure 5.8: CNOT gate

Thus, the action on the second bit is determined by the state of the first bit. The truth table of CNOT gate is

input 1	input 2	output 1	output 2
0	0	0	0
0	1	0	1
1	0	1	1
1	1	1	0

Table 5.2: Truth table of CNOT gate

There are many other gates, such as OR, NOR, AND, and NAND. These logic gates are the building blocks of the classical computer. All these gates do some kind of transformation or processing of information in bits. And if we need to build a quantum computer, we will need quantum analogs of these classical gates for our qubits.

5.2.1 Bloch Sphere

Let's now move to the quantum realm. We know that Eq.(6.4) is the mathematical representation of a qubit. To make our lives easier, I would like to draw a picture of a qubit. It is not an actual qubit but something that you can associate with a qubit for better understanding. This representation of a qubit is called the Bloch sphere.

I define a quantum state on the Bloch sphere as a vector from the center of this sphere to any point on the surface of the sphere such a vector is known as a Bloch vector. The sphere is of unit radius.

If the vector points along the northern hemisphere, this represents the state $|0\rangle$ and if the vector points to the southern hemisphere, this represents the state $|1\rangle$. This is the Bloch sphere representation of two orthogonal qubits.

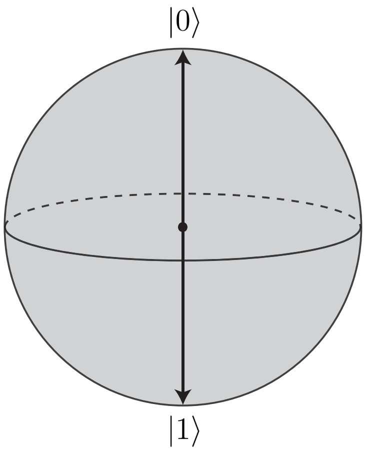


Figure 5.9: $|0\rangle$ and $|1\rangle$ on Bloch Sphere

The superposition state $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ will point to the equatorial plane in $+y$-direction. If we draw its orthogonal state $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, it will be 180° out of phase as compared to the aforementioned state represented by a vector pointing in the $-x$-directon. Similarly, we define the state $\frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)$ as the vector pointing in $+y$ direction and $\frac{1}{2}(|0\rangle - \iota|1\rangle)$ as the vector pointing in $-y$-direction shown in Fig. (6.10).

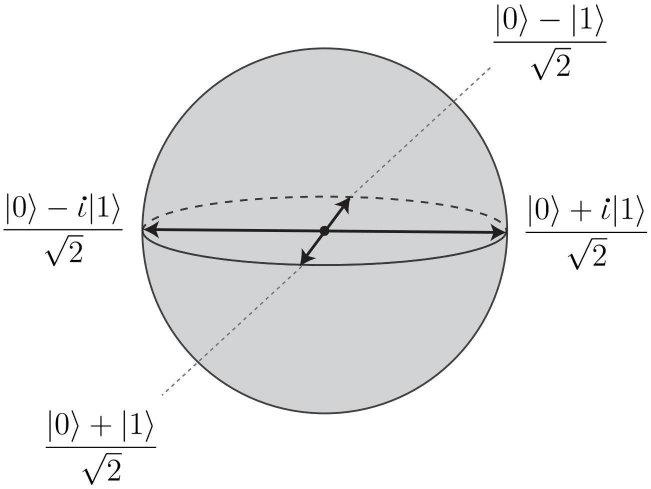


Figure 5.10: Quantum States on the Equatorial Plane

To find the overlap between $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $\frac{1}{\sqrt{2}}(|0\rangle + \iota|1\rangle)$, we can take their inner product.

$$\begin{aligned}
\frac{1}{\sqrt{2}}(\langle 0| + \langle 1|)][\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)] &= \frac{1}{2}(\langle 0|0\rangle + \langle 0|1\rangle - \langle 1|0\rangle - \langle 1|1\rangle) \\
&= \frac{1}{2}(1-1) \\
&= 0
\end{aligned}$$

This shows that the two states don't overlap i.e. they are orthogonal.

But how will we define some arbitrary state from the Bloch sphere, that is not in some specific direction, and lies at a random point on the sphere? Let's draw a Bloch vector for such a state.

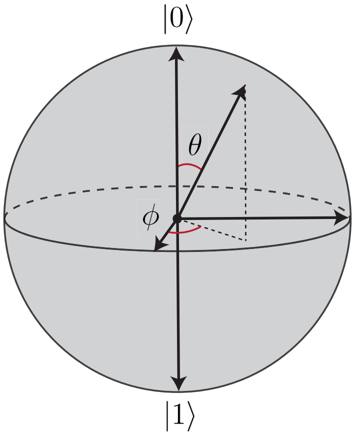


Figure 5.11: Representation of an arbitrary quantum state on the Bloch sphere

This general quantum state and its orientation can be determined from two angles. One of them is the angle θ that it makes with the $z-axis$, this is called the polar angle. It can vary from 0° to 180°. Polar angle is 0° at north pole and 180° at the south pole. The other angle made by the quantum state is called Azimuthal angle ϕ. It is the angle between the $x-axis$ and the projection of the Bloch vector on the $xy-plane$. Mathematically, the universal representation of a quantum state on an arbitrary point on the Bloch sphere is defined as

$$|\psi\rangle = \cos(\frac{\theta}{2})|0\rangle + \sin(\frac{\theta}{2})e^{i\phi}|1\rangle \tag{5.5}$$

We can insert different values of θ and ϕ in this equation and obtain the corresponding quantum states.

If our quantum state is along the $x - axis$ i.e. at the equatorial plane, the polar angle $\theta = 90°$ and Azimuthal angle $\phi = 0°$, and the corresponding quantum state is $|\psi_1\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$

If we have the state vector in the positive y-axis, the corresponding quantum state representation will be $|\psi_2\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}e^{\iota\frac{\pi}{2}}|1\rangle = \frac{1}{\sqrt{2}}|0\rangle + \iota|1\rangle$

5.2.2 Q-NOT Gate

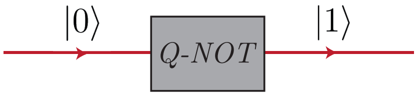


Figure 5.12: Q-NOT gate

Now that we have the physical representation of a qubit in our minds, let's see how can we manipulate our qubit. For that, we will be needing quantum gates. Let us make a quantum NOT gate. Luckily, the Bloch sphere contains all the information to achieve this task. Any single-qubit operation can be described as a transformation of the Bloch vector.

The truth table of Q-NOT gate is

input	output
$\|0\rangle$	$\|1\rangle$
$\|1\rangle$	$\|0\rangle$

Table 5.3: Truth table of Q-NOT gate

This defines the operation or transformation of a Q-NOT gate. If I input the state $\frac{1}{\sqrt{3}}|0\rangle + \frac{2}{\sqrt{3}}|1\rangle$; then my output will be $\frac{1}{\sqrt{3}}|1\rangle + \frac{2}{\sqrt{3}}|0\rangle$ according to the above table.

To put it simply, the logic gates in a quantum computer rotate the quantum states on the Bloch sphere. Q-NOT gate rotates a state by 180° about $x - axis$. The direction of rotation on the Bloch sphere is also important; if my axis of rotation is $x - axis$ then the quantum state must be rotated clockwise.

If we have a quantum state $|\psi_{in}\rangle$ pointing along the $+y$-direction $|\psi_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + \iota|1\rangle)$, and we pass this state through a Q-NOT gate, the output will be $|\psi_{out}\rangle = {}^1/\!\sqrt{2}(\iota|0\rangle + |1\rangle)$. But we know that when we rotate our vector pointing along the $+y$-direction by 180°, we get to the state pointing in $-y$ direction, and the state of this Bloch vector is represented as $|\kappa\rangle = \frac{1}{\sqrt{2}}(|0\rangle - \iota|1\rangle$.

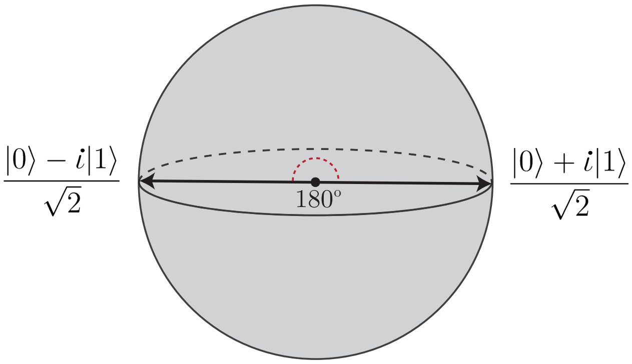


Figure 5.13: NOT gate Implementation on $|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm \iota|1\rangle)$

The output state that we get from the Q-NOT gate is seemingly different from the state represented on the Bloch sphere. But allow me to prove that both these states are, in fact, equal.

$$\begin{aligned} |\psi_{out}\rangle &= \frac{1}{\sqrt{2}}(\iota|0\rangle + |1\rangle) \\ &= \iota\left(\frac{1}{\sqrt{2}}(|0\rangle - \iota|1\rangle)\right) \\ &= \iota|\kappa\rangle \\ &= e^{\frac{\iota\pi}{2}}|\kappa\rangle \end{aligned}$$

Where the factor $e^{\frac{\iota\pi}{2}}$ is the phase factor and $\frac{\pi}{2}$ is called the global phase. If we have two states $|\psi\rangle$ and $e^{i\theta}|\psi\rangle$ then both these states are the same. We cannot differentiate between them. To prove this, let us have a beam splitter and let the input beam in state $|\psi\rangle$ come in. The BS splits the beam to two outputs $|0\rangle$ and $|1\rangle$. The probability that the state $|0\rangle$ is detected by D_1 is

$$\mathrm{P}(\mathrm{D}_1) = |\langle 0|\psi\rangle|^2$$

Similarly, the probability of the state $|1\rangle$ detected by D_2 is

$$P(D_2) = |\langle 1|\psi\rangle|^2$$

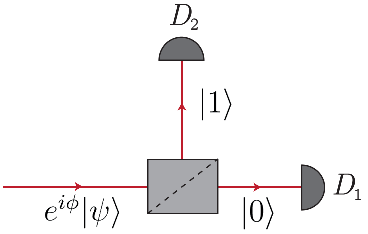


Figure 5.14: Beam splitter

If the incoming beam was in state $e^{i\phi}|\psi\rangle$ instead, the probability of detection by detectors D_1 and D_2 will be. For D_1

$$\begin{aligned} P(D_1) &= |\langle 0|e^{\iota\theta}\psi\rangle|^2 \\ &= |e^{\iota\theta}\langle 0|\psi\rangle|^2 \\ &= e^{\iota\theta}e^{-\iota\theta}|\langle 0|\psi\rangle|^2 \\ &= |\langle 0|\psi\rangle|^2 \end{aligned}$$

The final probability is the same as the above probability. Similarly, for D_2,

$$\begin{aligned} P(D_2) &= |\langle 1|e^{\iota\theta}\psi\rangle|^2 \\ &= |e^{\iota\theta}\langle 1|\psi\rangle|^2 \\ &= e^{\iota\theta}e^{-\iota\theta}|\langle 1|\psi\rangle|^2 \\ &= |\langle 1|\psi\rangle|^2 \end{aligned}$$

These probabilities are exactly the same as the probabilities we got for $|\psi\rangle$. Hence, these states are physically indistinguishable.

