स्वाध्याय स्वमन्थन स्वावलम्बन

UTTAR PRADESH RAJARSHI TANDON OPEN UNIVERSITY

(Established vide U.P. Govt. Act No. 10, of 1999)

UGPHS-08
MODERN PHYSICS

FIRST BLOCK
The Special Theory of Relativity

INDIRA GANDHI NATIONAL OPEN UNIVERSITY

UTTAR PRADESH
RAJARSHI TANDON OPEN UNIVERSITY

Shantipuram (Sector-F), Phaphamau, Allahabad-211013

UTTAR PRADESH
RAJARSHI TANDON OPEN UNIVERSITY

UGPHS-08
Modern Physics

Block

1

THE SPECIAL THEORY OF RELATIVITY

UNIT 1

Emergence of Special Relativity 5

UNIT 2

Relativistic Kinematics 22

UNIT 3

Relativistic Dynamics 50

MODERN PHYSICS

So far your study of physics has been restricted to the classical domain – the laws of nature that you have studied are the laws of classical physics. This means that you have studied physical phenomena in the macroscopic world, and in particular, their gross features. For example, using these laws, the motion of a macroscopic system consisting of pulleys, flywheels, levers, etc. can be described if the relevant parameters, viz., the density and modulus of elasticity of the material are given. However, if you ask why the densities and elastic constants have the values they have, the laws of classical physics are silent. Similarly, if we wish to know why sodium vapour emits yellow light, what makes the sun shine, why the uranium nucleus disintegrates spontaneously or what happens when objects travel at speeds close to the speed of light, classical physics does not provide us the answers. We then turn to new areas of knowledge, namely, the special theory of relativity, quantum mechanics and nuclear physics. These areas of physics embody concepts which are foreign to our everyday experience. They are the most remarkable intellectual creations of the early twentieth century physics.

Through this course on Modern Physics we extend you an invitation to join us in an adventure of the mind. We promise to take you on a guided tour of these areas which can be regarded as the fundamental conceptual structures of twentieth century physics. The going may be a little tough but the rewards will be rich – intellectually stimulating and deeply satisfying!

To give you a glimpse of the intellectual 'feast' awaiting you, we recount here an adventure of Mr. Tompkins, the hero of George Gamow's book entitled Mr. Tompkins in Wonderland. George Gamow is a well- known physicist of this century. He was a proponent of the Big Bang theory of the origin of the universe.

Mr. Cyril George Henry Tompkins, after listening to a popular lecture on the theory of relativity, dreams of a visit to a fantastic city in which the speed of light is only 25 km h^{-1}. What does he observe there? At first, nothing unusual seems to happen around him– a policeman standing on the corner looks as policemen usually do! The streets are nearly empty. But when a cyclist coming down the street approaches Mr. Tompkins, he is absolutely astonished. For, the bicycle and the man on it appear unbelievably flattened to him. When the clock strikes twelve, the cyclist, who seems to be in a hurry, pedals harder. Though he does not gain much in speed, he appears flattened even more. Mr. Tompkins decides to overtake the cyclist and ask him about it. He borrows a bicycle and pedals on it hoping to get flattened. But amazingly, nothing happens to him. Instead, the picture around him changes completely. The streets grow shorter, the windows of the shops begin to look like narrow slits and the policeman on the corner becomes the thinnest man he had ever seen!

What a topsyturvy world Mr. Tompkins visits in his dreams! And what weird experiences he has! We are sure you will be interested in finding an explanation for these happenings in Mr. Tompkins' dream world. This course will help you do so. We hope you enjoy studying this course and wish you good luck.

BLOCK 1 THE SPECIAL THEORY OF RELATIVITY

Albert Einstein (1877–1955)

What is it that you think of when you hear or see the word **relativity**? The name of Albert Einstein? Or the equation $E = mc^2$? Or a vision of astronauts who return young from their trips to space lasting many years? This signifies the enormous intellectual impact (even after almost a hundred years) of what Einstein called his **special theory of relativity**. The development of this theory is rightly regarded as one of the greatest strides ever made in our way of understanding the physical world. And yet the basic concept of relativity is as old as the mechanics of Galileo and Newton. So what did Einstein do to make his name almost synonymous with relativity?

At the beginning of the twentieth century, two great and beautiful theories were known in the physical sciences – Newtonian mechanics and Maxwell's electrodynamics. Both of them gave a unified explanation of countless physical phenomena. These theories were expressed in concise mathematical language within a certain conceptual framework. You have studied both these theories in Block 1 of PHE-01 (Elementary Mechanics) and Block 4 of PHE-07 (Electric and Magnetic Phenomena), respectively. You have also learnt of the numerous applications of these theories. Both theories have been confirmed many a times; and they have been extraordinarily successful in their predictions. And yet these two theories were *conceptually* in contradiction with one another!

What was this contradiction? Unit 1 of this block explains it as well as the dilemma which occupied the best minds of the time in the physical sciences. How was this contradiction resolved? Its solution was provided by none other than Albert Einstein. He resolved the contradictions as he saw them and formulated a new theory based on two new principles. These two principles led Einstein to a new view of space and time about which you will study in Unit 2. Naturally, a radically new attitude to the notions of space and time resulted in changes in well-established areas of physics. In Unit 3, we shall discuss the new mechanics that replaced Newtonian mechanics as a result of these changes.

What we had, in effect, was an Einsteinian revolution. Its impact far exceeds that of the Copernican revolution and has rarely been equalled in the history of physics. We have tried our best to bring to you the beauty and logic of the special theory of relativity. We hope that you will appreciate and enjoy studying it as much as we did presenting it. We wish you good luck.

UNIT 1 EMERGENCE OF SPECIAL RELATIVITY

Structure

1.1 Introduction
Objectives

1.2 Classical Relativity
Galilean Coordinate Transformations
Galilean Principle of Relativity

1.3 Electromagnetism and Classical Relativity
Problems of Relativity vis-a-vis Laws of Electromagnetism
Galilean Relativity and the Speed of Light
Attempts to Locate the Absolute Frame – The Michelson-Morley Experiment

1.4 The Special Theory of Relativity
The Principle of Relativity
The Principle of Constancy of Speed of Light

1.5 Summary

1.6 Terminal Questions

1.7 Solutions and Answers

1.1 INTRODUCTION

You have studied Newtonian mechanics in your school science courses and in your first physics elective entitled 'Elementary Mechanics' (PHE-01). You are familiar with the concept of inertial frames of reference. You know that Newton's laws of motion are the same in all inertial frames of reference. You must have recognized the validity of this statement in everyday life. An object moves in the same way in a uniformly moving train or an aeroplane as it does on earth. For instance, when you walk, drop a coin or throw a ball up in the air while riding in such a train or aeroplane, the bodies move just as they do on earth. Both Galileo and Newton were deeply aware of this principle that the laws of mechanics are the same in all inertial reference frames; this is the classical principle of relativity. So the classical notion of relativity is not new to you. However, you have not encountered this terminology before. Therefore, we shall begin this unit with a brief review of the classical notion of relativity as embodied in the works of Galileo and Newton.

You know that Newtonian Mechanics was highly successful in describing motion in the world of our everyday experiences. Then why did the need arise for re-examining Newtonian mechanics and the notion of relativity it contained? The need arose when the classical principle of relativity was applied to the propagation of electromagnetic waves and that led to certain inconsistencies. In Sec. 1.3, you will learn about some of these inconsistencies and find that the Newtonian relativistic world view could not easily incorporate the laws of electromagnetism. The question is: What replaced it? It was replaced by a radically different way of understanding the world when, in 1905, Albert Einstein proposed his special theory of relativity. In the last section (1.4) of this unit you will study the main features of this theory.

"What I see in Nature is a magnificent structure which we can comprehend only very imperfectly, and that must fill a thinking person with a feeling of humility."

– Albert Einstein, 1944/5

In brief, what we intend to say in this unit is this: Einstein was not the first to introduce relativistic notions in physics. What he did was to generalise the classical notion of relativity (applicable only to mechanics) to all physical phenomena. Although we will go into some detail of the background in which Einstein's special relativity emerged, it will not necessarily be a historical description. We will simply bring out the factors which induced scientists to change their concepts in so radical a manner. In this process we hope that you will be able to appreciate and understand the special theory of relativity much better.

In the next unit, you will learn about the consequences of the special theory of relativity. In particular, you will understand in what way special relativity altered the established notions of space and time.

Objectives

After studying this unit you should be able to

- use the Galilean coordinate transformations to describe events in different inertial frames of reference
- explain the Galilean principle of relativity and state why it became necessary to generalise it
- state the postulates of special theory of relativity
- apply the principle of relativity to physical phenomena
- compare the nature of time in classical relativity and the special theory of relativity.

Study Guide

This unit presents the background out of which the special theory of relativity emerged. Therefore, we shall be using many concepts and ideas from our earlier physics courses. We strongly advise you to go through the Block 1 of PHE-01 (Elementary Mechanics), Block 4 of PHE-07 (Electric and Magnetic Phenomena) and Block 2 of PHE-09 (Optics) before studying this unit. It will help you in understanding the ideas presented in Sec. 1.2 and 1.3 better, and in less time. In our estimate, you should take about 6 to 7h to complete this unit.

1.2 CLASSICAL RELATIVITY

You may like to refer to Sec. 2.2.1 of Unit 2, PHE-01 where the concept of an inertial observer has been discussed in detail. What we have said there for inertial observer applies to an inertial frame of reference.

You have studied the concepts of inertial frames of reference and relative motion in Unit 1 of PHE-01. You are familiar with the relationship between the velocity and acceleration of an object measured with respect to two inertial frames in uniform relative motion. You have also studied the laws of Newtonian Mechanics in Unit 2 of PHE-01. Here we shall use these concepts to explain briefly the notion of classical or Galilean relativity.

Let us begin by considering a physical event. An idealised version of an event is that it is something that happens at a point in space and at an instant in time. While discussing the theory of relativity, Einstein often used this dramatic example of an event—lightning strikes the ground. A small explosion is an equally dramatic event. You can think of several other examples of an event. There are two basic questions that we can ask about any event:

Where did it take place?

When did it take place?

How do we answer these questions? As you are well aware, we specify an event by four measurements in a particular frame of reference — three for the position and one for the time t. We usually fix the position of the event by the Cartesian coordinates (x, y, z). You have used the Cartesian coordinate system quite often in your physics elective courses. For example, two particles may collide at $x = 1\text{m}$, $y = 2\text{m}$, $z = 3\text{m}$ and at time $t = 4\text{s}$ in one frame of reference such as a laboratory on the earth. Then the four numbers (1, 2, 3, 4) specify the event in that reference frame; the first three numbers specify its position and the fourth the time at which it occurred.

Thus, we must first establish a **frame of reference** to accurately describe where and when an event happens. You know that for describing an event we are free to use any frame of reference we wish. In this course we shall restrict our study to what are called **inertial** reference frames. Recall that

> An inertial frame is a frame of reference in which Newton's first law holds true.

So in an inertial frame of reference, objects at rest remain at rest and objects moving uniformly in a straight line continue to do so, unless acted upon by a net external force. From this concept you can readily conclude that

> **Any frame that moves with constant velocity relative to an inertial frame is also an inertial frame.**

Would you like to test whether you understand the concept of an inertial frame of reference before studying further? If so, try the following SAQ.

Spend 2 min.

SAQ 1

Classify the following frames of reference as inertial and non-inertial (i.e., frames which are not inertial). The frames attached to

a) a car in circular motion

b) spaceships cruising uniformly

c) an electron accelerating in an electric field

d) a boat moving at a constant speed in a river

e) an apple at rest on a fixed table in your room.

Suppose now that we have made space and time measurements describing an event in one inertial frame of reference. We want to describe the same event in another inertial frame of reference. For example, consider the following event. A boy throws a ball vertically upwards in a train moving at a uniform velocity with respect to the ground. In the frame of reference attached to the train, the ball goes straight up and comes down along the same path. Now how do we describe the ball's motion in another frame of reference attached to the ground?

We can use the Galilean coordinate transformations to describe an event in different inertial frames of reference. Let us briefly study the Galilean coordinate transformations.

1.2.1 Galilean Coordinate Transformations

Consider an inertial frame S and another frame S' which moves at a constant velocity u with respect to S (Fig. 1.1). We define the x-axis (and the x'-axis) to be along the direction of motion. We assume the other two axes (y, z) and (y',z') to be parallel to each other — y parallel to y' and z parallel to z'. Further, we define the origin of time, $t = 0$, to be the instant when the origins of the two coordinate systems coincide, i.e., when point O' coincides with O.