5.2.3 Hadamard gate

It's time to define a quantum gate that generates superpositions from orthogonal states. Consider the quantum circuit shown in the figure below. A wire carries a qubit $|0\rangle$ or $|1\rangle$ and we have a device that creates a superposition.

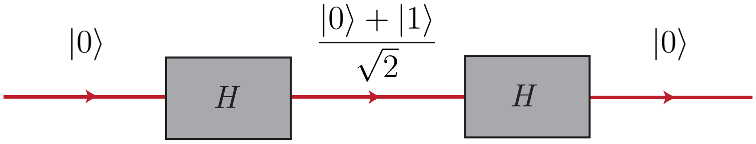


Figure 5.15: Hadamard gate

This device is a quantum gate called the Hadamard gate and the transformation is called the Hadamard transformation. The action of the Hadamard gate is to convert an orthogonal state to a superposition state and vice versa. Truth table for the Hadamard gate is:

input	output
$\|0\rangle$	$\frac{1}{\sqrt{2}}(\|0\rangle + \|1\rangle)$
$\|1\rangle$	$\frac{1}{\sqrt{2}}(\|0\rangle - \|1\rangle)$

Table 5.4: Truth table of Hadamard gate

You might remember that the 50:50 BS created the same superposition states. Yes! Our 50:50 BS was an implementation of the Hadamard gate.

Let's implement this circuit using our 50 : 50 beam splitters.

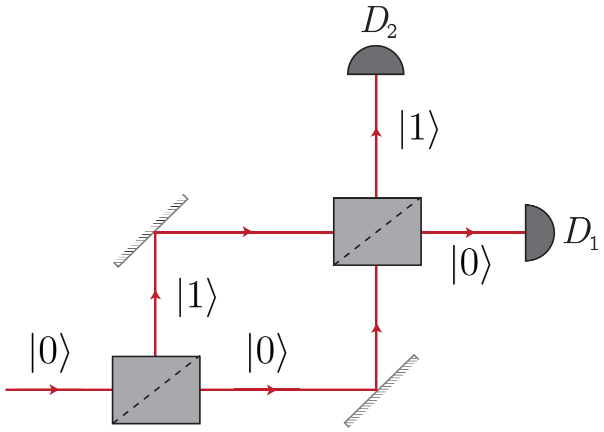


Figure 5.16: Hadamard gates demonstrated using 50:50 Beam Splitters

In this figure, a photon of state $|0\rangle$ comes into the BS_1 and will be either transmitted or reflected with a 50% probability each. And if we decide to not make any measurements after BS_1, the region between the two beam splitters will contain the superposition state $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle$. When this superposition state reaches BS_2, it will act on it as if it was acting on $|0\rangle$ and $|1\rangle$ states independently. But the output will be the state $|0\rangle$ only. Mathematically it can be shown like this:

$$|0\rangle \xrightarrow{H} \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \xrightarrow{H} \frac{1}{\sqrt{2}}(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) + \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)) \tag{5.6}$$

By simplifying the terms, we can easily see that the output state $|\psi_{out}\rangle = |0\rangle$.

5.2.4 Phase Gate

Let's take two Hadamard gates and place another quantum gate in between these gates. Let's call it the Phase Gate (P).

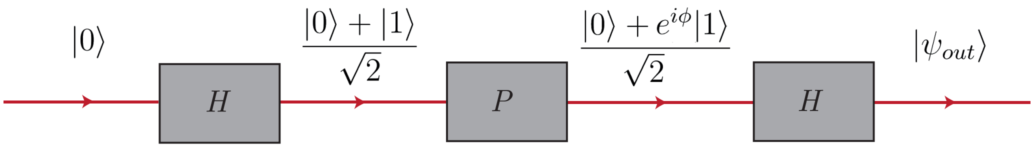


Figure 5.17: Phase gate in between two Hadamard gates

The truth table of the phase gate can be written as follows

input	output
$\|0\rangle$	$\|0\rangle$
$\|1\rangle$	$e^{i\phi}\|1\rangle$

Table 5.5: Truth table of phase gate

What happens if my input state is $|0\rangle$ and it passes through the above circuit?

$$|0\rangle \xrightarrow{H} \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \xrightarrow{P} \frac{1}{\sqrt{2}}(|0\rangle + e^{\iota\phi}|1\rangle) \xrightarrow{H} \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle + \frac{1}{\sqrt{2}}e^{\iota\phi}|0\rangle - \frac{1}{\sqrt{2}}e^{\iota\phi}|1\rangle\right)$$

The output quantum state can be written as

$$|\psi_{out}\rangle = \frac{1}{2}\left((1 + e^{i\phi})|0\rangle + (1 - e^{i\phi})|1\rangle\right)$$

Let's physically implement the above system. To change the phase of the photon we place some medium after $|1\rangle$ output of BS_1.

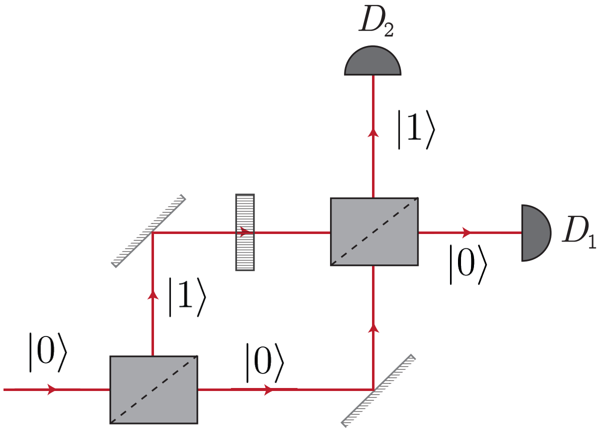


Figure 5.18: Phasor medium in path 1

This medium is known as the phasor medium]gives a phase to the photon due to its certain refractive index. When the photon enters, it slow down and a phase difference is introduced. This phase difference can be implemented by glass or other such material. We can calculate the probabilities that the detectors D_1 and D_2 click.

For detector D_1

$$\begin{aligned}
P(D_1) &= |\langle 0||\psi_{out}\rangle|^2 \\
&= \frac{1}{4}|1+e^{i\phi}|^2\langle 0||1\rangle \\
&= \frac{1}{4}(1+e^{-i\phi})(1+e^{i\phi}) \\
&= \frac{1}{4}(2+e^{i\phi}+e^{-i\phi}) \\
&= \frac{1}{4}(2+2\cos\phi) \\
&= \frac{1}{2}(1+\cos\phi) \\
&= \cos^2\frac{\phi}{2}
\end{aligned}$$

Similarly the probability that detector D_2 clicks is

$$P(D_2) = \sin^2\frac{\phi}{2}$$

Plot for $P(D_1)$ and $P(D_2)$ against ϕ will be

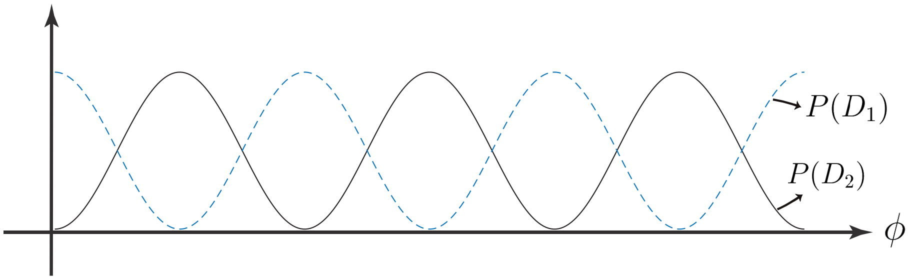


Figure 5.19: Probabilities of detection

If we put the phase element after $|0\rangle$ output of BS_1 as shown in figure below.

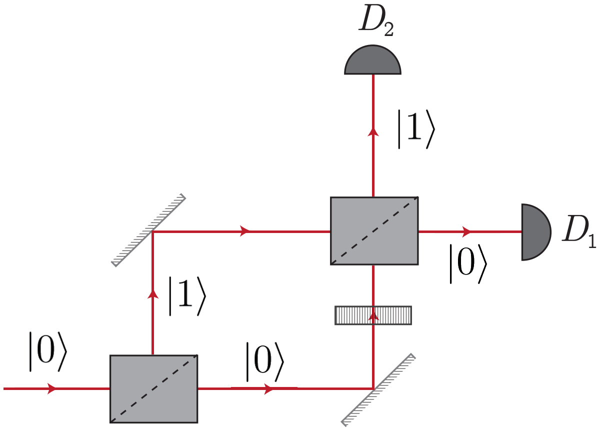


Figure 5.20: Phaser medium in path 2

The truth table changes to

input	output
$\|0\rangle$	$e^{i\phi}\|0\rangle$
$\|1\rangle$	$\|1\rangle$

Table 5.6: Truth table of phase gate

The phase of the state $|0\rangle$ changes with the factor $e^{\iota\phi}$, but our calculations remain the same. It doesn't matter where the phase element is placed.

5.3 Implementations of Qubit

There are many quantum mechanical systems and the path of a photon after passing through a beam splitter is only one of them. This means that there must be more ways to implement a qubit. Let's spend some time exploring some of these systems.

5.3.1 Polarization of light waves

First up in our list is another property of a light wave, its polarization. We know that light waves comprise of a varying electric field. The strength of the electric field changes constantly in an oscillatory fashion. This electric field lies in a plane as shown in the diagram.

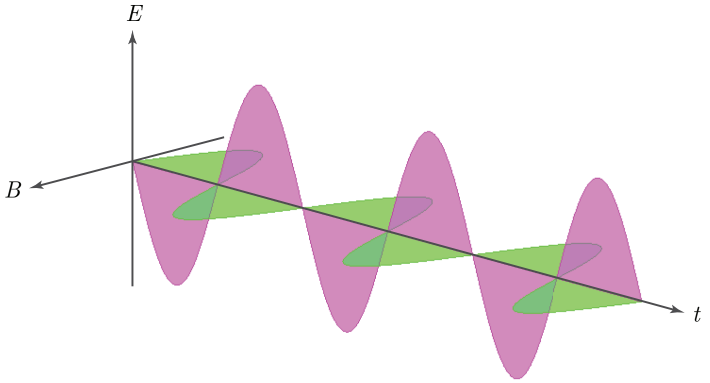


Figure 5.21: Normal wave inside a medium

The plane in which the electric field is polarized is called the plane of polarization. If the direction of the electric field is vertical, the electric field is vertically polarized. Similarly, it is horizontally polarized if the direction of the electric field is horizontal. We define the vertical polarization of photons as state $|0\rangle$ and the horizontal polarization of photons as state $|1\rangle$. This way, we have used the polarization of the photon to physically implement the qubit. The same Bloch sphere representation is used, i.e $|0\rangle$ on the north pole and $|1\rangle$ on the south pole.

We also need a measuring instrument to detect this qubit. Consider a normal wave inside a medium and it passes through an object with a vertical slit in it. If we oscillate the rope vertically, it will pass through the slit, but the slit is horizontal, it will be blocked. This is how the detector for polarization works. Such a detector is called a polarizer. If the axis of the polarizer is vertical, the output will be vertically polarized light and if the polarizer axis is horizontal, the output will be horizontally polarized light.

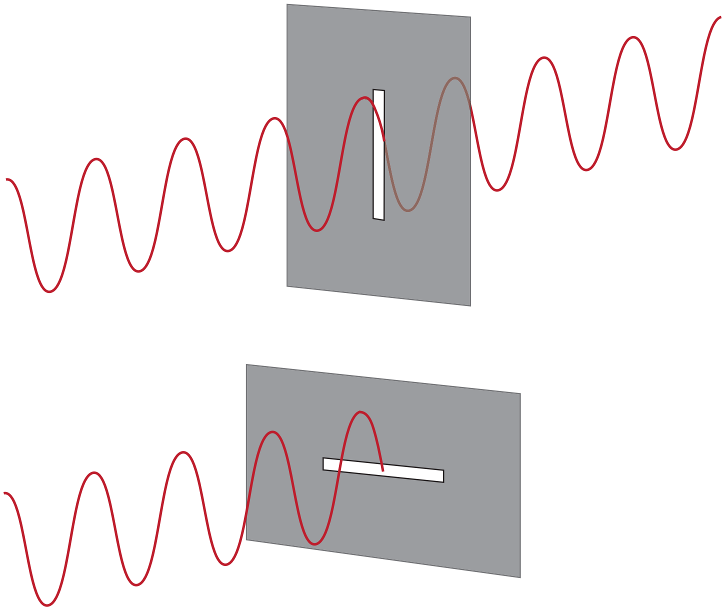


Figure 5.22: Polarization through vertical and horizontal slits

Consider a single photon coming into a vertical polarizer P and reaching the detector. The detector is very sensitive and is able to detect a single photon (hence the name *single photon detector*).

If my input state is $|0\rangle$, which corresponds to a vertically polarized photon, the probability that the detector clicks is $|\langle 0||0\rangle|^2 = 1$. Similarly, if my input state is $|1\rangle$, which corresponds to the horizontally polarized photon, the probability that the detector clicks is $|\langle 1||0\rangle|^2 = 0$. Now, if my input state is $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ then the probability that the detector clicks is $|\frac{1}{\sqrt{2}}\langle 0|(|0\rangle + |1\rangle)|^2 = \frac{1}{2}$ Similarly if my input state is $\frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)$ then the probability is $|\frac{1}{\sqrt{2}}\langle 0|(|0\rangle + i|1\rangle)|^2 = \frac{1}{2}$

I would now like to introduce to you, the polarizing beam splitter PBS. These beam splitters are also called horizontal-vertical beam splitters.

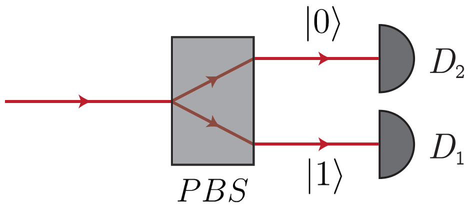


Figure 5.23: Polarizing Beam Splitter

This beam splitter has two output channels. One of these channels is $|0\rangle$ and the other one is $|1\rangle$. When a photon comes in, based upon the polarization of this photon, one of the two channels will be populated. The channels are terminated by detectors D_1 and D_2. If my input state is $|0\rangle$, detector D_1 can detect it with a 100% probability, and if my input state is $|1\rangle$, detector D_2 will detect it with 100% probability

Let's perform an experiment a bit and block the $|0\rangle$ output channel. A photon of polarization state $|0\rangle$ comes into a polarizing beam splitter PBS_1 and I don't observe the photon. Instead, I place another beam splitter PBS_2 in place of the detector D_1. I have two detectors D_1 and D_2 after PBS_2 which will detect the output photons of polarization state $|0\rangle$ and $|1\rangle$ respectively.

What is the probability that the photon is detected by D_1 and D_2? The answer is 0 because the output from PBS_1 has been wasted. Similarly, if the input photon has the state $|1\rangle$ then the probability of its detection by D_1 is 0 and by D_2 is 1.
Now suppose my input state is $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. In this case, the probability of the detection of the photon by D_1 and D_2 is $P(D_1) = 0$ and $P(D_2) = 1$

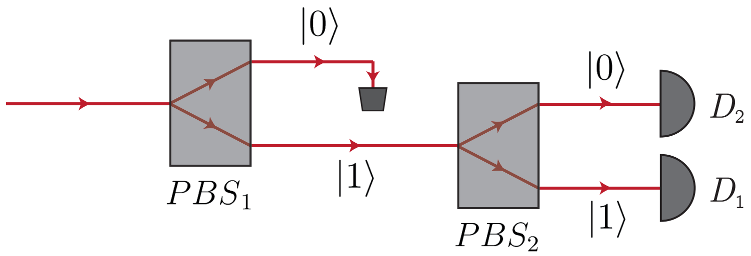


Figure 5.24: One PBS followed by another.

Let's modify this experiment a bit. Let's have the same polarizing beam splitter 1 and 2, having output channels $|0\rangle$ and $|1\rangle$. In between these, we place another kind of beam splitter that has two output channels $\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)$ and $\frac{1}{\sqrt{2}}(|0\rangle-|1\rangle)$.

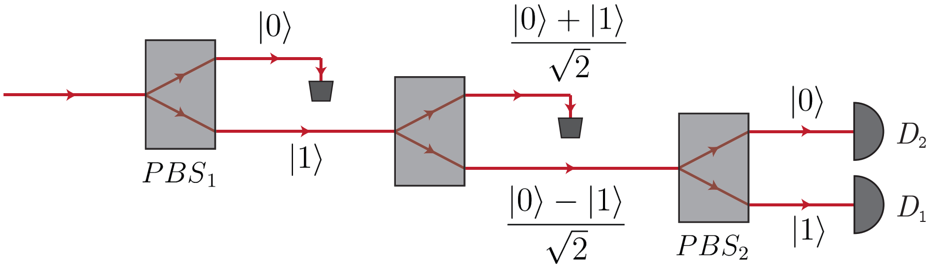


Figure 5.25: Superposition making BS between two PBS

Channel 1 allows +45° photon and channel 2 allows −45° photon through it. Let us pass a photon $|1\rangle$ through these beam splitters. It will pass through channel 2 of PBS_1 because my state is a horizontally polarized state. Now the photon will enter the second beam splitter. Since we can write the state $|1\rangle$ as

$$|1\rangle=\frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)+\frac{1}{\sqrt{2}}(|0\rangle-|1\rangle)\right)$$

the photon will pass through both of the channels with a 50% probability. I dump the state $\frac{|0\rangle+|1\rangle}{2}$, and pass the other channel's output through PBS_2. PBS_2 will now receive the state $\frac{1}{\sqrt{2}}(|0\rangle-|1\rangle)$ and split it into two parts again; half $|1\rangle$ and half $|0\rangle$. Finally, the probability of detection of the photon by the detectors D_1 and D_2 is $P(D_1)=\frac{1}{4}$ and $P(D_2)=\frac{1}{4}$

Previously, if a photon in state $|1\rangle$ was supplied to the system, there was 0 probability of detection on D_1. Now, although the new polarizer wasted 50% of photons, it created a new state which is the superposition of horizontally and vertically polarized states. As a result, detector D_1 clicks with a non-zero probability.

5.3.2 Spin of an Electron

An electron has certain properties, like mass and charge. One such property is of great interest to us as we can use it to make qubits: The spin of electrons. We all know that spin is the rotation or revolution of a body about some axis is called spin, like the spin of earth about its axis. In quantum mechanics, however, spin is not the rotation of electrons.

Before we can further look into the spin of an electron, I shall like to review some electromagnetism. Suppose I have an object of charge q which moves around in a circle with speed v. Due to the motion of the charge, a current will flow in the circle.