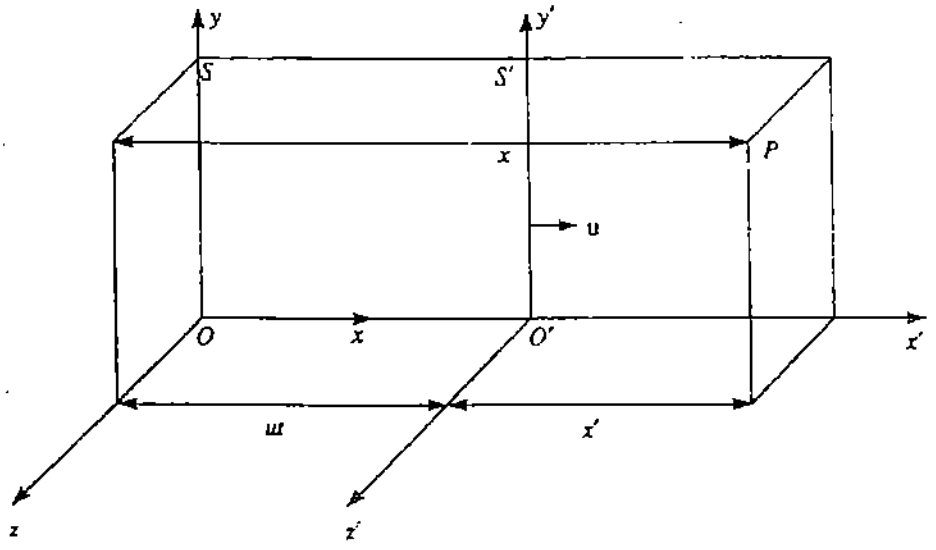


Fig.1.1: Two inertial frames of reference S and S'. S' moves with a constant velocity u (= u i) with respect to S so that the x-x' axis is common and the y-y', z-z' axes are parallel. As seen from frame S', S moves with a velocity $-u$, i.e., at a speed u in the negative x direction. Point P represents an event whose space-time coordinates can be measured by observers in S and S'. The origins O and O' coincide at time $t=0$ and $t'=0$. You can see that $x = x' + ut$, $y' = y$ and $z' = z$.

Suppose that an event E occurs at point P. Let us assume that any measurement in the two frames of reference are being made by observers who have jointly calibrated their metre

sticks and clocks. The observer attached to S ascribes the coordinates x, y, z, t to P and the observer attached to S' specifies the same event by x', y', z' and t'. The coordinates (x, y, z) give the position of P relative to O as measured by observer S and t is the time at which E occurs as recorded by the clock of S. The coordinates (x', y', z') give the position of P with respect to O' and t' is the time at which E occurs according to the clock of S'. For simplicity, the clocks of each observer read zero at the instant that the origins O and O' of the frames S and S' coincide.

What is the relationship between (x, y, z, t) and (x', y', z', t')? The **Galilean coordinate transformation** (see Fig. 1.1) relates these measurements as follows:

$$\begin{aligned} x' &= x - ut \\ y' &= y \\ z' &= z \\ t' &= t \end{aligned} \tag{1.1}$$

We can write the generalized Galilean transformation in vector notation as follows:

$$\mathbf{r}' = \mathbf{r} - \mathbf{u}t \tag{1.2a}$$

$$t' = t \tag{1.2b}$$

where $\mathbf{r}$ is the position vector of P in S and $\mathbf{r}'$ in S'. Under our simplifying assumptions $\mathbf{u} = u\hat{\mathbf{i}}$, and Eqs. (1.2a, b) reduce to Eqs. (1.1).

Differentiating Eq. (1.2a) with respect to t gives

$$\frac{d\mathbf{r}'}{dt} = \frac{d\mathbf{r}}{dt} - \mathbf{u} = \mathbf{v} - \mathbf{u}$$

But since $t = t'$, $\frac{d\mathbf{r}'}{dt} = \frac{d\mathbf{r}'}{dt'} = \mathbf{v}'$. Hence, we get

$$\mathbf{v}' = \mathbf{v} - \mathbf{u} \tag{1.3a}$$

Differentiating Eq. (1.3a) with respect to t and using Eq. (1.2b) we get

$$\mathbf{a}' = \mathbf{a}, \text{ since } \mathbf{u} \text{ is constant.} \tag{1.3b}$$

You have encountered Eqs. (1.1a and b) in Sec. 1.5 of Unit 1, PHE-01. These are the Eqs. (1.53) and (1.38) given there.

The equation of motion is then given as

$$m\,\mathbf{a}' = m\,\mathbf{a} = \mathbf{F} \tag{1.4}$$

This means that we obtain the same law of motion in the frame S' as in the frame S.

In relation to Eq. (1.4) we would like to ask another question. How does the force $\mathbf{F}$ transform when we go from one frame to another? You know that forces considered in mechanics depend either on distance (gravitational forces, elastic forces) or on relative velocity (friction forces) and on time interval. So let us find out how distance, relative velocity and a time interval change under the Galilean coordinate transformations.

Suppose we investigate two objects P and Q. Let the force of their interaction depend on the distance between them, their relative velocity and time. From Eqs. (1.1) we can at once see that the distance between P and Q, measured at the same instant is the same in S and S':

$$x'_P - x'_Q = x_P - x_Q, \quad y'_P - y'_Q = y_P - y_Q, \quad z_P' - z_Q' = z_P - z_Q$$

or in vector notation

$$\mathbf{r}'_P - \mathbf{r}'_Q = \mathbf{r}_P - \mathbf{r}_Q \tag{1.5a}$$

On differentiating Eq. (1.5a) with respect to time we find that the relative velocity of P with respect to Q remains the same in both the frames of reference.

$$\mathbf{v}'_P - \mathbf{v}'_Q = \mathbf{v}_P - \mathbf{v}_Q \tag{1.5b}$$

Remember that in arriving at Eq. (1.5b) we have also used the fact that Galilean transformation does not change time and so also the time interval between any two events say A, and B:

$$t'_A - t'_B = t_A - t_B \qquad (1.6)$$

Hence, we can conclude that forces occurring in mechanics, that depend on time intervals, distance and relative velocity, do not change under the Galilean transformation. We say that forces remain invariant under Galilean transformation. Thus, all the quantities appearing in Eq. (1.4) do not change under the Galilean transformation. Therefore, the fundamental equation of classical mechanics – Newton's second law – has the same form in a stationary frame S as in a frame S' moving with constant velocity with respect to S. With this information at our command, we are now ready to present the classical principle of relativity. It is also called the Galilean principle of relativity, since it was Galileo who first enunciated it, although its mathematical basis given above was provided only later by Isaac Newton.

1.2.2 Galilean Principle of Relativity

Eqs. (1.5a and b) and (1.6) tell us that according to the Galilean transformations the time interval, space interval (distance) and relative velocity measurements and hence the force law in mechanics is the same in all inertial frames. The relative velocity of the frames can be arbitrary and does not affect these results. Implicit in Eq. (1.4) is the basic postulate of classical mechanics that the mass of a body is constant, i.e., it is an invariant quantity.

So what do Newtonian mechanics and Galilean transformation put together imply? The length, mass and time – the three basic quantities in mechanics – as well as forces (which depend upon time interval, space interval and relative velocity) are independent of the relative motion of an inertial observer. The laws of mechanics hold good in all inertial frames of reference. Thus, we arrive at the classical or Galilean principle of relativity:

> **The laws of mechanics can be written in the same form in all inertial frames. If they hold in one inertial frame, they will also hold in all other inertial frames.**

It is a limited principle of relativity in that it applies to only the laws of mechanics. Let us understand with the help of a simple example what the Galilean principle of relativity means. Suppose you are in a car which is moving at a constant speed and you cannot look out. Then to you, all mechanical experiments performed and all mechanical phenomena occurring in the car will appear the same as if the car were not moving. For instance, a ball thrown vertically upwards will always fall down along the same path.

No mechanical experiment performed in the car could help you determine whether it was moving uniformly or was at rest provided, of course, you did not look out. This is what we mean when we say that if the laws of mechanics are true in one inertial frame of reference, they will be true and of the same form in any other inertial frame as well. Thus, as far as mechanics is concerned there is no preferred inertial frame of reference in which alone the classical laws have the most basic form. Therefore, there is no absolute frame of reference.

You may like to pause for a while and find out whether you have understood these ideas. Try the following SAQ.

SAQ 2

a) Does the fact that Eq. (1.4) is invariant under Galilean transformation mean that all inertial observers will measure the same values for the position, time, velocity, energy and momentum corresponding to an event?

b) Are the laws of conservation of linear momentum and energy invariant under Galilean transformation?

An interesting aspect of the classical principle of relativity pertains to the nature of space and time. And we would like you to know about it.

Absolute Space and Absolute Time

You have just studied that according to Newtonian mechanics and Galilean relativity, the measurements of length (relative position), mass, time and their relationship are independent of the relative motion of an inertial observer. They do not depend on which

inertial observer measures them. This implies that there exist absolute space intervals and absolute time intervals in Newtonian mechanics. In other words we may say that space and time exist in themselves and have properties that do not depend upon anything else.

In Newton's own words: "Absolute space, in its own nature, without relation to anything external, remains always similar and immovable." So according to Newton, space represents a giant empty box which contains material objects and various physical phenomena take place in it, and it does not get affected by these. Likewise time is thought to flow absolutely, uniformly and evenly without being affected by any actual events that happen as time passes. Once again we quote Newton: "Absolute, true and mathematical time, of itself, and from its own nature, flows equably without relation to anything external, and is otherwise called duration".

Thus, in Newton's world-view space and time are things external in relation to nature. Further, there is no relationship between space and time – the properties of space are determined independently of movements of objects with the passage of time; the flow of time is independent of the spatial properties of such objects.

In a nutshell, as per Newtonian ideas, space and time exist by themselves, independent of each other and do not depend on material bodies located in space or physical phenomena occurring therein.

We have briefly introduced these ideas here so that you are able to appreciate the difficulties that arose when the classical principle of relativity was applied to electricity, magnetism and optics. Let us now study how classical relativity proved to be inconsistent with the laws of electromagnetism.

1.3 ELECTROMAGNETISM AND CLASSICAL RELATIVITY

You have thus far studied that the Galilean principle of relativity applies to mechanical phenomena. The question that arises next is: Do other laws of physics (e.g., the laws of electromagnetism and optics) have the same form in all inertial frames? In other words, are they invariant under a Galilean transformation?

In fact, when the principle of relativity was applied to Maxwell's equations, certain problems arose immediately — they did not seem to obey it. Let us briefly outline some of these problems.

1.3.1 Problems of Relativity vis-a-vis Laws of Electromagnetism

Let us first consider a simple example of two equal, positive point charges carrying charge q, as shown in Fig. 1.2a. We will first examine the system as seen by an observer in the reference frame S'.

As you can see in Fig. 1.2a one charge rests at the origin of S' and the other rests at a distance y_2' on the y' axis of S'. From Maxwell's equations we can determine the electromagnetic force that the charges at rest exert on each other in S': it is just the electrostatic Coulomb force of magnitude $F_C = \frac{1}{4\pi\epsilon_0}\frac{q^2}{y_2'^2}$.

Let us now consider the electromagnetic force from the point of view of S. This observer sees the charge q unchanged and $y_1 = y_2'$. So Coulomb's force law is unchanged. However, the observer in S also sees both charges moving to the right at a speed v. Now two positive charges moving to the right constitute two conventional parallel currents which attract each other. Therefore, the total force in S has two components – the electrostatic force of repulsion and the attractive force between parallel currents. We find that it is different from the force in S'. But according to Newtonian physics, these forces should be the same. This is an inconsistency (Fig. 1.2b).

Another problem arises when we try to transform Maxwell's equations according to Galilean coordinate transformations — they change their form. For example, the wave equation for electromagnetic fields, deduced from Maxwell's equations does not remain the same (See Unit 14 of PHE-07 entitled Electric and Magnetic Phenomena for

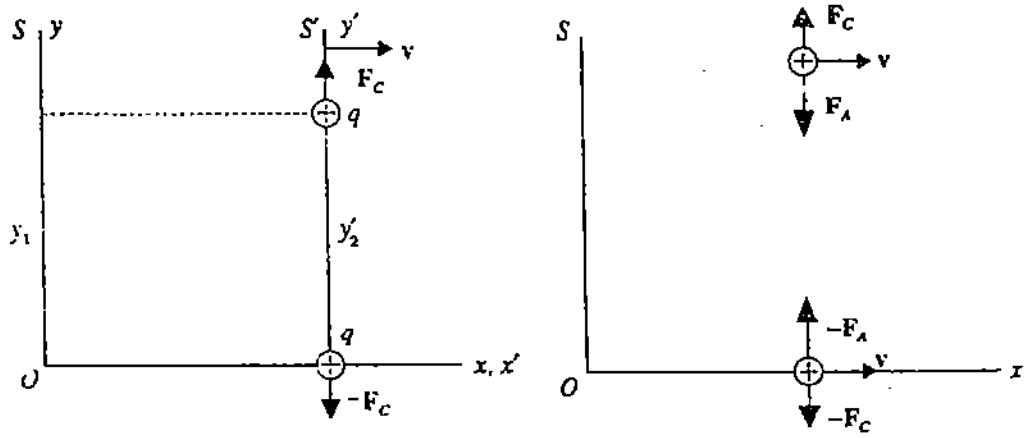


Fig. 1.2 : (a) Two equal positive point charges (carrying charge q) at rest on the y' axis of the frame of reference S'. In S' the charges repel each other with a force of magnitude F_C; (b) as seen in S the charges are moving to the right with a velocity v and attract each other with an additional force of magnitude F_A, giving a total force of magnitude $|F_C - F_A|$.