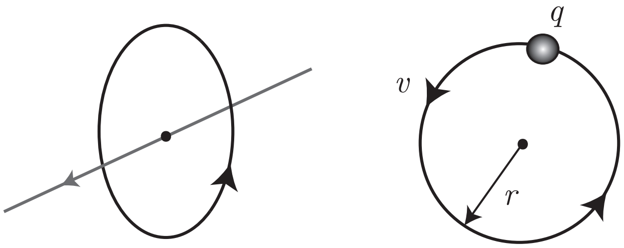


Figure 5.26: Charged particle moving in a circle.

If the radius of the circle is r, the current can be calculated as

$$I = \frac{dq}{dt} = \frac{vq}{2\pi r} \tag{5.7}$$

where the time period of the charge for one complete revolution is $\frac{2\pi r}{v}$.

A moving charge produces a magnetic field B in the direction that can be found from the right-hand rule. In our case, the moving charge produces a magnetic field that points out of the page. Our system starts acting like a magnetic dipole, and attains a dipole moment.

Dipole moment is represented by μ and has the units of ampere square-meter (A.m^2).

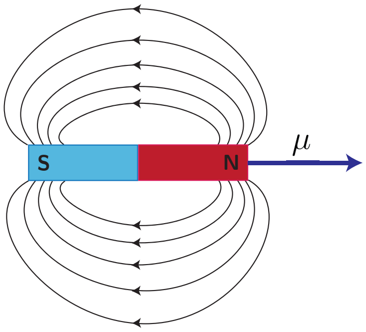


Figure 5.27: Magnetic Dipole

μ depends on the current I and the area A of the circle in which charge is revolving.

$$\vec{\mu} = I\pi r^2 \tag{5.8}$$

By putting the value of the I in the equation above, we have

$$\begin{aligned}
\vec{\mu} &= I\pi r^2 \\
&= \frac{qv\pi r^2}{2\pi r} \\
&= \frac{qvr}{2} \\
&= \frac{q}{2m}(mvr) \\
\vec{\mu} &= \frac{q}{2m}\vec{L} \qquad (5.9)
\end{aligned}$$

Bohr's Atomic model and Spin of an Electron

According to Bohr's atomic model, the electrons are revolving around the nucleus with an angular momentum called orbital angular momentum $\vec{L}$. As $\vec{L}$ is a mechanical quantity and $\vec{\mu}$ is a magnetic quantity, Eq. (6.9) shows the relation between a mechanical quantity and a magnetic quantity of the charged particle. $\vec{L}$ depends on the orbital in which the electron is present. Orbitals with higher energy have electrons with higher orbital angular momentum. Since we only have four orbitals, we only have four quantized values of orbital angular angular momentum. Angular momentum is given by the formula

$$\vec{L} = \sqrt{l(l+1)}\hbar \tag{5.10}$$

The factor $\frac{q}{2m}$ is constant because the charge on an electron (1.6×10^{-19}) and mass of electron (9.11×10^{-31}) are constant quantities. Magnetic moment $\vec{\mu}$ is just the product of this constant and the quantized quantity vecL, hence $\vec{\mu}$ is also quantized. All these values are summarized in the table below.

Orbital	s	p	d	f
l	0	1	2	3
$L = \sqrt{l(l+1)}\hbar$	0	$\sqrt{2}\hbar$	$\sqrt{6}\hbar$	$2\sqrt{3}\hbar$
$\mu = \frac{q}{2m}L$	0	$\frac{q}{\sqrt{2}m}\hbar$	$\frac{q}{m}\sqrt{\frac{3}{2}}\hbar$	$\frac{q}{m}\sqrt{3}\hbar$

Table 5.7: Orbital angular momentum and magnetic moments of different orbitals

Stern-Gerlach Experiment

If we have a stream of particles coming, how we can distinguish the magnetic moment of the different particles? Suppose that we have a magnetic field $\vec{B}$ pointing upwards and two magnetic moments $\vec{\mu_1}$ and $\vec{\mu_2}$ which are parallel and anti-parallel to the field $\vec{B}$.

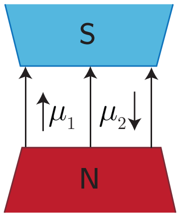


Figure 5.28: Uniform magnetic field

If we put a dipole inside a magnetic field, the magnetic field tends to align the dipole inside the magnetic field. The energy of the dipole parallel to the magnetic field is lower, while the energy of the dipole anti-parallel to the magnetic field is higher. Inside the magnetic field, it is

$$E = -\vec{\mu}.\vec{B} \tag{5.11}$$

The force acting on the dipole inside the magnetic field is $\vec{F} = -\vec{\nabla}E$ Where $\vec{\nabla}E$ is called the gradient of the energy and the component of the force are

$$F_x = -\frac{d}{dx}(\vec{\mu}.\vec{B})$$
$$F_y = -\frac{d}{dy}(\vec{\mu}.\vec{B})$$
$$F_z = -\frac{d}{dz}(\vec{\mu}.\vec{B})$$

It means that the variation in the magnetic field gives the force acting on the magnetic dipole.

Let's now perform an experiment. Consider the north pole and south pole of two magnets facing each other. The magnets are of non-uniform shape as shown in figure below. The magnetic field lines start from the north pole and bunch together in the south pole which produces a non-uniform magnetic field. There is a field gradient from north to south poles. The magnetic field can be written as

$$\vec{B} = B_x\hat{\mathrm{i}} + B_y\hat{\mathrm{j}} + B_z\hat{\mathrm{k}} \tag{5.12}$$

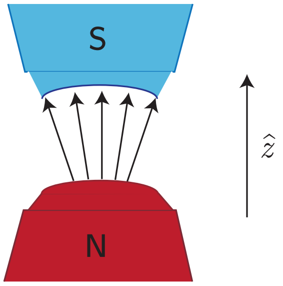


Figure 5.29: Non-uniform magnetic field

Let's suppose that the magnetic field in this experiment is oriented along the z-axis. Thus, the field is stronger in the bottom and weaker on top i.e. the field is changing along the z-axis. We must, however, remember that this assumption is purely theoretical and is not physically achievable. In this case, the magnetic field can be written as $\vec{B} = B_z \hat{\mathrm{k}}$ and force on the dipole as $F_z = -\mu_z \frac{dB_z}{dz}$

This experiment was originally performed by two scientists, Stern and Gerlach in 1932. For this experiment they used Silver atoms in which the electrons exist in the s-orbital. They produced a non-uniform field like the one shown above, baked the Silver in an oven at high temperature (T). Raising the temperature increased the energy of the Silver atoms and they started popping out of the oven in the form of an atomic beam. The beam passed through the non-uniform magnetic field.

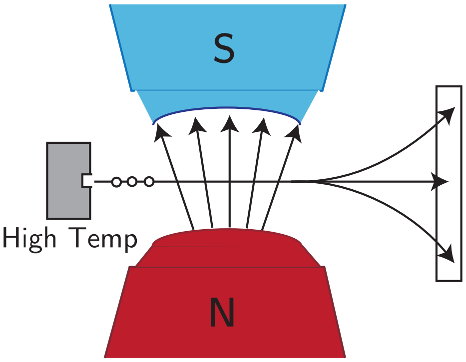


Figure 5.30: Stern-Gerlach experiment.

At the right end, a screen was placed to observe the orientation of the beam that passed through the non-uniform magnetic field. Silver atom did not have any magnetic moment because $L = 0$ for s orbital, and the atoms should not interact with magnetic field. However, Stern and Gerlach observed that the beam transmitted in two distinct paths and created distinct spots on the screen. This implied that the magnetic moment μ_z can take up two values, a positive one and a negative one $\mu_z = \pm\mu$.

This accidental theory led to the discovery of another property of the particle: the spin. Spin was theoretically predicted by Dirac and the experiment by Stern and Gerlach proved its existence. This experiment proved that apart from orbital angular momentum, there is another moment associated with the particle, called spin angular momentum of the particle. It is the property of a particle and gives the same magnetic moment with an additional factor g i.e

$\vec{\mu} = g\frac{q}{2m}\vec{S}$

Where g is called Lande G-factor and $\frac{q}{2m}$ is called gyro-magnetic ratio. As particles have both orbital angular momentum $\vec{L}$ and spin angular momentum $\vec{S}$, the total dipole moment $\vec{\mu}$ can be written as

$$\vec{\mu} = \frac{q}{2m}(\vec{L} + g\vec{S}) \tag{5.13}$$

In our case, the orbital angular momentum of the silver atom is zero $\vec{L} = 0$ and it only possesses spin angular momentum. Hence, the dipole moment is purely spin dipole moment.

$\vec{\mu} = g\frac{q}{2m}\vec{S}$

But we are only concerned with the z-component of the magnetic moment $\mu_z = g\frac{q}{2m}S_z$.

Since μ_z is quantized, the spin is quantized and is a pure quantum mechanical property of a particle. From the experiment, it was found that μ_z has two values. This implies that S_z has two values. Hence, this system is a two-level system and the spin magnetic moment has two values that make a qubit. Therefore, the spin on an electron is another possible materialization of a qubit.

S_z takes two possible values for an electron $S_z = \pm\frac{\hbar}{2}$. Factor $\frac{1}{2}$ is the spin of the electron. Hence, the electron is a spin-half particle with $g \approx 2$. The magnetic moment of the electron is therefore

$\mu_z = \pm\frac{q}{2m}\hbar$

μ_z is quantized and there are only two possibilities of μ_z. These two possibilities act like $|0\rangle$ and $|1\rangle$. If the direction of the magnetic moment is along the magnetic field, it is represented by $|0\rangle$ if direction of magnetic moment is against the magnetic field, it is represented by $|1\rangle$. In this case, $|0\rangle$ represents the spin-up ($|\uparrow\rangle$) and $|1\rangle$ represents the spin-down ($|\downarrow\rangle$). Both states are orthogonal and represented by the Bloch sphere at the north pole and south pole.

Originally these spins are randomly aligned. If we do not have a magnetic field, the magnetic moment is aligned randomly in all possible directions. The magnetic field in the Stern-Gerlach experiment orients the moments along the $\pm z - axis$ and creates two paths. One of these paths is $|0\rangle = |\uparrow\rangle$ and the other is $|1\rangle = |\downarrow\rangle$. In this way, we have encoded the quantum information within the quantum state.