Maxwell's equations and the electromagnetic wave equation). It is a simple exercise and you could try it out yourself.

SAQ 3 *Spend 10 min*

Show that the electromagnetic wave equation

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} = 0$$

does not retain its form (i.e., it is not invariant) under the Galilean transformations (Eq. 1.1).

Hint: Use the chain rule in which if ϕ is a function of (x', y', z', t'), then for any function f

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial x'}\cdot\frac{\partial x'}{\partial x} + \frac{\partial f}{\partial y'}\cdot\frac{\partial y'}{\partial x} + \frac{\partial f}{\partial z'}\cdot\frac{\partial z'}{\partial x} + \frac{\partial f}{\partial t'}\cdot\frac{\partial t'}{\partial x}$$

Thus, there seems to be a fundamental disagreement between Maxwell's theory of electromagnetic fields, Newtonian mechanics and Galilean principle of relativity. Historically, this disagreement centred around the 'problem of light'. We too would like to focus on this problem. However, we shall confine ourselves only to one aspect of light, namely its propagation. You know that one of the consequences of Maxwell's equations is that light is an **electromagnetic** wave which propagates in all directions at the same speed $c = 3 \times 10^8 \text{m s}^{-1}$. Another consequence of the equations is that if the source of light is moving, the light emitted still travels at the same speed c. This brings up an interesting problem when Galilean relativity is applied to the propagation of light. Let us examine it in some detail.

1.3.2 Galilean Relativity and the Speed of Light

The wave nature of light was recognised even before Maxwell described its electromagnetic nature (for example, in the works of Young, Huygens and Fresnel). A search was on to find the medium in which light propagated. Sound waves require air to propagate and ocean waves travel on water. So what was the medium for propagation of light? Nineteenth century physicists believed that light propagated through a rarefied, all-pervasive (space filling) elastic medium called **luminiferous ether**. It was assumed to be so fine that planets and other heavenly bodies passed through it without appreciable friction. When Maxwell described the electromagnetic nature of light waves, the

associated electric and magnetic fields were visualised as stresses and strains in the ether. It was in this medium that light propagated with a speed $c = \frac{1}{\sqrt{\varepsilon_0 \mu_0}}$. When the Galilean principle of relativity was applied to the propagation of light in ether, it led to an inconsistency.

To understand the inconsistency, let us consider a frame S with respect to which light travels with velocity c. What is the velocity of light in a frame S' moving with a constant velocity u with respect to S? We can carry out the Galilean velocity transformation and obtain

$$\mathbf{v}' = \mathbf{c} - \mathbf{u}\ ,\quad |\mathbf{v}'| = \left(c^2 + u^2 - 2\mathbf{c}\ .\ \mathbf{u}\right)^{1/2} \tag{1.7}$$

where $|\mathbf{v}'|$ is the speed of light in S'. Clearly the speed of light in S' depends on the direction in which it is travelling. If c is in the direction of u, the speed of light in S' is $c - u$. In the direction opposite to u, the speed of light in S' is $c + u$. In any other direction it has a value between $c - u$ and $c + u$ as given by Eq. (1.7). We can also see that according to Galilean relativity principle, the speed of light would be different in different inertial frames of reference. In other words, Maxwell's equation would have to be of different forms in different inertial frames of reference, to give different speeds of light in those frames. So it appears that the Galilean principle of relativity is incompatible with the laws of electromagnetism, which give a constant speed of light.

Now suppose we accept both the Galilean transformation and the laws of electromagnetism (or Maxwell's equations) as basically correct. Then it follows that there is one unique privileged inertial frame of reference (the absolute frame) in which Maxwell's equations are valid. In this unique frame the speed of light would be $c = 1/\sqrt{\mu_0\ \varepsilon_0}$ whereas in other frames it would be different.

Let us now put all these developments in physics which led to the special theory of relativity in a perspective. The situation towards the end of nineteenth century seems to be as follows: The Galilean relativity principle does apply to Newton's laws of mechanics but not to Maxwell's laws of electromagnetism. This requires us to choose the correct consequences from among the following possible alternatives.

1. **Retain the Relativity Principle for mechanics but not for electrodynamics.** This would mean that Newton's mechanics remains unchanged. But the laws of electromagnetism hold only in one privileged frame of reference, that is the ether frame. If this alternative were correct we should be able to locate the ether frame experimentally.	2. **Retain the Relativity Principle for both mechanics and electrodynamics but hold the laws of electromagnetism as not correct.** If this alternative were correct, we should be able to do experiments that show deviations from the electromagnetic theory. Then we would need to reformulate the laws of electromagnetism so that the Galilean transformations apply to the new Laws.	3. **Retain the Relativity Principle for both mechanics and electrodynamics: but hold that the Newtonian mechanics is not correct.** If this alternative were correct then we should be able to do experiments which show deviations from Newtonian mechanics. Then we would need to reformulate Newtons's Laws. We would also have to give up the Galilean transformation because they do not give us an invariant form of Maxwell's equations. We shall need to look for some other transformation which is consistent with classical electromagnetism and the new laws of mechanics.

Several investigations were carried out to decide which of the three alternatives was correct. Their net outcome was to provide an experimental basis for rejecting the alternatives 1 and 2. The most famous of these experiments is the one performed by Michelson and Morley in 1887 to locate the absolute frame. You have studied this experiment in Unit 7, Block 2 of the course PHE-09 on optics. However, let us briefly study this celebrated historic experiment.

1.3.3 Attempts to Locate the Absolute Frame – The Michelson-Morley Experiment

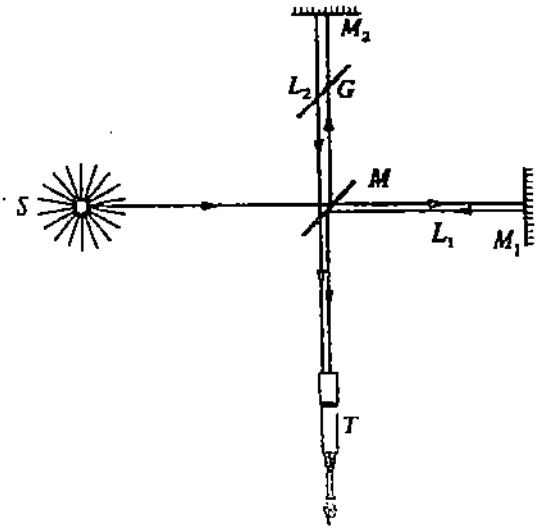
Let us first understand what was being investigated through this experiment. Consider a simple example. When we say that sound travels at 340 m s^{-1}, we are referring to the speed of sound with respect to air through which it propagates. If we move through still air towards an oncoming sound wave at a speed of 30 m s^{-1} (relative to the air), we observe the speed of sound to be 310 m s^{-1}. Clearly, the speed of sound relative to us varies with our speed relative to air.

Now the ether hypothesis suggests that the earth is moving in the ether medium as it orbits the sun. Therefore, in analogy to the example above, we can say that the speed of light relative to an observer on the earth varies with the earth's speed relative to the ether. The speed at which the earth orbits the sun is 30 km s^{-1}, about 0.01% (10^{-4}) of the speed of light. This is the maximum change which we can observe in the speed of light on earth as it moves through ether. Michelson, in 1881, and then in collaboration with Morley, in 1887, performed an experiment designed to detect such a change in the speed of light.

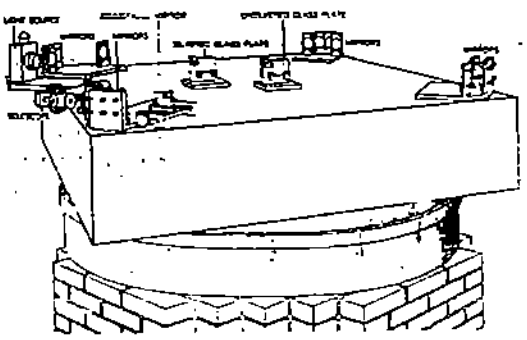
The essential principle of the experiment was to send a light-signal from a source to a mirror and back, noting the total time taken. The experiment was to be done twice:

(i) in the direction of earth's motion in ether, and

(ii) at right-angles to it.



(a)



(b)

Fig.1.3: (a) Schematic representation of Michelson- Morley experiment; (b) the apparatus.

In the experiment a beam of light from a source S (fixed with respect to the apparatus) is separated into two coherent beams by a partially silvered mirror M inclined at 45° to the beam direction (Fig. 1.3a). Two mirrors M_1 and M_2 are placed at nearly equal distances from M and at right angles to each other. These reflect the beams back to M. A part of each of the two beams reflected by M_1 and M_2, respectively, are reunited at M and the recombined beams are observed through a telescope T. A glass plate G is placed between M and M_2 to compensate for the extra distance travelled by light through M to M_1. Now if you have studied Block 2 of the course PHE-09, you would realise that when the two parts of the split beam recombine, they will interfere. Now suppose the time taken for light to travel from M to M_1 and back is t and the time required to travel from M to M_2 and back is t'. Then the interference will be constructive at a given point if at that point the difference

$$t - t' = n T, \qquad n = 1, 2, 3, \ldots \tag{1.8a}$$

and destructive if

$$t - t' = n + \frac{1}{2} T, \quad n = 1, 2, 3, \ldots \tag{1.8b}$$

where T is the time period of the light wave. So the existence of a time difference between the two paths travelled by light influences the illumination of a given point (say A) viewed through the telescope. It is either bright (constructive interference) or dark (destructive interference). If M_1 and M_2 are very nearly at right angles, the fringe pattern consists of nearly parallel lines.

Now, suppose we rotated the entire apparatus by 90° in the plane of M, M_1 and M_2. Then the orientation of MM_1 and MM_2 relative to the direction of Earth's motion through ether would change. Such a rotation would alter the time taken along each path and thus change the illumination of the point A. Or we could say that the fringe pattern would shift. It was this change in illumination of a given point (or a shift in the fringe pattern), as a result of rotation, that Michelson actually tried to detect. The expected shift was of the order of four-tenths of a fringe.

Michelson and Morley took utmost care in eliminating all possible sources of error, such as stresses and temperature effects. And this shift should have been clearly observable. Nevertheless,

no fringe shift was observed.

If you wish to go into the historical details you may read the first book listed as further reading.

One could say that at the time when the experiment was being done, the earth was at rest relative to the ether. However, the result did not change when it was repeated after a gap of six months. Indeed, this experiment was repeated many times by many workers over a 50-year period, in more sophisticated ways, at different times of the year. But the result was always the same. As far as Michelson was concerned, its implication was clear as he wrote at that time

'The result of the hypothesis of a stationary ether is shown to be incorrect.'

Needless to say, the ether hypothesis was not given up immediately. Several interpretations of the null result of this experiment to preserve the concept of ether were suggested. We will not go into the details of all these interpretations because, with time, (as evidence accumulated) it turned out that these were either inconsistent with observations and experiment or lacked sound conceptual bases.

Einstein was motivated by both the problem of the ether as a preferred reference frame and his thoughts on Maxwell's electrodynamics and, in particular, on Faraday's law of induction. Recall from Unit 13 of PHE-07 that Faraday's law of electromagnetic induction refers to the relative motion of a wire loop and a magnet. That motion together with the magnetic field of the magnet results in a current flowing in the loop. The detailed explanation of this effect as given at that time was not symmetrical. It was not the same when looked at from the reference frame of the magnet and from the reference frame of the loop. Einstein felt that this phenomenon should be exactly symmetrical since only relative motion is involved. He resolved all these problems by a bold leap, postulating the principle of relativity. He insisted that all laws be valid for all inertial observers. The second postulate stated what he believed nature was trying to tell us all along. The constancy of the speed of light was for him not something that needed to be explained. It was a new law of nature that had to be accepted.

Various experiments performed to measure the speed of light over the years have confirmed this result. Indeed, the speed of light in free space has been found to be constant at all times. It is independent of the place where measurements were carried out. It does not depend on its frequency, nature and motion of its source, direction of its propagation. It is also constant with respect to all inertial frames of reference. Thus, experiments help us to accept, indisputably, the following principle.