It is also possible to create a superposition if these states. As an example, let's create a state which lies around the x-axis on the Bloch sphere. To measure the spin of electrons, we use a special detector called the Mott detector.

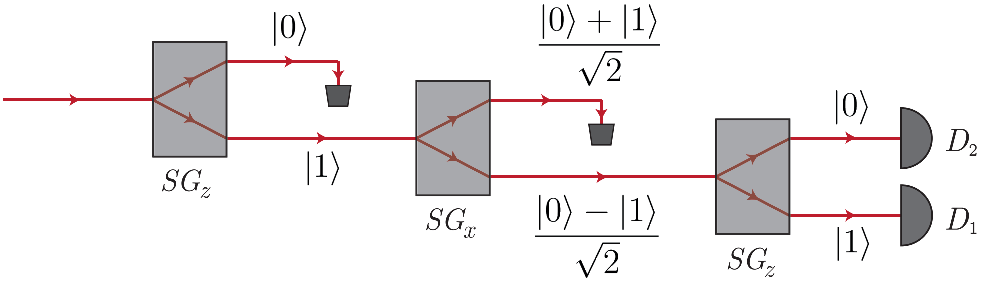


Figure 5.31: State around $x - axis$.

First, we take the above-shown apparatus and block the output channel of $|0\rangle$ state. Then we pass the output state of the $|1\rangle$ channel through another Stern-Gerlach apparatus which is now oriented along the x-axis. The direction of the magnetic field in this apparatus is thus rotated by 90°. To detect the spin states, we will use another Stern-Gerlach apparatus oriented along the z-axis.

Now that our apparatus is all set, let's pass some particles with randomly orientated spins through it. The first Stern-Gerlach apparatus will split these particles into two paths. $|0\rangle$ is blocked and $|1\rangle$ passes through the second Stern-Gerlach apparatus. One of the outputs of this second SG-apparatus is $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and the other is $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$.

Hence, both channels are populated. If I block $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, the probability of electron in the blocked state is $\left|\langle 1| \left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\right)\right|^2 = \frac{1}{2}|\langle 1||1\rangle|^2 = \frac{1}{2}$

And the probability in the transmitted state is $\left|\langle 1|\left(\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\right)\right|^2 = \frac{1}{2}|\langle 1||1\rangle|^2 = \frac{1}{2}$

The SG-apparatus creates a new state which is a superposition of two spins.

As the last SG-apparatus is oriented along the $z - axis$, both the output states should be populated with $\frac{1}{2}$ probability. Note that the $\frac{1}{2}$ probability indicates that there is some uncertainty about the projection along $z - axis$. If the last SG-apparatus was along the $x - axis$, there would have been no uncertainty about the final state of the electron and only one of the channels would have been populated. In principle, we can only measure one of the two orientations at a time. If we measure the x-state, we will be uncertain about z-projection. Similarly, if we measure z-state, we will not be certain about x-projection anymore. Hence, we cannot be sure about x-orientation and z-orientation at the same time.

5.3.3 Other Implementations

We can have many more implementations of a qubit. Some of these are briefly discussed here.

Spin of a Proton

The proton also has a spin associated with it and it also acts as a tiny, magnetic dipole. If we pass this dipole through a magnetic field $\vec{B}$, and it aligns parallel the field, we say that it is in state $|0\rangle$, and if it aligns anti-parallel to the field we say that it is in state $|1\rangle$. Between these two orthogonal states, there will be infinitely many superposition states.

Energy Levels of Electron Inside an Atom

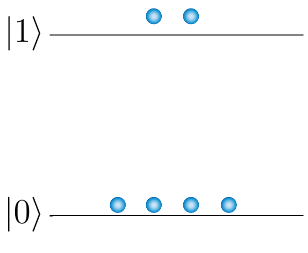


Figure 5.32: Direction of magnetic dipoles in magnetic field.

A qubit can also be implemented using quantized energy levels of electrons inside an atom. We can refer to the ground state as $|0\rangle$ and the excited state as $|1\rangle$. We can create a superposition of these states too. If we put energy into this system and this energy exactly matches the energy gap between two energy levels i.e. $\Delta E = \hbar f$, then this energy can move the electron from the ground state to the excited state. In this way, we can implement different operations on qubits and construct gates for this qubit system.

Superconductor Qubit

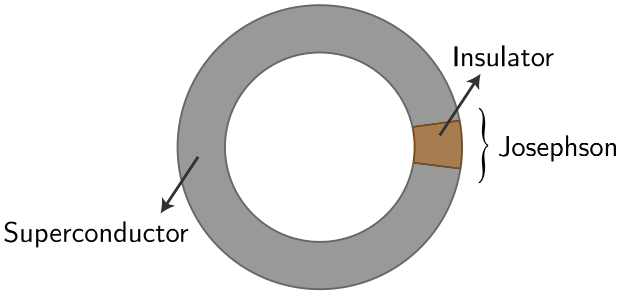


Figure 5.33: Direction of magnetic dipoles in magnetic field.

This device shown in the diagram above is known as a SQUID (Superconducting Quantum Interference Device). A SQUID is a ring of superconductor through which a portion is cut out and filled with an insulator. In the region with the insulator, we have got a superconductor-insulator-superconductor junction. This kind of junction is called a Josephson[1]

There are two ways of implementing qubits using SQUIDs. One way is to consider the flux passing through the SQUID. If a current passes through the superconductor, a magnetic field $\vec{B}$ is produced. The product of this magnetic field $\vec{B}$ with the area A of the SQUID gives us the magnetic flux ϕ, which will be a quantized value. Presence of flux will be denoted by $|0\rangle$ and absence of flux will be denoted by $|1\rangle$

The other way is to consider the induced current passing through the SQUID in presence of a changing magnetic field. If the current is in anticlockwise direction, we say that the state is $|0\rangle$, and $|1\rangle$ if the direction of induced current is clockwise.

---

[1] Josephson won Nobel prize in 1973 for his work on SQUIDs.

5.4 Quantum Computer

Let's now make a practically useful quantum computer. The key point is that there are many physical implementations of the qubit. To build a quantum computer, we need to choose one particular technology first. We need to identify the quantum states of the qubit used in that technology. If we have to encode information, we will need to make quantum gates because we have to manipulate the states somehow. With all of these recipes in place, we can now build a quantum computer.

We need to have at least two qubits and a correlation between them. We want them to interact with one another so that they can do something useful. We need an ascendant machine that is beyond the classical computers.

5.4.1 Quantum Entanglement

So far, we have seen quantum circuits containing a single qubit and they were independent of what other qubits do or what their state is. Quantum entanglement is the phenomenon in which the state of one qubit depends on the state of some other qubit. Let's start by considering two qubits. One of them goes through the first channel while the other one goes through the second channel. Suppose both of them are in state $|0\rangle$. The quantum state in the shaded region will be written as $|0\rangle \otimes |0\rangle$, or in short $|0\rangle \otimes |0\rangle$.

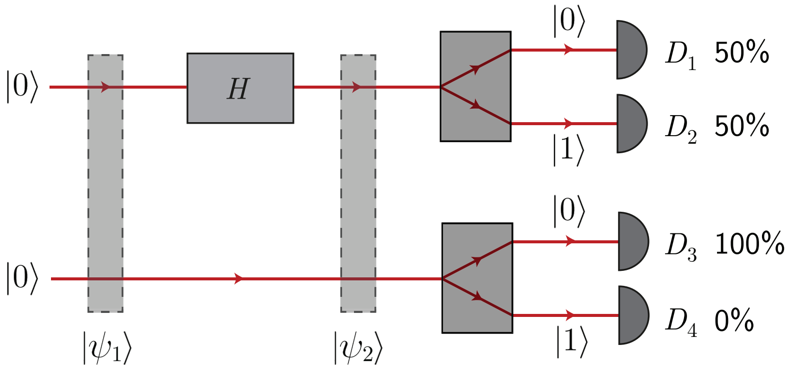


Figure 5.34: Qubits to study quantum entanglement.

$|\psi_1\rangle = |0\rangle|0\rangle$ represents the initial state of the two-qubit system.[1] Let's now put a Hadamard gate at the first channel and leave the second state

---

[1] Note that this is not the multiplication of two states.

unchanged. The Hadamard gate creates a superposition of the qubit, and the state $|0\rangle$ is converted to state $\frac{|0\rangle+|1\rangle}{\sqrt{2}}$. Now the quantum state in the second shaded region is

$$\begin{aligned} |\psi_2\rangle &= \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right)|0\rangle \\ &= \frac{1}{\sqrt{2}}(|0\rangle|0\rangle + |1\rangle|0\rangle) \end{aligned}$$

Let's measure the outputs of both channels. For this, we'll connect each channel with a measuring device. This device has two output channels $|0\rangle$ and $|1\rangle$. Thus, if we place detectors at the outputs of this measuring device, we can measure the probabilities of states $|0\rangle$ and $|1\rangle$.

If the first qubit passes through the device, both the detectors will click with a 50% probability. In the second device, only the detector D_3, which detects $|0\rangle$, will click with a 100% probability. Detectors D_1 and D_3 detect the same state with different probability, and there is no correlation between them. Such a qubit can not be used to build our quantum computer.

Let's now replace this experiment with another one, in which we have a Hadamard gate in the path of both the qubits. The rest of the experiment remains unchanged. This time, all the detectors will click with the same probability.

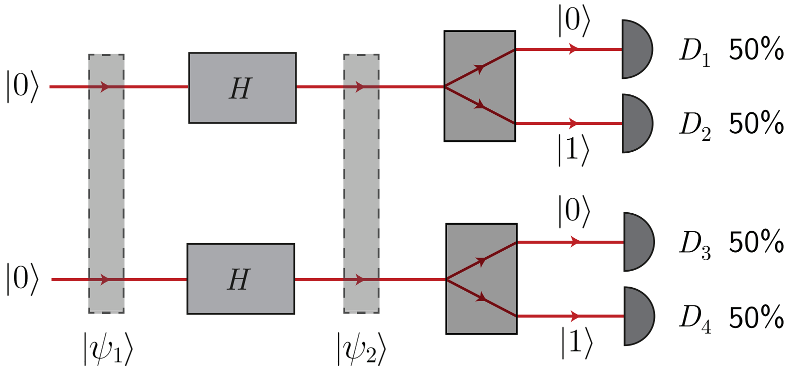


Figure 5.35: Hadamard gate in the path of both the qubits

But there is still no interaction between the qubits. So, we will have to come up with some other design.