The speed of light in free space is a universal constant.

This result obviously contradicts the Galilean principle of relativity. At the same time the laws of electromagnetism are upheld by experiment. Moreover in certain other experiments performed in the early twentieth century departures from Newtonian mechanics were observed. In 1902, the motion of electrons (emitted by radioactive sources) in electric and magnetic fields was investigated experimentally. It was found that Newton's second law did not correctly describe the motion of these electrons which moved with velocities close to that of light. To sum up, we have found that the classical principle of relativity is incompatible with the laws of electromagnetism. Michelson-Morley experiment fails to detect ether (i.e., an absolute frame of reference). Thus, the ether hypothesis is untenable. It is experimentally established that the speed of light, in free space, is a constant.

What is more, experiments done on electrons moving at speeds close to that of light in electric and magnetic fields show a breakdown of Newton's laws of motion. Hence we can see that a relativity principle, applicable to both mechanics and to electromagnetism, is operating. Clearly it is not the Galilean principle, since that requires the speed of light to depend on the frame of reference in which it is measured. We conclude that the Galilean transformations should be replaced. Hence, the laws of mechanics, which are consistent with these transformations, need to be modified.

The discussion so far gives you an idea of the background in which Einstein's special theory of relativity emerged. Let us now study the special theory of relativity.

1.4 THE SPECIAL THEORY OF RELATIVITY

You have studied in the previous section that the constancy of the speed of light in all inertial frames stands in contradiction with the Galilean transformations. In 1905, Albert Einstein (1879-1955) presented a revolutionary proposal which resolved this contradiction. Rather than modifying electromagnetic theory, he rejected the ether hypothesis and generalised Galilean principle of relativity. In his paper "On the Electrodynamics of Moving Bodies", Einstein formulated the two postulates of the special theory of relativity, which we rephrase here:

Postulates of the Special Theory of Relativity

Postulate 1 — The Principle of Relativity

The laws of physics are the same in all inertial frames of reference.

Postulate 2 — The Principle of Constancy of Speed of Light

The speed of light (in vacuum) has the same constant value in all inertial reference frames.

These two assumptions led Einstein to a new theory of physics which is now known as **the special theory of relativity.** It is special because it only deals with observations made in inertial frames. For example, it does not say anything about the relationship between two frames undergoing relative acceleration. Non-inertial frames are the subject matter of another of Einstein's theories — the general theory of relativity.

Let us now understand the meaning of these postulates.

1.4.1 The Principle of Relativity

You have briefly studied the Galilean principle of relativity which applied to the Newtonian laws of mechanics. This limited principle has now been generalised to all laws of physics — **any law of physics that is true in one inertial frame will also be true in all other inertial frames.** Let us consider an example to understand this statement.

Suppose a positive electric charge q was fixed at a point $(X, 0, 0)$ in a stationary inertial frame S (Fig.1.4). If another positive charge q' was released at some point on the x-axis, it would accelerate away from the fixed charge at $x = X$. We could experimentally determine the x component of the acceleration of the moving charge as a function of its distance from the fixed charge. The relationship would be of the following form:

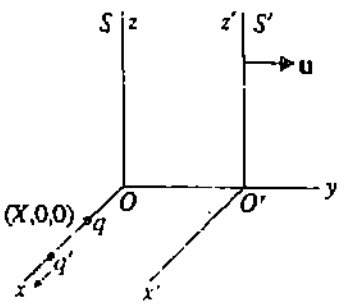


Fig. 1.4: A positive charge q is fixed at $x = X$, $y = 0$, $z = 0$ in frame S. Another positively charged particle q' released from a point on the x-axis experiences an acceleration away from the charge.

For a charged particle moving away from a fixed charge

$$\frac{d^2x}{dt^2} = \frac{k}{(x - X)^2} \tag{1.9a}$$

where k is a constant.

Now suppose that another observer is stationed in an inertial frame of reference S' moving with respect to S with a constant velocity $\mathbf{u}$. The principle of relativity tells us that if Eq. (1.9a) is really a law of physics, then the observer in the S' frame should find that

$$\frac{d^2x'}{dt'^2} = \frac{k'}{(x' - X')^2} \tag{1.9b}$$

where X' is the location of the fixed charge q on the x'-axis of the S' frame. Thus, even though the values of x', t', X' and the constant k' are different from x, t, X, and k, the relationship between them is of the same form as Eq. (1.9a). Conversely, we can also say that any equation that cannot be written in the same form in all inertial frames cannot be a law of physics. So the principle of relativity also allows us to determine which relationships (or equations) can or cannot be laws of physics.

In summary, the principle of relativity implies that the laws of nature do not depend upon the choice of an inertial frame of reference or the position or motion of an observer – they

Incidentally, using this principle Einstein discovered that many of the equations which were thought to be 'laws of physics' in his day could not be laws at all, even though they agreed with all existing experiments. He proposed modifications or alternatives to many of these 'laws' – alternatives that could be written in the same form in all inertial frames, in accordance with the principle of relativity. Experiments done later have shown that the old 'laws' though adequate in their own time, do not describe the results that are now available.

will always retain their form in any inertial frame of reference. Indeed, the measurements of various quantities, like the positions, times, velocities, energies, momenta, electric and magnetic fields may be different in different inertial frames. However, the relationships between these quantities governed by various laws would remain the same in all inertial frames.

In philosophical terms, we can say that the principle of relativity underscores the objective character of the laws of nature and not the relativity of knowledge.

The principle of relativity is also stated in the following form which you will often encounter: The laws of physics do not allow us to distinguish between different inertial frames.

In other words, **you will not be able to distinguish through any experiment whether you are at rest or in a state of uniform motion.** For, if there were such an experiment, it would mean that the laws of physics depended in some way on your velocity and were different from the laws of physics when you were at rest.

You should understand that the principle of relativity does not claim that all inertial frames are the same in all respects. To appreciate this point, consider two different spacecrafts, each travelling with a different constant velocity with respect to S. The principle of relativity tells us that the two frames cannot be distinguished as far as the laws of physics are concerned. However, if one could look outside each spacecraft through a window, it would be easy to know that they are moving at different velocities, relative to S. Does this contradict the principle of relativity? No, because the velocity of the spacecraft relative to S is not determined by a law of physics. Besides, in formulating this form of the principle of relativity the essential condition was that the spacecrafts were completely isolated.

The ideas presented here are worth careful thought and you may need to go through them more than once. You may now like to attempt the following SAQ to know whether you have grasped the principle of relativity or not.

SAQ 4

(a) Suppose the observer in S' frame finds that Eq.(1.9b) is supported by experiment. Does it automatically follow that Eq. (1.9b) is a law of physics?

(b) Suppose you observe the motion of a particle in an inertial frame. You find that the x-component of the acceleration of the particle is described by the following equation:

$$\frac{d^2x}{dt^2} = -k_1 \frac{dx}{dt} + k_2 [(x - X)^2 + (y - Y)^2 + (z - Z)^2]$$

where k_1 and k_2 are constants and (X, Y, Z) are the coordinates of a second particle in your frame. If this equation is to be regarded as a possible law in physics, what relationship must an observer in a different inertial frame find to be experimentally valid?

Let us now study the second postulate of the special theory of relativity.

We have chosen to highlight this aspect because of its historical significance. When asked how long he had worked on the Special Theory, Einstein replied that he had started at age 16 and worked for ten years. He had to abandon many fruitless attempts 'until at last it came to him that time was suspect' — in particular, the assumption that there exists a universal time which is the same. In this section, we shall briefly study this aspect.

1.4.2 The Principle of Constancy of Speed of Light

The second postulate about the constancy of speed of light is crucial in leading to the theory of relativity. It is very important because it radically alters the classical notions of absolute space and time. Here we shall briefly discuss the implications of the second postulate of special relativity theory, particularly regarding the notion of time.

The Nature of Time in Special Relativity

The basic premise of Newtonian mechanics was that the same time scale applied to all inertial frames of reference (recall the equation $t' = t$ in Galilean transformation). Using this universal time scale, we must be able to give meaning to statements such as "Events A and B occurred at the same time", without referring to any inertial frame of reference. To use the example given by Einstein, when we say that a train arrives at 7 o'clock, what we mean is this: The pointing of the clock hand to 7 and the arrival of the train are simultaneous events.

Thus, the assigning of time to events involves judging whether they are simultaneous or not. So, if all observers, independent of their position and velocity, agreed that any two events (e.g., the arrival of the train at the station and the pointing of the clock hand to 7) are simultaneous, we could certainly say that the absolute Newtonian time scale existed.

We shall certainly not have an absolute time scale if different inertial observers disagree about two events being simultaneous, i.e., one inertial observer says that two events occur at the same time and another inertial observer says that they do not. This is precisely what happens if we uphold the constancy of the speed of light. Let us understand this idea with the help of an imaginary experiment.

Consider a train compartment travelling at a very high constant velocity V to the right of an observer S at rest on the earth (Fig. 1.5). A highspeed flashbulb is situated at the exact centre of the compartment. It sends out light pulses to the right and left when it flashes. There are photocells at each end of the compartment, so that an observer S', in the compartment can detect when the light pulses strike its ends. Now, suppose by some ingenious device, the observer S on the earth is also able to measure the progress of the two pulses. Let the positions of S and S' coincide with that of the bulb when it flashes (Fig. 1.5a).

The flashbulb is at rest relative to the observer S' in the compartment. Since it is at the centre, when the bulb flashes, two light pulses travel equal distances to the two ends of the compartment in equal times. Hence, S' observes that the light pulses hit the two ends of the compartment at the same time.

Is the same conclusion drawn by S, who is stationary on the earth? Refer to Figs. 1.5b and 1.5c. The light pulses travel equal distances to the right and left in equal time. But in the frame of S the compartment is moving to the right. So in the frame of S, the distance

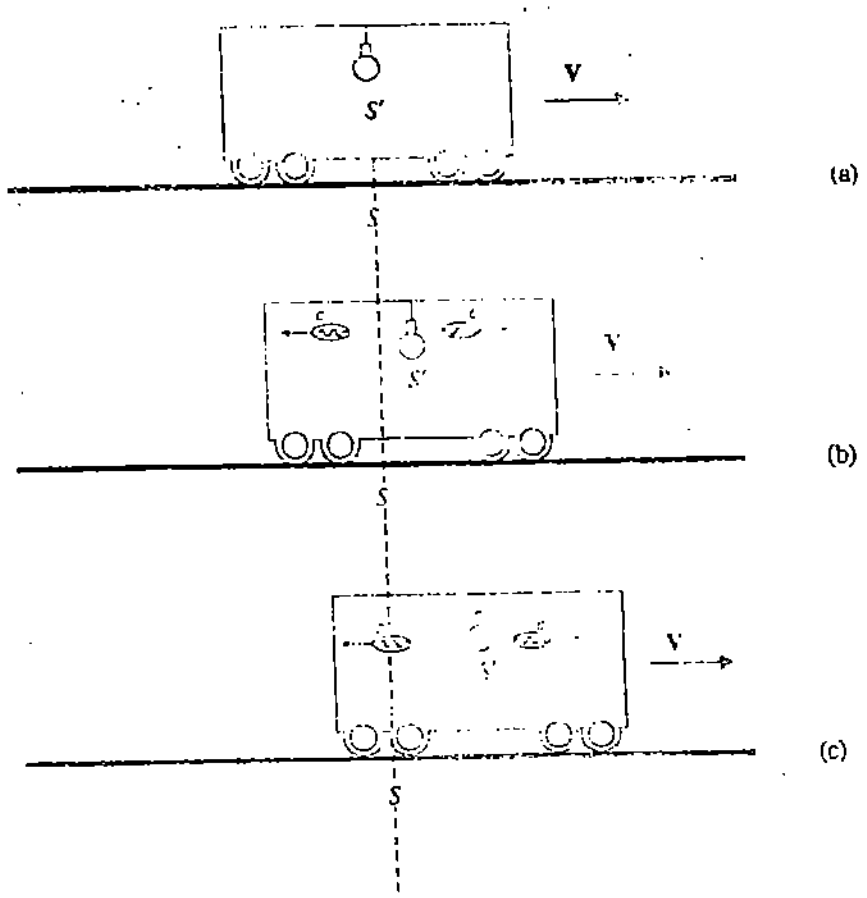


Fig.1.5: Unlike the inertial observer S' in the moving train compartment, the stationary observer S on the earth observes that the light pulses do not strike the ends of the car simultaneously. The figures are drawn with respect to the inertial observer S.

between the point at which S observes the bulb flashing and the left end of the compartment is shorter compared to its distance from the right end. As a result, S measures that the light pulse moving to the left strikes the end of the compartment before the other pulse strikes the opposite end. In the frame of reference of S, the light pulses do not hit the two ends of the compartment at the same time.

Now, if Newtonian mechanics were valid, this difference in the distance travelled in the frame of S would be compensated by the different speeds of light measured by S. The observer S would assign a lower speed $(c - V)$ to the pulse travelling to the left (opposite to the direction of the train's motion); the pulse travelling to the right would travel a greater distance but at a greater speed of $(c + V)$. Thus, S would measure both the times to be equal and conclude that the two pulses hit the ends of the compartment at the same time.

But the speed of light is constant. Therefore, in the frame of S, the two events (the light pulses hitting the two ends of the compartment) do not occur simultaneously.

This signifies a major break with the older ideas of absolute time, because different observers do not agree on what is the same time. Of course, remember that this result is arrived at for events occurring at different locations (the two ends of the compartment for instance). In the next unit we shall come back to this discussion and also consider events occurring simultaneously at the same point in space.

To sum up, we can conclude that the notion of absolute time is contradicted by the second postulate because

> **Events (occurring at different points in space) that are simultaneous in one inertial system may not be simultaneous in another.**

This conclusion is termed the **relativity of simultaneity**. It is the fundamental difference between Newtonian relativity and special relativity. In Newtonian relativity, observers in S and S' always agree everywhere about events occurring at the same time. This is also the origin of other features of space and time that follow from the special theory of relativity, namely, the phenomena of length contraction, time dilation, twin paradox, etc. In this brief discussion we have given you a flavour of what follows in the next unit. We end this section with an SAQ for you.

Spend 10 min

SAQ 5

a) The speed of light [illegible] a constant speed [illegible] frame at rest). [illegible]

b) One observer [illegible] speed. Considering [illegible] what do the observers measure to be the same and what [illegible]

c) Suppose the compartment of the train in Fig. 1.5 would shrink so [illegible] between its two ends approaches zero. Can you give a simple argument [illegible] events not separated in space are simultaneous for all inertial observers.

Let us now summarise what you have studied in this unit.

1.5 SUMMARY

- An event is an occurrence that happens at a point in space and at an instant in time. The Galilean coordinate transformations from an inertial frame of reference S to another inertial frame S' moving at a velocity $\mathbf{u} = u\hat{\mathbf{i}}$ relative to S are given by

$$\begin{aligned} x' &= x - ut \\ y' &= y \\ z' &= z \\ t' &= t \end{aligned}$$

Here (x, y, z) are the coordinates of an event and t is the time at which it occurs, as

measured in the frame S. The coordinates (x', y', z') and the time t' are measured in S'. S' moves with respect to S so that the x-x' axes are common and the y-y', z-z' axes are parallel.

- The **Galilean or classical principle of relativity** states that the laws of mechanics can be written in the same form in all inertial frames of reference. If they hold in one inertial frame, they also hold in all other inertial frames.
- **Galilean coordinate transformations** predict that the velocity of light should be different in different inertial frames and do not preserve the form of Maxwell's equations. Thus, the Galilean principle of relativity does not apply to the laws of electromagnetism.
- Experiments, especially the **Michelson-Morley experiment,** indicate that the speed of light is a universal constant and is independent of the relative uniform motion of the observer, the transmitting medium, and the source. The laws of electromagnetism are also upheld by experiments. Newtonian mechanics is experimentally observed to break down for particles moving at speeds close to that of light.
- In his **special theory of relativity** Einstein affirms the classical principle of relativity and generalises it to include all laws of physics. This also means that the speed of light should be the same in all uniformly moving systems.
- The **postulates of special relativity** are as follows:

  Postulate 1 — The Principle of Relativity

  The laws of physics are the same in all inertial frames of reference.

  Postulate 2 — The Principle of Constancy of Speed of Light

  The speed of light (in vacuum) has the same constant value in all inertial frames of reference.

1.6 TERMINAL QUESTIONS

Spend 30 min

1. Linear momentum and kinetic energy are conserved in an elastic collision. Use the Galilean transformation equations to show that if a collision is elastic in one inertial frame it is elastic in all inertial frames.
2. a) Does the Michelson-Morley experiment indicate that ether is an unnecessary concept, or does it prove that there is no such thing?

   b) One of Maxwell's equations is $\oint_C \mathbf{E}.d\mathbf{l} = -\frac{\partial \Phi_B}{\partial t}$ in an inertial frame S.

   According to the principle of relativity, what will its form be in another inertial frame S'?
3. How is the nature of time as enunciated in classical relativity different from that in the special theory of relativity?

1.7 SOLUTIONS AND ANSWERS

SAQs (Self-Assessment Questions)

1. a) Non-inertial since the car is accelerating
   b) Inertial
   c) Non-inertial since the electron is accelerating
   d) Inertial
   e) Inertial
2. a) No. Different inertial observers can measure different values of physical quantities but the relationship between them remains the same.

   b) Yes, since these laws follow from Newtonian mechanics.

3. Using the chain rule we get

$$\frac{\partial \phi}{\partial x} = \frac{\partial \phi}{\partial x'}\frac{\partial x'}{\partial x} + \frac{\partial \phi}{\partial y'}\frac{\partial y'}{\partial x} + \frac{\partial \phi}{\partial z'}\frac{\partial z'}{\partial x} + \frac{\partial \phi}{\partial t'}\frac{\partial t'}{\partial x}$$

From Galilean transformations given by Eq. (1.1)

we have $\frac{\partial x'}{\partial x} = 1, \frac{\partial y'}{\partial x} = 0, \frac{\partial z'}{\partial x} = 0, \frac{\partial t'}{\partial x} = 0.$

Thus, $\frac{\partial \phi}{\partial x} = \frac{\partial \phi}{\partial x'}$ and $\frac{\partial^2 \phi}{\partial x^2} = \frac{\partial^2 \phi}{\partial x'^2}$.

You can also show that, $\frac{\partial^2 \phi}{\partial y^2} = \frac{\partial^2 \phi}{\partial y'^2}$ and $\frac{\partial^2 \phi}{\partial z^2} = \frac{\partial^2 \phi}{\partial z'^2}$.

Now $\frac{\partial \phi}{\partial t} = \frac{\partial \phi}{\partial x'}\frac{\partial x'}{\partial t} + \frac{\partial \phi}{\partial y'}\frac{\partial y'}{\partial t} + \frac{\partial \phi}{\partial z'}\frac{\partial z'}{\partial t} + \frac{\partial \phi}{\partial t'}\frac{\partial t'}{\partial t}$.

From Eqs. (1.1) $\frac{\partial x'}{\partial t} = -u, \frac{\partial y'}{\partial t} = 0, \frac{\partial z'}{\partial t} = 0, \frac{\partial t'}{\partial t} = 1.$

Thus $\frac{\partial \phi}{\partial t} = -u\frac{\partial \phi}{\partial x'} + \frac{\partial \phi}{\partial t'}$

and $\frac{\partial^2 \phi}{\partial t^2} = u^2\frac{\partial^2 \phi}{\partial x'^2} - 2u\frac{\partial^2 \phi}{\partial x'\,\partial t'} + \frac{\partial^2 \phi}{\partial t'^2}$

Thus there are two extra terms in the expression of $\partial^2\phi/\partial t^2$ in the S' frame. Hence the wave equation does not retain its form in the S' frame.

4. a) No, the Eq. (1.9b) has to satisfy the principle of relativity as well, i.e., it has to hold in all inertial frames.

   b) The relationship in another inertial frame S' should be of the form

$$\frac{\partial^2 x'}{\partial t'^2} = k'_1\frac{\partial x'}{\partial t'} + k'_2\,[(x'-X')^2 + (y'-Y')^2 + (z'-Z')^2]$$

   where k'_1 and k'_2 are constants, (X', Y', Z') are the coordinates of the second particle in S' frame, and (x', y', z', t') are the space-time coordinates of the particle in S' frame.

5. a) It is 3.0×10^8 m s^{-1} since it is a universal constant.

   b) The speed of light is measured to be the same, its frequency and wavelength are measured to be different by the two observers.

   c) The loss of simultaneity in S is due to the finite length of the compartment, which makes the observer S measure the distance to the left to be smaller. If the compartment's length shrunk to zero, the two pulses would strike its ends simultaneously for all inertial frames. Thus, two events occurring at the same position would occur at the same time for all inertial observers.

Terminal Questions

1. Let an object of mass m and velocity v_1 collide elastically with an object of mass M and velocity V_1 in frame S. Let their velocities after collision be v_2 and V_2, respectively in S. Then since linear momentum and kinetic energies are conserved in an elastic collision, we have

$$mv_1 + MV_1 = mv_2 + MV_2 \qquad (A)$$

and $$\frac{1}{2}mv_1^2 + \frac{1}{2}MV_1^2 = \frac{1}{2}mv_2^2 + \frac{1}{2}MV_2^2 \qquad (B)$$

Now let S' frame move with a velocity v with respect to S. Then the Galilean velocity transformation gives us the velocities of m and M before and after collision as

$$v_1' = v_1 - v,\ V_1' = V_1 - v,\ v_2' = v_2 - v,\ V_2' = V_2 - v, \qquad (C)$$

Substituting for v_1, V_1, v_2, V_2 from (C) in (A) and (B) we get

$$m(v_1' + v) + M(V_1' + v) = m(v_2' + v) + M(V_2' + v)$$

or $$mv_1' + MV_1' = mv_2' + MV_2' \qquad \text{(D)}$$

Thus linear momentum is conserved in the S' frame. For the conservation of kinetic energy, we proceed as follows:

$$m\,|v_1' + v|^2 + M\,|V_1' + v|^2 = m|v_2' + v|^2 + M\,|V_2' + v|^2$$

or $$m(v_1'^2 + v^2 + 2v_1'.v) + M(V_1'^2 + v^2 + 2V_1'.v) = m(v_2'^2 + v^2 + 2v_2'.v) + M(V_2'^2 + v^2 + 2V_2'.v)$$

or $$mv_1'^2 + 2mv_1' \cdot v + MV_1'^2 + 2MV_1' \cdot v = mv_2'^2 + 2mv_2' \cdot v + MV_2'^2 + 2MV_2' \cdot v$$

or $$mv_1'^2 + MV_1'^2 + 2(mv_1' + MV_1') \cdot v = mv_2'^2 + MV_2'^2 + 2(mv_2' + MV_2') \cdot v$$

Using (D) we get

$$mv_1'^2 + MV_1'^2 = mv_2'^2 + MV_2'^2,$$

or $$\frac{1}{2}(mv_1'^2 + MV_1'^2) = \frac{1}{2}(mv_2'^2 + MV_2'^2)$$

Thus, K.E. is conserved in S'.

2. a) The Michelson-Morley experiment indicates only that the concept of ether is unnecessary.

   b) The form in frame S' will be

$$\oint_C \mathbf{E}' \cdot d\mathbf{l}' = -\frac{\partial \Phi_B'}{\partial t'}$$

3. In classical relativity observers in different inertial frames will always agree about the time at which an event occurs. If two events occur simultaneously for one inertial observer, then according to classical relativity, these two events will be simultaneous for all other inertial observers. This need not be so according to special theory of relativity. Thus, in special relativity two inertial observers need not measure the time at which an event occurs to be the same, if the events occur at different positions in space.

UNIT 2 RELATIVISTIC KINEMATICS

Structure

2.1. Introduction
Objectives

2.2 Lorentz Transformation

2.3 Implications of Special Relativity
Relativity of Simultaneity
Length Contraction
Time Dilation

2.4 Relativistic Transformation of Velocity

2.5 Relativistic Doppler Effect

2.6 Summary

2.7 Terminal Questions

2.8 Solutions and Answers

2.1 INTRODUCTION

In Unit 1 you have studied briefly about the developments in physics which led to the emergence of special relativity. You have seen that the Galilean coordinate transformation stands in contradiction to the laws of electromagnetism. In particular, the speed of light turns out to be different in different inertial frames of reference according to Galilean velocity transformation. However, experiments show that the speed of light in vacuum is a universal constant. You have studied postulates of special theory of relativity and learnt about the need for a new transformation.

"I sometimes ask myself why I was the one to develop the theory of relativity. The reason, I think, is that a normal adult never stops to think about problems of space and time. These are things... thought of as a child. But I began to wonder about space and time only when I had grown up. Naturally, I could go deeper into the problem than a child."

—Albert Einstein

In this unit we shall use the postulates of special relativity to deduce the new coordinate transformation, called the **Lorentz transformation** (Sec. 2.2) and the implications of these postulates for particle kinematics. In Sec. 1.4.2 (Unit 1) we have briefly discussed how the special theory of relativity alters our notion of time based on classical ideas. In Sec. 2.3 we shall examine in somewhat greater detail how the established ideas of absolute space and absolute time are revised radically in a world view based on special theory of relativity. In particular, we shall revisit the concept of relativity of simultaneity. You will study how this concept raises serious doubts about the Newtonian assumptions that the measurements of distance between two points and the time intervals are the same for all observers. This will become clear when you study the phenomena of length contraction and time dilation. In Sec. 2.4 you will learn how the velocity of an object transforms in going from one inertial frame to another under the Lorentz coordinate transformation. Finally, in Sec. 2.5, we have discussed the relativistic Doppler effect as an application of special relativity to Optics.