Let's modify our first experiment again. This time, we will place a CNOT gate in the path of the second qubit. This CNOT gate will be controlled by the output of the Hadamard gate of the first qubit. Let's see what happens now.

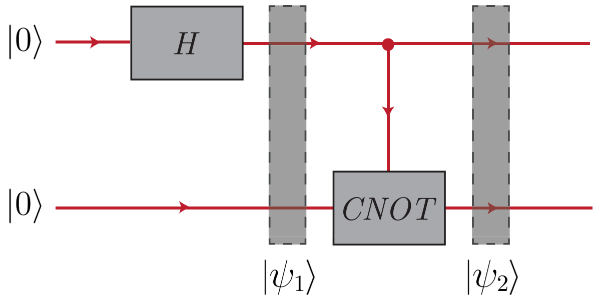


Figure 5.36: Quantum entanglement

Consider two qubits to be in state $|0\rangle$ initially. The Hadamard gate changes the state of the first qubit to $\frac{|0\rangle+|1\rangle}{\sqrt{2}}$. This controls the NOT gate on the second qubit. When the state of the first qubit is $|1\rangle$, the state of the second qubit will be inverted; the state remains unchanged otherwise.

The quantum state in the first shaded region is

$$\begin{aligned} |\psi_1\rangle &= \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right)|0\rangle \\ &= \frac{1}{\sqrt{2}}(|0\rangle|0\rangle + |1\rangle|0\rangle \end{aligned}$$

The state $|1\rangle$ of the Hadamard gate inverts the state of the second qubit, and the state of the second shaded region becomes

$$|\psi_2\rangle\frac{1}{\sqrt{2}}(|0\rangle|0\rangle + |1\rangle|1\rangle)$$

In this experiment, the state of the second qubit depends on the first qubit. Hence, we have finally achieved quantum entanglement.

Since you have accompanied me in obtaining quantum entanglement, I must now explain how we have just achieved is one of the weirdest phenomena. Before I continue, let me introduce you to Alice and Bob. They want to verify quantum entanglement with us, and we could always use some help. Alice decides to observe the first qubit and Bob decides to observe the second one. If we place detectors at the output channels of both the qubits, both

D_1 and D_2 will click with a 50% probability. Similarly, D_3 and D_4 will also click with a 50% probability. This means that Alice and Bob are observing completely random events.

Alice decides to take her qubit and go to Pluto, far away from Bob. They both measure their qubits and note the results. As one measurement is not enough to verify the results, they take numerous measurements. Alice then decides to come back, and they compare the results. Alice's string of the random measurements is 001011000110101.... When Bob shows his string to Alice and they compare their results, they find that Bob's string exactly matches with Alice's string i.e 001011000110101.... This correlation does not make much sense, as both of them measured random events. Actually, once Alice and Bob made measurements, the states collapsed to $|0\rangle|0\rangle$ with a probability $\frac{1}{2}$ and into state $|1\rangle|1\rangle$ with the same probability. If Alice measures $|0\rangle$, Bob also measures $|0\rangle$. So, their final result is perfectly correlated, which verifies quantum entanglement. The two qubits are exact copies of each other.

What if our initial state of the two-qubit system was $|1\rangle|0\rangle$ instead of $|0\rangle|0\rangle$? Let's check. To prepare this state, we start off with qubits in state $|0\rangle|0\rangle$ and pass the first qubit through a quantum NOT gate. Our state is now $|1\rangle|0\rangle$ which is separable and the qubits are not co-related.

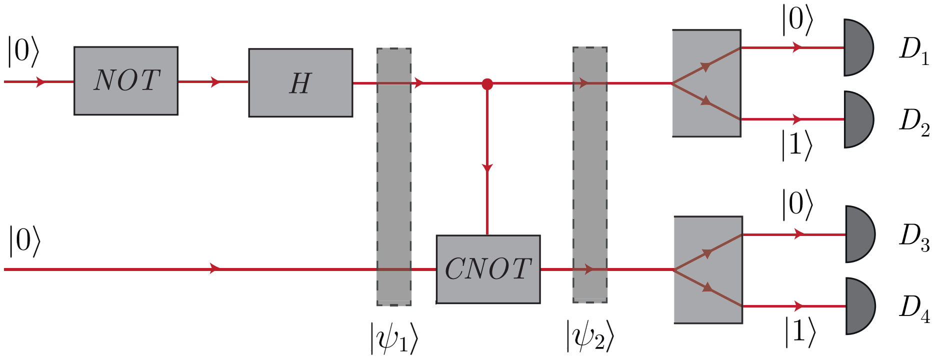


Figure 5.37: Quantum circuit

After the application of the Hadamard gate, the state becomes

$$\left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right)|0\rangle = \left(\frac{|0\rangle|0\rangle - |1\rangle|0\rangle}{2}\right)$$

This state is again separable. However, after the C-NOT gate is applied, the state becomes $\left(\frac{|0\rangle|0\rangle-|1\rangle|1\rangle}{2}\right)$ which is not separable. In this case, the final state is entangled.

$$|0\rangle|0\rangle \xrightarrow{Q-NOT} |1\rangle(|0\rangle) \xrightarrow{H} \left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right)|0\rangle \xrightarrow{CNOT} \frac{|0\rangle|0\rangle-|1\rangle|1\rangle}{\sqrt{2}}$$

5.4.2 Quantum Teleportation

All the data in the world whether classical or quantum is of no use if it can not be transferred or communicated. Quantum information transfer is called quantum teleportation and unlike classical data transfer, it is not the transfer of the particle from one spatial location to another. Quantum teleportation is the transfer of the quantum state from one particle to another particle which are spatially separated at a distance from one another.

Let's build a proper quantum circuit that can achieve quantum teleportation. This circuit will build upon the entanglement circuit from the previous section. For this new circuit, we will add another qubit. We now have three qubits in total, and the new qubit is in an unknown state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$. Our goal in this section is to create a replica of this state $|\psi\rangle$ at some other location.

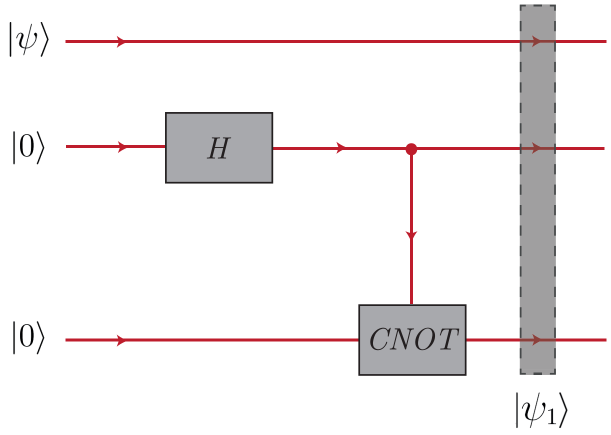


Figure 5.38: Entangled Qubit

In order to teleport the state $|\psi\rangle$, we will first need two qubits in an entangled state. If our initial qubits are $|0\rangle|0\rangle$, our entangled state will be $\frac{1}{\sqrt{2}}(|00\rangle +$

$|11\rangle$). The state of our three-level system at region 1 will be

$$\begin{aligned}|\psi_1\rangle &= \frac{1}{\sqrt{2}}(\alpha|0\rangle + \beta|1\rangle)(|00\rangle + |11\rangle) \\ &= \frac{1}{\sqrt{2}}(\alpha|000\rangle + \alpha|011\rangle + \beta|100\rangle + \beta|111\rangle)\end{aligned}$$

Alice and Bob join us again! This time, Bob takes one of the two entangled qubits and moves to a far-off planet. Alice has the other qubit of the entangled pair and the unknown qubit that we want to transfer. Let's name the unknown qubit's channel as channel 1 and Alice's entangled half's channel as channel 2.

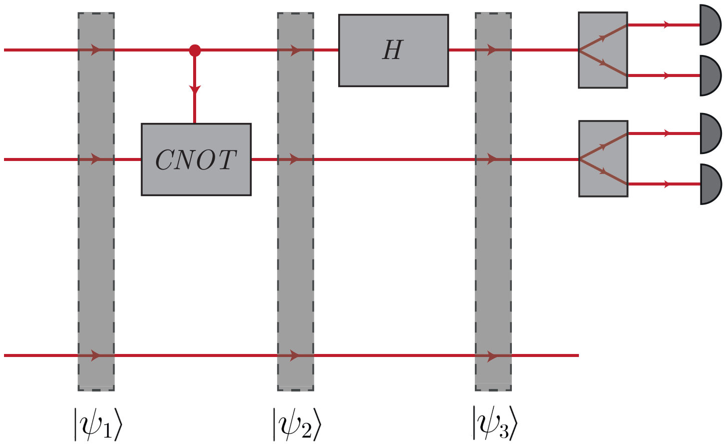


Figure 5.39: Quantum teleportation circuit

Alice applies a controlled not gate with qubit 1 (unknown qubit) as control and qubit 2 (entangled qubit) as target qubit $CNOT_{12}$. So the state of our three-qubit system in region 2 will be

$$|\psi_2\rangle = \frac{1}{\sqrt{2}}(\alpha|000\rangle + \alpha|011\rangle + \beta|110\rangle + \beta|101\rangle) \tag{5.14}$$

She then applies Hadamard gate on qubit $1(H_1)$ and the state in region 3 will be

$$\begin{aligned}|\psi_3\rangle &= \frac{1}{\sqrt{2}\sqrt{2}}(\alpha(|0\rangle + |1\rangle)|00\rangle + \alpha(|0\rangle + |1\rangle)|11\rangle + \beta(|0\rangle - |1\rangle)|10\rangle + \beta(|0\rangle - |2\rangle)|01\rangle) \\ &= \frac{1}{2}(\alpha|000\rangle + \alpha|100\rangle + \alpha|011\rangle + \alpha|111\rangle + \beta|010\rangle - \beta|110\rangle + \beta|001\rangle - \beta|101\rangle)\end{aligned}$$