In the next unit we shall turn our attention to relativistic dynamics. You will learn the modifications carried out in Newtonian mechanics to make it compatible with the special theory of relativity and Lorentz transformations.

Objectives

After studying this unit you should be able to

- use the Lorentz transformation equations
- explain the phenomena of length contraction and time dilation
- transform the velocity of an object from one inertial frame to another
- perform relativistic velocity addition
- compute the relativistic Doppler shift, and

- solve numerical problems based on special theory of relativity.

Study Guide

The concepts presented in this unit are entirely new for you. In working through the unit you will find that Sec. 2.3 is the longest and the most intellectually demanding. You should study this section very carefully. You may need to devote more than half the study time for this unit to Sec. 2.3. You can go through Sec. 2.2 quickly and concentrate more on Sec. 2.3. Secs. 2.4 and 2.5 should not be too difficult to understand. Yet do not rush through them. In our estimate it may take you about 9h to 10h to complete this unit.

2.2 LORENTZ TRANSFORMATION

Recall from Unit 1, Sec. 1.4.2 that we have to find a new transformation of space and time coordinates of a given event in the inertial frames S and S' moving with a velocity $\mathbf{v} \equiv (v\,\hat{\mathbf{i}})$ relative to one another. This new transformation should be consistent with the postulates of special relativity.

We begin as usual by considering two inertial frames of reference, moving with velocity v relative to one another. We take S to be the rest (laboratory) frame of reference and S' to be the moving frame which has velocity v in the positive x- direction of S. Let both frames be rectangular with their axes parallel. Finally, we define the origin of time $t = 0$ in S and $t'=0$ in S' when the two origins of S and S' coincide.

Before we determine this new coordinate transformation we should be clear about the method of assigning coordinates to an event in an inertial frame S. For this purpose, we assume that every observer is equipped with a standard clock and a standard of length, e.g., a metre stick. The observer can then assign Cartesian right-handed rectangular space coordinates (x, y, z) to any event in S. Knowing the distance of the event and noting the time at which the observer receives the light signal from it, a time coordinate t can be assigned to the event. Such coordinates (x, y, z, t) are called *standard coordinates*. Let us now find the new coordinate transformation.

Now consider a light wave spreading out from a point source stationed at the origin of the frames at $t = t' = 0$. The wavefront (i.e., the surface of equal phase) will be a sphere if observed in the reference frame S in which the source is at rest. But according to the postulates of special relativity the wavefront must also be a sphere when observed in the frame S'. For, if the shape of the wavefront changes in S' then we can know that the source is moving. But this would violate the first postulate of special relativity theory which tells us that if S and S' are isolated then no experiment can help us know which one of these frames (S or S') is moving. Therefore, we should not be able to tell from the shape of the wavefront whether the source is at rest or in uniform motion. In other words, the shape of the wavefront as observed from S and S' has to be the same. The equation of the spherical wave front in S emitted at the origin at $t = 0$ is

$$x^2 + y^2 + z^2 = c^2 t^2 \tag{2.1a}$$

Therefore, the equation of the wavefront in S' must be

$$x'^2 + y'^2 + z'^2 = c^2 t'^2 \tag{2.1b}$$

where the speed of light c is the same in both S and S', according to the second postulate.

Does the Galilean transformation satisfy Eqs. (2.1a and b)? You could check it out quickly.

You have found that the Galilean transformation fails to satisfy Eqs. (2.1a) and (2.1b). So

we have to look for another coordinate transformation consistent with the special theory of relativity. We will impose a few conditions to make our task easier. Firstly, the new coordinate transformation should be consistent with the postulates of special relativity, i.e., we must have a transformation which satisfies Eqs. (2.1a and b) simultaneously.

Secondly, we assume that space and time are homogeneous, i.e., all points in space and time are equivalent. To understand its implication, let us suppose that we measure a length or time interval of a specific event in a reference frame. The results of the measurement should not depend on where or when the event occurs–they should be the same at whichever point in space or time that event occurs. (This assumption simplifies our task considerably because this leads to a linear transformation.) For example, suppose x' depends on the square of x, i.e., $x' = ax^2$. Then the length of a rod would be related as follows in the frames S and S':

$$x'_2 - x'_1 = a(x_2 - x_1)^2$$

If $x_1 = 1$ and $x_2 = 2$, then $x'_2 - x'_1 = 3a$. However, if $x_1 = 4$ and $x_2 = 5$, then $x'_2 - x'_1 = 9a$. Thus, for a quadratic or higher order transformation the measured length of rod in S' would depend on where it was situated in S. This is inconsistent with the homogeneity of space that we have assumed. We assume the following relations

You may wonder why we have chosen this particular form of transformation equations. For a detailed explanation, we refer you to the reference 1 listed in Further Reading at the end of this block.

$$x' = a_1 x + a_2 t \tag{2.2a}$$

$$y' = y \tag{2.2b}$$

$$z' = z \tag{2.2c}$$

$$t' = b_1 x + b_2 t \tag{2.2d}$$

Now, consider a point for which $x' = 0$. In the S frame, it is moving along the positive x-axis with speed v. So its coordinate in S is $x = vt$. Thus,

$$\text{for } x' = 0\,, \quad \frac{dx}{dt} = v \tag{2.3a}$$

Similarly, a point for which $x' = 0$, seen from S' moves along the negative x' axis with speed v. So its coordinate in S' is $x' = -vt'$. Thus

$$\text{for } x = 0\,, \quad \frac{dx'}{dt'} = -v \tag{2.3b}$$

For $x' = 0$, Eq. (2.2a) yields

$$a_1 x + a_2 t = 0, \quad \text{where } \frac{dx}{dt} = -\frac{a_2}{a_1} = v \tag{2.4a}$$

For $x = 0$, Eqs. (2.2a) and (2.2d) reduce to

$$x' = a_2 t \quad \text{and} \quad t' = b_2 t$$

or

$$x' = \frac{a_2}{b_2} t',$$

so that

$$\frac{dx'}{dt'} = \frac{a_2}{b_2} = -v \tag{2.4b}$$

From Eqs. (2.4a) and (2.4b) we get

$$\frac{a_2}{b_2} = \frac{a_1}{a_2}$$

or

$$a_1 = b_2 \tag{2.5}$$

Let us now substitute the coordinate transformation given by Eqs. (2.2) and (2.5) in Eq. (2.1b). The result is

$$(a_1 x + a_2 t)^2 + y^2 + z^2 = c^2 (b_1 x + a_1 t)^2$$

or $\quad a_1^2 x^2 + 2a_1a_2xt + a_2^2t^2 + y^2 + z^2 = c^2 (b_1^2x^2 + a_1^2t^2 + 2a_1b_1x)$

This result should be consistent with Eq. (2.1a). Therefore, on comparing these we obtain

coefficient of xt is zero $\quad \Rightarrow \quad 2a_1a_2 = 2c^2a_1b_1$

coefficient of x^2 is 1 $\quad \Rightarrow \quad a_1^2 - c^2b_1^2 = 1$

coefficient of t^2 is $-c^2$ $\quad \Rightarrow \quad a_2^2 - c^2a_1^2 = -c^2$

Now the task of arriving at the transformation equations is a matter of simple algebra. Why don't you give it a try?

SAQ 2

Obtain the transformation equations in terms of x, t and v and verify Eqs. (2.2a) to (2.2d).

On solving SAQ 2, you have arrived at the

Lorentz Transformation

$$x' = \frac{x - vt}{(1 - v^2/c^2)^{1/2}} \qquad (2.6a)$$

$$y' = y \qquad (2.6b)$$

$$z' = z \qquad (2.6c)$$

$$t' = \frac{t - (v/c^2)x}{(1 - v^2/c^2)^{1/2}} \qquad (2.6d)$$

The Lorentz transformation is linear in x and t. It has another interesting property. You may like to discover this property. For this you should simplify Eqs. (2.6a to d) when $v/c << 1$. Work it out in the SAQ given below.

It is customary to cast the Lorentz transformation in a more compact form by introducing two factors $\beta = v/c$ and $\gamma = \frac{1}{(1 - \beta^2)^{1/2}} = \frac{1}{(1 - v^2/c^2)^{1/2}}$. Note that

(a) $\gamma \geq 1$ since v is finite

(b) $\gamma \to 1$ as $v \to 0$

(c) $\gamma \to \infty$ as $v \to c$

In the compact notation the Lorentz transformation Eqs. (2.6) take the form

Lorentz Transformation

$$x' = \gamma(x - vt) = \gamma(x - \beta ct) \quad (2.7a)$$

$$y' = y \quad (2.7b)$$

$$z' = z \quad (2.7c)$$

$$t' = \gamma(t - vx/c^2) = \gamma(t - \beta x/c) \quad (2.7d)$$

You should commit these equations to memory as they will be used very often in our discussion on special relativity. You can immediately see from Eqs. (2.6) as well as (2.7) that v can never exceed c. For $v > c$, the space and time coordinates become imaginary, which is a physical impossibility. Thus, we arrive at the following conclusion: **We cannot measure speeds greater than the speed of light;** c **is the limiting speed in the physical universe.** We shall return to this point in Sec. 2.4 of this unit.

The inverse transformation is obtained by interchanging

$$x \leftrightarrow x'$$

$$y \leftrightarrow y'$$

$$z \leftrightarrow z'$$

$$t \leftrightarrow t'$$

and $$v \leftrightarrow -v$$

In SAQ 4 you will also prove the **Inverse Lorentz Transformation.**

Inverse Lorentz Transformation

$$x = \gamma(x' + \beta ct') \quad (2.8a)$$

$$y = y' \quad (2.8b)$$

$$z = z' \quad (2.8c)$$

$$t = \gamma(t' + \beta x'/c) \quad (2.8d)$$

SAQ 4

Verify Eqs. (2.8) algebraically.

Thus, we have arrived at the Lorentz transformation and its inverse transformation which is consistent with the postulates of special relativity.

Let us now study some interesting implications of postulates of special relativity and Lorentz transformation.

2.3 IMPLICATIONS OF SPECIAL RELATIVITY

You have studied in Sec. 1.4.2 of Unit 1 how radically the nature of time in the Einsteinian world view differs from that in the Newtonian world view. In particular, according to the special theory of relativity two events assumed to occur simultaneously in one inertial frame are not necessarily simultaneous in another inertial frame moving with respect to the first. We used a thought experiment to illustrate the relativity of simultaneity. In this section we shall make use of Lorentz transformation to outline this feature of the special theory of relativity. You will also learn how the breakdown of simultaneity leads to other relativistic phenomena like the relativity of length and time interval measurements. In relativistic parlance these are termed as length contraction and time dilation. Let us begin by re-examining the notion of simultaneity and its breakdown in special theory of relativity.

2.3.1 Relativity of Simultaneity

Consider two simultaneous events occurring at two different points in reference frame

S. Let us assign the coordinates (x_1, y_1, z_1, t_1) and (x_2, y_2, z_2, t_2) to the events. In this case $x_1 \neq x_2$ and $t_1 = t_2$. Using Lorentz transformation we now show that, in general, those events which are simultaneous in S are not simultaneous in another inertial frame S' moving uniformly relative to S at a speed v. Let the coordinates of the two events in S' be (x'_1, y'_1, z'_1, t'_1) and (x'_2, y'_2, z'_2, t'_2). From the inverse Lorentz transformation (Eqs. 2.8) we can write

$$t_1 = \gamma (t'_1 + \beta x'_1/c), \qquad t_2 = \gamma (t'_2 + \beta x'_2/c) \tag{2.9}$$

Since these events are simultaneous in S, $t_1 = t_2$. Hence Eq. (2.9) yields the result

$$(t'_1 + \beta x'_1/c) = (t'_2 + \beta x'_2/c)$$

or

$$t'_1 = t'_2 + \frac{\beta}{c} (x'_2 - x'_1) \tag{2.10a}$$

We can obtain $(x'_2 - x'_1)$ in terms of x_2 and x_1 from Eqs. (2.7). Hence

$$x'_1 = \gamma (x_1 - \beta c t_1) \qquad \text{and} \qquad x'_2 = \gamma (x_2 - \beta c t_2)$$

Therefore,

$$x'_2 - x'_1 = \gamma (x_2 - x_1) - \gamma \beta c (t_2 - t_1)$$

$$= \gamma (x_2 - x_1) \qquad (\because t_2 = t_1)$$

Thus

$$t'_1 = t'_2 + \frac{\beta}{c} \gamma (x_2 - x_1) \tag{2.10b}$$

Since $x_2 \neq x_1$, Eq. (2.10b) yields that $t'_1 \neq t'_2$. Hence simultaneous events occurring at different positions in one frame are not simultaneous in another frame in uniform relative motion.

However, for $x_2 = x_1$ and $t_1 = t_2$, Eq. (2.10b) tells us that $t'_1 = t'_2$. Thus, if two events occur simultaneously at the same position in one inertial frame, then they are simultaneous and occur at the same position in every other inertial frame of reference (since $t'_1 = t'_2$ and $x'_1 = x'_2$).

To sum up, observers in S and S' agree on the simultaneity of events occurring at the same point in space. But, they will disagree on simultaneity of events occurring at different positions in space: If two events occurring at different positions in space are simultaneous in S, they will not be simultaneous in any other inertial frame moving uniformly with respect to S.

Similarly, we can show that events occurring at the same point in space but at different times in S will appear to be occurring at different points in S'. Thus, if $x_1 = x_2$ and $t_1 \neq t_2$ in S, then $x'_1 \neq x'_2$ in S'. We have left this as an exercise for you in the following SAQ.

In the above discussion we have illustrated the fundamental difference between classical relativity and special relativity. In classical relativity, observers in S and S' always agree about simultaneity at all points. You can see that for $v/c \ll 1$ or $\gamma \to 1$, we find $t' = t$ everywhere. This is not so in special relativity. An observer in S may record two events to occur at different positions at the same time $(x_2 \neq x_1, \ t_2 = t_1)$. But an observer in S' will record these events as occurring at different times $(t'_2 \neq t'_1)$. Similarly, if an observer in S measures two events occurring at the same position but at different times $(x_2 = x_1, t_2 \neq t_1)$, the observer in S' will measure the two at different positions $(x'_2 \neq x'_1)$. You must understand thoroughly this breakdown of simultaneity at all points in space-time. It is the

origin of the phenomena of **length contraction** and **time dilation**, i.e., the relativity of length and time-interval measurements. Let us now study these phenomena.

2.3.2 Length Contraction

You have just studied that events that are simultaneous in one inertial frame may not be simultaneous in another. You have also deduced from Lorentz transformation equations that two observers in S and S' who agree about simultaneity of events occurring at one point in space-time will disagree about the simultaneity of events occurring at different points. This has an interesting consequence for the measurement of lengths in two inertial frames in uniform relative motion.

The problem is this : Suppose a body's length is measured to be L_0 in an inertial frame in which it is at rest. Will an inertial observer in relative motion with respect to the body, measure the length of the body to be the same, i.e., L_0 ? To find the answer, consider a rod R_1 (Fig. 2.1a) lying along the x-axis and at rest in reference frame S. Since the rod is at rest in S, the position coordinates of its ends, say x_2 and x_1, are independent of time. Its length in S is

$$L_0 = x_2 - x_1 \tag{2.11a}$$

The length L_0 of the rod in the frame in which it is at rest is also called its **rest length** or **proper length**. Similarly, consider a rod R_2 lying along the x' axis (Fig. 2.1b) and at rest in frame S' with length

$$L_0 = x'_2 - x'_1 \tag{2.11b}$$

Since the rod is at rest with respect to S', L_0 is the rest length or proper length of the rod in S'.

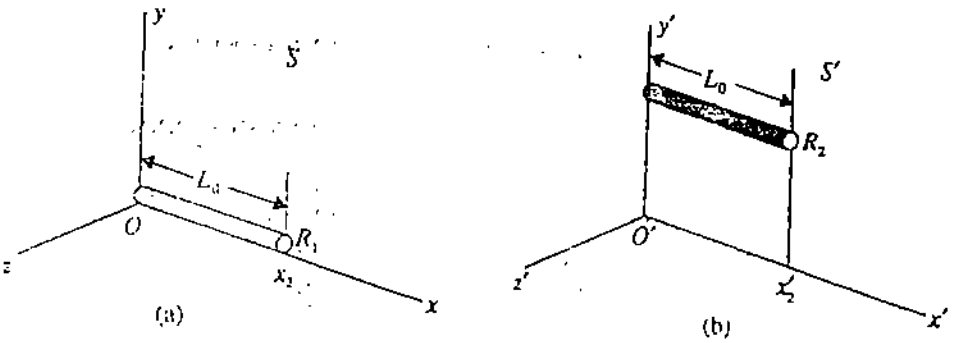


Fig.2.1 : (a) A rigid rod R_1 of proper length L_0 in its rest frame S; (b) a similar rigid rod R_2 of proper length L_0 in its rest frame S'. Note that $x_1 = 0$ and $x'_1 = 0$ in the figures.

We now wish to measure the lengths of these rods from a moving reference frame. To do so, let us assume that S' moves with a uniform velocity $v\,\hat{i}$ with respect to S. Then our problem reduces to measuring the length of the rod R_1 (at rest in S) from S'. For this we have to determine in S' the positions x'_1 and x'_2 that coincide with the ends of the rod at a given time t'; that is the time t' is the same for measuring both x'_1 and x'_2. In other words, we are defining the length L of the rod R_1 in the moving frame S' as follows: It is the distance between positions x'_1 and x'_2 in S' which coincide simultaneously (in S') with the end points of the rod.

Since we are making the space and time measurements in S', we will make use of the inverse Lorentz transformation. We are essentially comparing x and x' when $\Delta t' = 0$. Remember that in choosing the relevant set of transformation equations, it is important to determine in which frame the measurements of the ends of the length are simultaneous.

$$x_1 = \gamma (x'_1 + v t'_1)$$

$$x_2 = \gamma (x'_2 + v t'_2)$$

$$\therefore \quad x_2 - x_1 = L_0 = \gamma (x'_2 - x'_1) + \gamma v (t'_2 - t'_1)$$

At this point we would like to introduce a note of caution. It is easy to fall into the following trap "In the rest system, the ends of the rod have coordinates x_1 and x_2 at some time $t = 0$. To find the length in the moving frame we use the transformation $x' = \gamma (x - vt)$ and obtain $L = \gamma L_0$. Hence, the length of the moving stick is greater." The error is introduced by ignoring the fact that end points must be measured simultaneously in the moving frame, i.e., $t'_1 = t'_2$. These measurements will, of course, not be simultaneous in the rest frame, i.e., $t_1 \neq t_2$ as you have studied in Sec. 2.2.1

$$\begin{aligned} L &= x'_2 - x'_1 \\ &= \gamma (x_2 - vt_2) - \gamma (x_1 - vt_1) \\ &= \gamma [(x_2 - x_1) - v (t_2 - t_1)] \end{aligned}$$

If you erroneously assume $t_1 = t_2 = 0$, say, you arrive at the faulty conclusion that $L = \gamma L_0$. In fact, the length is being measured in S'. Hence $t'_1 = t'_2$ and $t_1 \neq t_2$.

Now since we have to measure x_1' and x_2' at the same time in S', we must let $t_2' = t_1'$. Then we get

$$L_0 = \gamma(x_2' - x_1') = \gamma L$$

or

$$L = \frac{L_0}{\gamma} = L_0\left(1 - \frac{v^2}{c^2}\right)^{1/2} \tag{2.12}$$

Since $\gamma > 1$, $L < L_0$. In other words, the measurement in a moving frame S' gives a shorter length than the proper length measured in a stationary frame (Fig.2.2).

Alternatively, we can measure the length of rod R_2 (at rest in S') from S. Now S moves with a velocity $-v\hat{i}$ with respect to the rod R_2 at rest in S'. The length of R_2 in S is the distance between the positions x_1 and x_2 in S which coincide simultaneously with the end points of the rod R_2. Now we are comparing x' with x when $\Delta t = 0$. From the Lorentz transformation Eq. (2.7a), we have

$$x_1' = \gamma(x_1 - vt_1)$$

and

$$x_2' = \gamma(x_2 - vt_2)$$

so that

$$x_2' - x_1' = L_0 = \gamma(x_2 - x_1) - \gamma v(t_2 - t_1)$$

Letting $t_2 = t_1$, we get

$$L_0 = \gamma(x_2 - x_1) = \gamma L$$

or

$$L = L_0(1 - v^2/c^2)^{1/2} \tag{2.12}$$

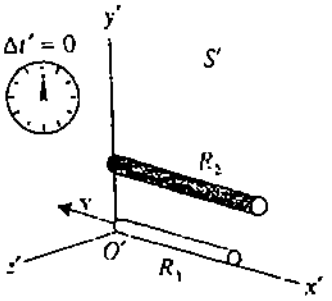


Fig. 2.2: The rod R_1 which is at rest in S but has a speed v in S' will be measured to have a length $L = L_0(1 - v^2/c^2)^{1/2}$ in S'. Note that $x_1 = x_1' = 0$ in the figure. Thus, the length of R_1 will be measured to be less than the length of R_2, the rod at rest in S'.

Once again we find that the measurement of length from a moving frame gives a lower value compared to the measurement in a stationary frame (Fig. 2.3). Thus, the length of an object is a relative quantity – it depends on the frame of reference in which it is measured.

This is called the **Lorentz-Fitzgerald contraction** of a rod moving parallel to its length with respect to the observer. Why does this happen?

Did you note that in both the cases we have emphasized that the observer measures the positions of the end points of the rod **simultaneously** in his (her) own reference frame. For instance, the observer in S' would measure the length of the rod at rest in S, in a way that x_2' and x_1' are measured at the same time ($t_2' = t_1' = t'$) in S'. But this act of simultaneously measuring x_1' and x_2' at time t' in S' does not transform into simultaneous measurement of the endpoints x_1 and x_2 in S, i.e., $t_2 \neq t_1$ in S. Here, t_2 and t_1 are the times measured in the frame S when $t_2 = t_1' = t'$ were measured in S'. From the Lorentz transformation equation (2.7d) we get a time interval in S for registering of the two end points that was done simultaneously in S' as follows :

$$t_2' = \gamma\left(t_2 - \frac{vx_2}{c^2}\right)$$

and

$$t_1' = \gamma\left(t_1 - \frac{vx_1}{c^2}\right)$$

For $t_2' = t_1'$ in S', we get

$$\gamma\left(t_2 - \frac{vx_2}{c^2}\right) = \gamma\left(t_1 - \frac{vx_1}{c^2}\right)$$

or

$$t_2 - t_1 = \frac{v}{c^2}(x_2 - x_1)$$

Since $x_2 \neq x_1$, it follows that $t_2 \neq t_1$.

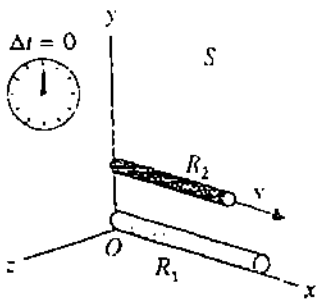


Fig. 2.3 : The rod R_1 which has speed v in S will be measured to have length $L = L_0\left(1 - \frac{v^2}{c^2}\right)^{1/2}$ in S. Note that $x_1 = x_1' = 0$ in the figure.

To sum up, the situation is as follows: There are two events – the measurements of two positions coincident with the end points of a rigid rod. These two events occurring at different spatial positions ($x_2' \neq x_1'$) are simultaneous in S' ($t_2' = t_1'$) but not in S with

respect to which S' is moving at a speed v. The result is that the measurement of length in S' is L_0/γ, i.e., it is shorter than the measured proper length of the rod in S. Thus, the relativity of simultaneity results in the fact that the measurement of length done in a moving inertial frame of reference yields a smaller value compared to the proper length measured in an inertial frame.

You may be wondering whether the rod has 'actually contracted'. Rest assured, the rod does not undergo any physical change. It is the process of measurement in the moving frame that has given a different result.

What about length measurement in the direction perpendicular to relative motion? For instance, what would the length of rod R_1 be in S', if it were stationary in S along y-axis (or z-axes)? You can readily see from Eqs. (2.7b and c) that

$$y' = y \qquad \text{and} \qquad z' = z$$

It means that in this case the measured length of the rod is the same in S and S'. Thus, if the rod moves perpendicular to its length, the measured length is independent of the motion of the observer.

It is common in relativity to speak of the frame in which the observed system (a rod in this case) is at rest as the proper frame. Thus S is a proper frame. The length of the rod in the frame in which it is at rest is called the rest length or proper length. Thus L_0 is the proper length of the rod in S.

Let us write this result in a more general form.

Length Contraction

An observer in motion, relative to two points at a fixed distance L_0 in a stationary frame, along the line joining the two points will measure that distance L to be L_0/γ, that is, shorter than the proper distance L_0.

$$L = L_0/\gamma$$

The effect of length contraction becomes particularly significant when the velocity of an object approaches the speed of light. For example, for $v = 0.9c$, the ratio L/L_0 reduces to 0.44. That is, the measured length is less than half of the proper length.