If we separate Bob's qubit channel along with the unknown probability amplitudes, we can write $|\psi_3\rangle$ as

$$|\psi_3\rangle = \frac{1}{2}(|00\rangle(\alpha|0\rangle+\beta|1\rangle)+|01\rangle(\alpha|1\rangle+\beta|0\rangle)+|10\rangle(\alpha|0\rangle-\beta|1\rangle)+|11\rangle(\alpha|1\rangle+\beta|0\rangle))$$

Now Alice makes measurements on her qubits i.e. the first two qubits in our diagrams. There are four possibilities of measurements shown by the table below

Qubit 1	Qubit 2
0	0
0	1
1	0
1	1

Table 5.8: Alice's measurements

Note that all these possibilities are shown in the final form of $|\psi_3\rangle$. From this form, we can decipher what state would Bob's qubit will be in based on what Alice measured. Let's tabulate this as well

Alice's measurement	Bob's qubit's state
$\|00\rangle$	$\alpha\|0\rangle + \beta\|1\rangle$
$\|10\rangle$	$\beta\|0\rangle + \alpha\|1\rangle$
$\|01\rangle$	$\alpha\|0\rangle - \beta\|1\rangle$
$\|11\rangle$	$\beta\|0\rangle - \alpha\|1\rangle$

Table 5.9: Alice's measurements and Bob's qubit

Now that we know what qubit state Bob has, we just need to convert it to the state that we want to send to him i.e. $\alpha|0\rangle + \beta|1\rangle$. Alice sends her measurement results to Bob via a classical channel and based on these results, Bob applies gates to his qubit and converts it into the required state. If Alice's first qubit's measurement was 1, then Bob applies the Phase gate of πrad and if she got a 1 on the second qubit, then Bob applies a NOT gate to his qubit. For ease, we will turn this into a table too.

Alice measured	Bob applies
$\|00\rangle$	Nothing
$\|10\rangle$	NOT
$\|01\rangle$	Phase
$\|11\rangle$	NOT and Phase

Table 5.10: Bob's operations

One thing to note here is that in all this process, we have not created a clone of our qubit. Our original qubit state had to be destroyed in order to teleport it to Bob. This is due to the no-cloning theorem of quantum mechanics, which simply says that it is impossible to make a copy of a quantum state.

5.5 Quantum Algorithm

In previous sections, we saw how a quantum computer works and built some hardware for it. Now it's time to put that hardware to good use by giving it an algorithm to solve. This specific example that we will explore is from the early days of quantum information processing, it is called the Deutsch algorithm[2]. This algorithm will help us understand the advantage that a quantum computer has over a classical computer.

Let me first define the problem at hand. Binary functions can be of four types based on how they act on an input. These four types have been elaborated in the table below

Type 1	Type 2	Type 3	Type 4
$0 \mapsto 0$	$0 \mapsto 1$	$0 \mapsto 0$	$0 \mapsto 1$
$1 \mapsto 0$	$1 \mapsto 1$	$1 \mapsto 1$	$1 \mapsto 0$

The first two types are called constant functions as the output stays the same no matter what the input is, while the remaining two types are called balanced functions as the output will have an equal number of ones and zeros.

Our task now is to check whether a given binary function is constant or balanced. If we want to check it classically, we will have to run the function

---

[2] proposed by David Deutsch and Richard Jozsa in 1992

twice with different inputs and see if the outputs are the same or different. Deutch algorithm, on the other hand, checks it in a single go!

XOR Gate

To implement Duetch algorithm, we will need a XOR logic gate. So let's have a look at the functioning of the XOR gate before moving ahead. XOR between two values can be shown by the symbol $\oplus$ e.g. $a \oplus b$. Classical symbol and truth table for the XOR gate are shown below.

Inputs	Output
0 0	0
0 1	1
1 0	1
1 1	0

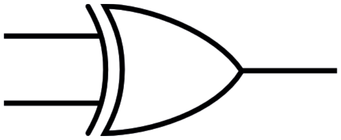


Figure 5.40: XOR Gate

Notice that if we have 0 on the first input, its XOR with the second input will give the value of the second input itself. While if we have 1 on the first input, the output will be the NOT of the second input.

Deutch Algorithm

Now we have a function $f(x)$, and we want to find out if it is a balanced or constant function. Let me show you the circuit diagram first, and then we will inspect it step by step.

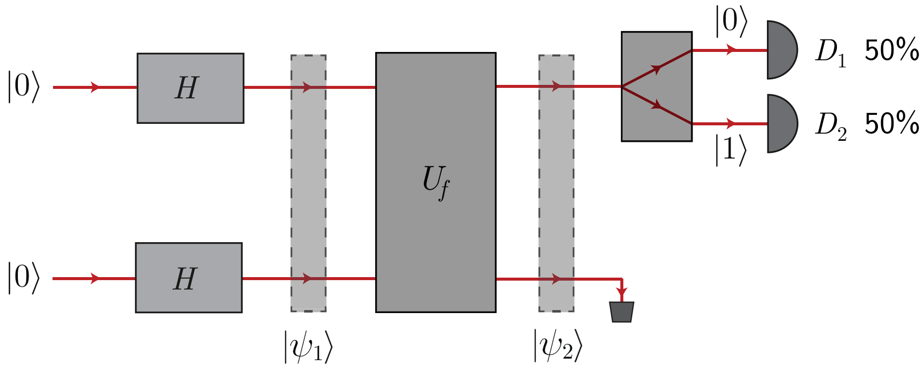


Figure 5.41: Circuit for Deutch Algorithm

In this diagram, U_f represents a controlled function like CNOT. Its task is to calculate $f(x)$ at the value of the first input qubit and XOR the result with the second input.

$$\begin{aligned}|00\rangle &\overset{U_f}{\longmapsto} |0\rangle|0\oplus f(0)\rangle\\ |01\rangle &\overset{U_f}{\longmapsto} |0\rangle|1\oplus f(0)\rangle\\ |10\rangle &\overset{U_f}{\longmapsto} |1\rangle|0\oplus f(1)\rangle\\ |11\rangle &\overset{U_f}{\longmapsto} |1\rangle|1\oplus f(1)\rangle\end{aligned}$$

Notice that the first qubit remains unchanged. In general, it can be written as

$$|ab\rangle \overset{U_f}{\longmapsto} |a\rangle|a\oplus f(b)\rangle$$

We always start with $|0\rangle$ as the input qubit and $|1\rangle$ as an ancillary qubit. After the application of Hadamard gate, state $|\psi_1\rangle$ at region 1 will be

$$\begin{aligned}|01\rangle \overset{H}{\longmapsto} |\psi_1\rangle &= \left(\frac{|0\rangle+|1\rangle}{\sqrt{2}}\right)\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right)\\ &= \frac{1}{2}(|00\rangle-|01\rangle+|10\rangle-|11\rangle)\end{aligned}$$

State $|\psi_2\rangle$ at region 2 after the application of U_f will be

$$\xrightarrow{U_f} |\psi_2\rangle = \frac{1}{2}(|0\rangle|0 \oplus f(0)\rangle - |0\rangle|1 \oplus f(0)\rangle + |1\rangle|0 \oplus f(1)\rangle - |1\rangle|1 \oplus f(1)\rangle)$$

But we know that XOR of 0 with anything is that thing itself, so $|\psi_2\rangle$ can be written as

$$|\psi_2\rangle = \frac{1}{2}(|0\rangle|f(0)\rangle - |0\rangle|1 \oplus f(0)\rangle + |1\rangle|f(1)\rangle - |1\rangle|1 \oplus f(1)\rangle)$$

We can write it in a compact form. First, we will consider the terms that depend on $f(0)$.

$$|0\rangle|f(0)\rangle - |0\rangle|1 \oplus f(0)\rangle$$

Function $f(x)$ is a binary function and can only have 0 or 1 as output. Let's consider first scenario where $f(0) = 0$. The two terms then become

$$\begin{aligned} &= |0\rangle|0\rangle - |0\rangle|1\rangle \\ &= |0\rangle(|0\rangle - |1\rangle) \end{aligned}$$

The second scenario would be $f(0) = 1$ and the first two terms would become

$$\begin{aligned} &= |0\rangle|1\rangle - |0\rangle|0\rangle \\ &= |0\rangle(|1\rangle - |0\rangle) \\ &= -|0\rangle(|0\rangle - |1\rangle) \end{aligned}$$

We see that the results of the two scenarios differ by only a negative sign. So, we can write the two scenarios in a general form

$$(-1)^{f(0)}|0\rangle(|0\rangle - |1\rangle)$$

Using the same analogy, we can write the terms containing $f(1)$ as

$$(-1)^{f(1)}|1\rangle(|0\rangle - |1\rangle)$$

State $|\psi_2\rangle$ can then be written as

$$\begin{aligned} |\psi_2\rangle &= \frac{1}{\sqrt{2}}\left((-1)^{f(0)}|0\rangle\left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) + (-1)^{f(1)}|1\rangle\left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right)\right) \\ &= \left(\frac{(-1)^{f(0)}|0\rangle + (-1)^{f(1)}|1\rangle}{\sqrt{2}}\right)\left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) \end{aligned}$$

The terms in first parenthesis of the above equation represent the state of first qubit and the terms in second parenthesis represent the state of second qubit. We dump the second qubit as we do not need it anymore and shift our focus to the second qubit only.

If our function was a constant function ($f(0) = f(1)$), the state of the first qubit at region 2 will be $\pm\left(\frac{|0\rangle+|1\rangle}{\sqrt{2}}\right)$. On the other hand, if the function was a balanced function ($f(0) \neq f(1)$), state of first qubit at region 2 will be $\pm\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right)$. We can ignore the $\pm$ as it is just a phase. So our first qubit's state is either $\left(\frac{|0\rangle+|1\rangle}{\sqrt{2}}\right)$ (constant case) or $\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right)$ (balanced case) at region 2. When we apply Hadamard gate on our first qubit, the state changes to $|0\rangle$ in case of constant function and to $|1\rangle$ in case of balanced function. Now we can easily measure this state and determine the property of our function. Notice that this took us only one shot to determine the nature of the function i.e. we only gave 0 as input and did not have to run the function for input 1. The fun part is that this circuit can easily be modified for n qubits.

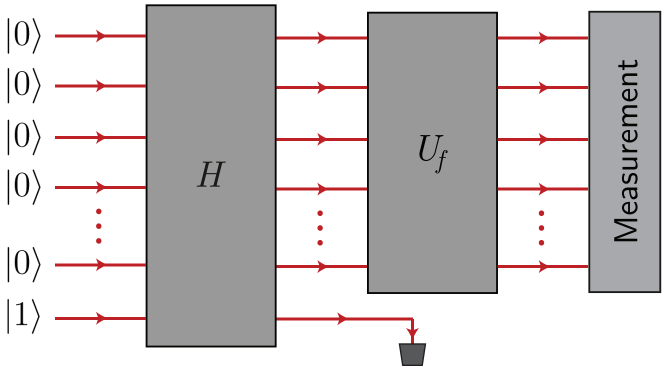


Figure 5.42: Circuit for Deutch Algorithm

Using this circuit we find the nature of function in just one try while classically, it would have taken us $2^{n-1} + 1$ experiments to determine it. This is one example where we can harvest the power of quantum computers.

5.6 Error Correction

So far, we have seen how to build simple quantum computers and designed an algorithm for them as well. But is it all so simple? Is it this easy to implement it in real life? No! Everything is prone to errors and so are our quantum circuits. In this section, we will explore three different types of errors and how to get rid of them.

5.6.1 Decoherence Free Subspace

Decoherence means the loss of information from a quantum system due to interaction with the environment.

Let's first observe a portion of a quantum circuit, a subspace, that is prone to decoherence. We need to implement logical 0 and 1 in the form of qubits $|0\rangle_L$ and $|1\rangle_L$. The problem with this subspace of ours is that if we use $|0\rangle$ and $|1\rangle$ as our logical 0 and 1, it induces decoherence in our subspace by adding a phase to $|1\rangle$ and leaving $|0\rangle$ as it is.

$$|0\rangle \xrightarrow{decoherence} |0\rangle$$

$$|1\rangle \xrightarrow{\text{decoherence}} e^{i\phi}|1\rangle$$

This means that if we want to start from the state $\alpha|0\rangle + \beta|1\rangle$, this state will change due to decoherence in the following fashion.

$$\alpha|0\rangle + \beta|1\rangle \longrightarrow \alpha|0\rangle + e^{i\phi}\beta|1\rangle$$

Both these states are distinct from each other. Our goal now is to chose our logical 0 and 1 qubit states in such a way that our subspace does not feel the effect of this decoherence. This means that the state we get after passing through the system is equivalent to state that we initially had. A decoherence-free subspace.

Here is one such set of logical qubit states.

$$|0\rangle_L = \frac{1}{\sqrt{2}}(|01\rangle + i|10\rangle) \tag{5.15}$$

$$and \tag{5.16}$$

$$|1\rangle_L = \frac{1}{\sqrt{2}}(|01\rangle - i|10\rangle) \tag{5.17}$$

Notice that we have used two qubits instead of one to simulate a logical 0 or 1 state; it's completely fine to do it this way. Now we apply the decoherence model to these particular states. If we have a $|1\rangle$ on the second or first qubit, we will pick up a phase.

$$|0\rangle_L \xrightarrow{decoherence} \frac{1}{\sqrt{2}}(e^{i\phi}|0\rangle|1\rangle + ie^{i\phi}|1\rangle|0\rangle)$$

and

$$|1\rangle_L \xrightarrow{decoherence} \frac{1}{\sqrt{2}}(e^{i\phi}|0\rangle|1\rangle - ie^{i\phi}|1\rangle|0\rangle)$$

The evolution of our overall state will be

$$\alpha|0\rangle_L + \beta|1\rangle_L \xrightarrow{decoherence} e^{i\phi}(\alpha|0\rangle_L + \beta|1\rangle_L) \tag{5.18}$$

The quantum state is only picking up a global phase when we encode the information into a bigger space. This global phase does not matter and no information is lost. This type of encoding does not see decoherence; this is uniform decoherence of phase and is called phase decoherence.

5.6.2 Robust Quantum Computation

The robust way to deal with errors is to design the gates and circuits in such a way that they compensate for the errors on a hardware level. Let's consider an example. We have a state $|0\rangle$ and we want to apply π *rad* rotation about $y - axis$ on it i.e. a NOT gate. However, the way we perform this rotation (like passing it through a magnetic field for a small amount of time), we always get some error in the amount of rotation. So there is some error in the term π in every experiment and it is never exactly π, but $\pi(1 - \epsilon)$. We are falling short by a factor of ϵ and the rotation is incomplete.

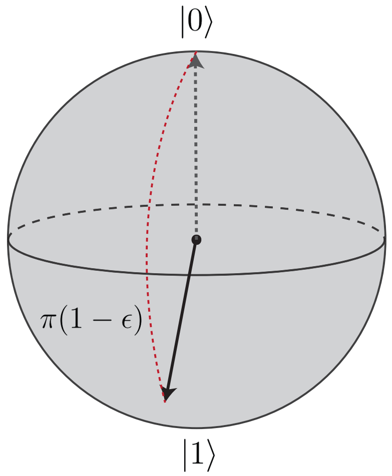


Figure 5.43: Incomplete rotation

To solve this problem, we will implement our rotation in such a way, that we bypass this error. We begin by rotating the state by $\frac{\pi}{2}$ instead of π, but because we have an error we do not reach the exact half and stay in the upper hemisphere. Next up, we apply a π rotation about the $x-axis$. This way, we switch to the lower hemisphere and are ahead of the exact half, by the distance that we were short of it previously. Now we again rotate about $y-axis$ by $\frac{\pi}{2}$ and complete the π rotation that we intended to do.

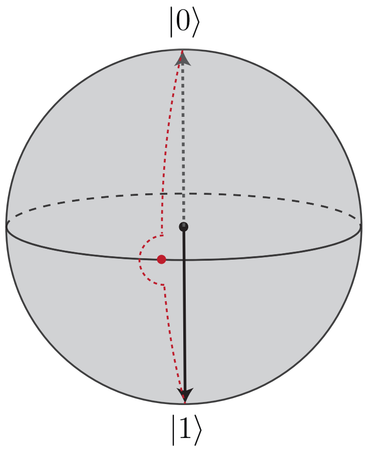


Figure 5.44: Complete rotation

$U_y(\frac{\pi}{2})U_x(\pi)U_y(\frac{\pi}{2})|0\rangle = |\phi\rangle$

This works for some particular initial states. If you want to implement a general quantum NOT gate that is robust against errors, you need to perform rotations like this.

$U_{60}(\pi)U_{300}(\pi)U_{60}(\pi)|0\rangle = |\phi\rangle$

5.6.3 Quantum Error Correction

During the transmission of an arbitrary qubit such as $\alpha|0\rangle + \beta|1\rangle$, different kinds of errors might occur. One possibility is that the bit flips, i.e. $|0\rangle$ becomes $|1\rangle$, and $|1\rangle$ becomes $|0\rangle$. Quantum error correction deals with these bit flips. In this method, we will encode our logical states $|0\rangle_L$ and $|1\rangle_L$ as $|000\rangle$ and $|111\rangle$ respectively. We give these states to Alice and ask her to send them to Bob. The problem is that when Bob receives these states, there is a possibility that one of the bits would have flipped on its way. It could be unchanged $(\alpha|000\rangle + \beta|111\rangle)$, or one of the qubits could have flipped and the state is $(\alpha|100\rangle + \beta|011\rangle)$, $(\alpha|010\rangle + \beta|101\rangle)$, or $(\alpha|001\rangle + \beta|110\rangle)$. To deal with it, Bob takes two ancillary qubits. He first carries out controlled-NOT operation from the first and second received qubits to the first ancillary qubit, then from the first and third received qubits to the second ancillary qubit.

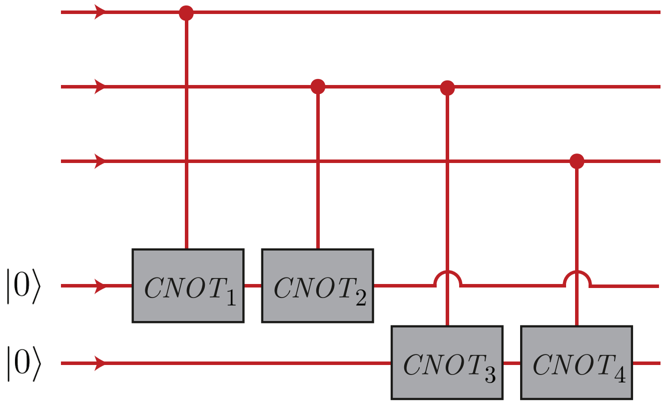


Figure 5.45: Quantum error correcting circuit

Let's see what will happen if the first qubit had flipped on its way.

$$\begin{aligned}
(\alpha|100\rangle + \beta|011\rangle) \otimes |00\rangle &\xrightarrow{CNOT_1} \alpha|100\rangle|10\rangle + \beta|011\rangle|00\rangle \\
&\xrightarrow{CNOT_2} \alpha|100\rangle|10\rangle + \beta|011\rangle|10\rangle \\
&\xrightarrow{CNOT_3} \alpha|100\rangle|10\rangle + \beta|011\rangle|11\rangle \\
&\xrightarrow{CNOT_4} \alpha|100\rangle|10\rangle + \beta|011\rangle|10\rangle \\
&= (\alpha|100\rangle + \beta|011\rangle)|10\rangle
\end{aligned}$$

Bob measures the ancillary qubits and if he gets $|10\rangle$, that the first qubit had flipped. Similarly, $|01\rangle$ would mean that the second qubit had flipped, $|11\rangle$ means that the third qubit had flipped, and $|00\rangle$ means that the state is in its original form. If we know what bit has flipped, its problem solved! Now we just need to apply NOT gate to the flipped qubit and get it back to its original state.

Appendix A

Linear Algebra