You may be wondering whether this consequence of special theory of relativity has been verified experimentally. The answer is yes. Let us know about it now.

Experimental evidence for length contraction

Detection of μ-mesons near the earth's surface is a direct evidence for length contraction. When cosmic rays interact with gaseous particles in the upper layers of the atmosphere at a height of about 10 km from the earth's surface, μ-mesons are produced in large numbers. (The μ-mesons are highly unstable with an average life time of 2.2×10^{-6} s.) Their speeds can be as high as 0.998 c. So in their life time the μ-mesons are able to travel a distance of $(2.2 \times 10^{-6}\text{s}) \times (3 \times 10^8\ \text{m s}^{-1}) \times 0.998 = 658$ m. However, some of the μ-mesons are detected near the earth's surface after travelling a distance of 10 km. How is this explained? To unfold this puzzle, we use the relation $L = L_0/\gamma$, where L is the distance travelled by a μ-meson in its own frame of reference and L_0 is the distance corresponding to the earth's frame of reference where we make the measurements. It is given by

$$L_0 = \gamma L = [1 - (0.998)^2]^{-1/2} \times 658\ \text{m} = 10.4\ \text{km}$$

Hence, the fact that in spite of their short life time, μ-mesons are able to reach the ground from great altitudes where they are produced provides experimental evidence for length contraction.

We end this discussion with an example.

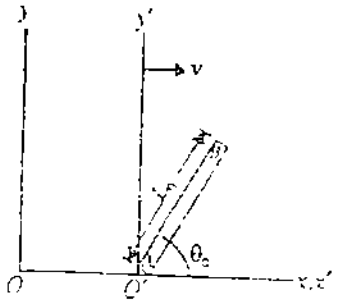


Fig. 2.4 : Orientation of a moving rod.

Example 1: The Orientation of a Moving Rod

A rod of proper length L_0 lies in the $x'y'$ plane of its rest frame S' and makes an angle θ_0 with the x'-axis. What is the length and orientation of the rod in the inertial frame S in which the rod moves to the right with a velocity $\mathbf{v} = v\hat{\mathbf{i}}$?

Solution

Let the end points of the rod have coordinates (x', y') in the frame S'. If one end of the rod is at the origin of S' (Fig 2.4), then we can write

$$x_1' = 0, \qquad y_1' = 0$$
$$x_2' = L_0 \cos\theta_0, \qquad y_2' = L_0 \sin\theta_0$$

We need to find out the coordinates of the end points of the rod at a time t in the frame S in which the rod is moving. From the Lorentz transformation Eqs. (2.7a and b) we can write

$$x_1' = 0 = \gamma(x_1 - vt), \qquad y_1' = 0 = y_1$$
$$x_2' = L_0 \cos\theta_0 = \gamma(x_2 - vt), \qquad y_2' = L_0 \sin\theta_0 = y_2$$

Hence,

$$x_2' - x_1' = L_0 \cos\theta_0 = \gamma(x_2 - x_1)$$

or
$$x_2 - x_1 = \frac{L_0 \cos\theta_0}{\gamma}$$

Similarly, we find that

$$y_2 - y_1 = L_0 \sin\theta_0$$

The length of the rod as measured in S is

$$L = [(x_2 - x_1)^2 + (y_2 - y_1)^2]^{1/2}$$
$$= L_0 \left[\left(1 - \frac{v^2}{c^2}\right)\cos^2\theta_0 + \sin^2\theta_0\right]^{1/2}$$
$$= L_0 \left(1 - \frac{v^2}{c^2}\cos^2\theta_0\right)^{1/2}$$

The angle that the rod makes with the x-axis is

$$\theta = \tan^{-1}\left(\frac{y_2 - y_1}{x_2 - x_1}\right) = \tan^{-1}(\gamma \tan\theta_0)$$

Thus $\theta \geq \theta_0$ since $\gamma \geq 1$.

Thus, the moving rod seems both contracted and rotated.

Can you now explain Mr. Tompkins' adventure described in the course introduction? You may like to try an SAQ to help you check your grasp of these ideas.

Spend 10 min

SAQ 6

(a) A scale is rotated from a position that is parallel to the direction of motion to a position that is perpendicular in a spaceship travelling at a very high speed. Will an observer in the spaceship measure a change in its length? Explain your answer. What change does an observer on the earth, with respect to which the spaceship is moving, measure in the scale's length?

(b) A rod of proper length 1 m measures only 50 cm in a reference frame that is moving with respect to the rod. What is the speed of the moving reference frame?

Let us next investigate the effect of relative motion on measurement of time.

The word 'dilate' literally means 'enlarge beyond normal size'. In connection with time, it means to lengthen an interval of time.

2.3.3 Time Dilation

Let us consider the measurement of a time interval in a frame S' in which the measuring device (the clock) is at rest. It is called the **proper time interval** and is denoted by τ. We can say that the proper time interval is the time interval between two events occurring at the same position in the rest frame of the clock. Then a *non-proper* (or *improper*) time interval would be a time interval measured by two different clocks at two different positions. Thus the proper time interval in S' is

$$\tau = t_2' - t_1' \tag{2.13}$$

where t_2' and t_1' are the instants of time at which two events occur at the same position in S'. Let us consider the situation when the events occur at the same position in S', $x_2' = x_1' = x'$. The question now is: What time interval is measured between these two events by a clock at rest in a frame S, with respect to which S' is moving at a velocity v ($= v\hat{i}$)? Using the inverse Lorentz transformation Eq. (2.8d), we get

$$t_1 = \gamma (t_1' + v x_1'/c^2) \tag{2.14a}$$

$$t_2 = \gamma (t_2' + v x_2'/c^2) \tag{2.14b}$$

Here we have used inverse transformation because the events occur at the same position in the S' frame. From Sec.2.2.1 you may recall that events occurring at the same position in S' need not occur at the same position in S, i.e., $x_2 \neq x_1$. Hence, in S we have to place two clocks at two different positions (x_1 and x_2) to measure the time interval $(t_2 - t_1)$ between the two events. Thus $(t_2 - t_1)$ is an improper time interval. From Eqs. (2.14a and b), it is given by

$$t_2 - t_1 = \gamma (t_2' - t_1') + \frac{\gamma v}{c^2} (x_2' - x_1')$$

$$= \gamma (t_2' - t_1') \quad (\text{since } x_2' = x_1')$$

Thus

$$t_2 - t_1 = \gamma \tau = \frac{\tau}{(1 - v^2/c^2)^{1/2}} \tag{2.15}$$

Since for non-zero v, $\gamma > 1$, the time interval measured by the S clocks between the same events is longer than the time interval between them measured by S' clocks. For instance, 10 s measured on the S' clock will be recorded as 11.5s on the S clock if S' is moving with a speed of $c/2$ with respect to S (since $\gamma = 1.15$ in this case). Since x' is arbitrary, this result applies to all S' clocks. This phenomenon is called **time dilation**. This also implies that to an observer in S, the moving S' clock is measured to slow down by a factor of $(1 - v^2/c^2)^{1/2}$ (see Fig.2.5b). Like length, the duration of time interval is a relative quantity. The rate at which a clock runs depends upon the frame in which it is measured.

You will often come across such statements: 'moving clocks run slow'. What this really means is explained ahead.

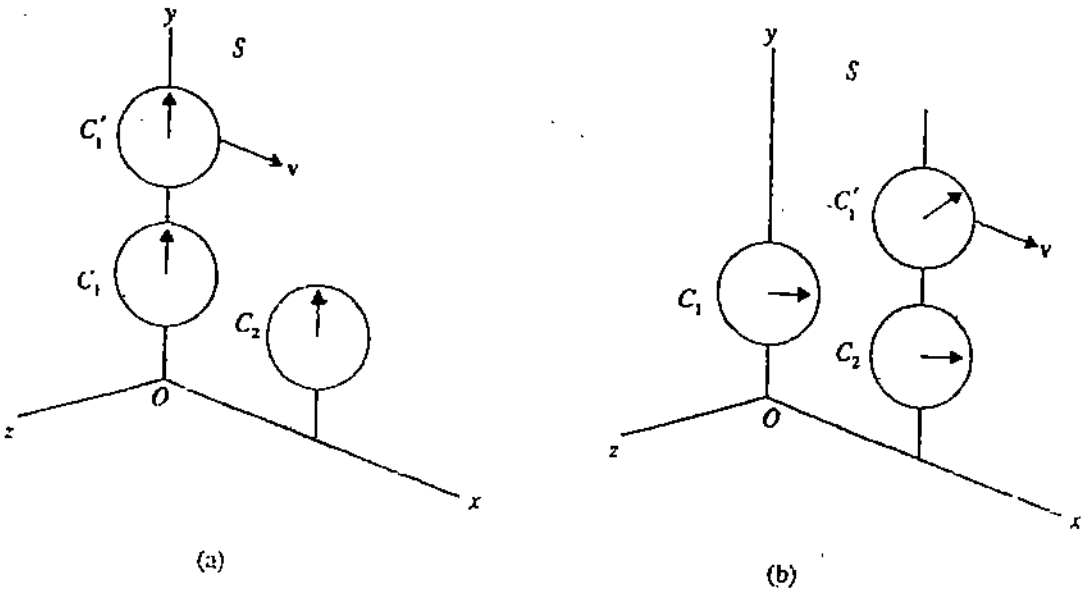


Fig.2.5 : (a) Clocks C_1 and C_2 are situated at fixed positions in S. Clock C_1' which is at rest in S', has a velocity v with respect to S along the x-axis. Suppose $t' = 0$ when $t = 0$; (b) Lorentz transformation yields that the time interval $\Delta t = \gamma \Delta t'$. Here Δt, the improper time interval, is measured by the clocks C_1 and C_2 in S. It is the time interval between two events occurring at the same position in S', the frame attached to the moving clock but at different positions in S. Thus, it takes a longer time in S to register the proper time interval $\Delta t'$ measured in S'. So, compared to the clocks C_1 and C_2 in S, the moving clock C_1' runs slow by a factor $(1 - v^2/c^2)^{1/2}$.

A clock is measured to go at its fastest when it measures the proper time interval, i.e., it measures the time interval between events occurring at a fixed position in a frame, say S'. Thus, the clock is at rest with respect to the frame S'. When it moves uniformly with a velocity v relative to an inertial frame, say S, the stationary clocks in S record a longer (improper) time interval between the two events than the moving clock. Thus, the clock moving with respect to S goes slow by a factor $(1 - v^2/c^2)^{1/2}$ relative to the stationary clocks in S.

You can argue that motion is relative and to an observer in S', th clock in S is moving. Hence, its rate should also slow down as measured by the stationary clocks in S'. This correct provided that the proper time interval between events measured at the same position in S $(x_1 = x_2)$. Then we use the Lorentz transformation (Eq. 2.7d) to ob

$$= \gamma (t_1 - vx_1)$$

$$= \gamma (t_2 - vx_2$$

whence $t_2' - \; = \; \gamma (t_2 - t_1)$ (since $x_2 = x_1$)

or $t_2' - \; = \; \gamma \tau = \frac{\tau}{(1 - v^2/c^2)^{1/2}}$ (2.

where τ is the proper time interval in S. important point to is that the clock measuring time interval of two events occurring at the same position gives the shortest time interval. In this case, the clock in S. This is what means ... say that a clock at rest with respect to observer goes at its fastest ... the clock in S is moving with a speed ... time interval recorded Δt in S ... recorded as $\gamma \Delta t$ by clock in S'.

Let us summarise what we have learnt in this sub-section.

---

tion

- Consider inertial reference ... uniform relative ... each frame ... clocks held in that frame.
- Suppose events occur at a ... ion in S' separated ... proper time interval Δt measured by a clock in S'. Then the improper time interval between the events measured in S, with respect ... S' is moving, will be long; will $\Delta t = \gamma$ ... with respect to the ... clock in S, the moving ... clock in S will run ...
- Alternatively, suppose ... over a fixed position in S ... ted by proper time interval Δt measured by clocks at rest in S ... the improper time interval between these events measured by clocks in S' ... respect to which S is moving, will be $\Delta t' =$ ... it will be longer ... respect to stationary clocks in S', the moving ... S will run slow.

A word of caution is needed here. ... understand time dilation in a ... way and you may be somewhat ... this notion ... To enable you ... up your ideas and understanding of ... further consider an example.

Example 2: Comparison of Time-Interval Measurements

Imagine a passenger sitting in a train that is moving with uniform velocity $\mathbf{v}$ ($= v\hat{i}$) with respect to ground. The passenger measures ... clock consisting of a light pulse reflected between two mirrors on a cylindrical tube. The clock is at rest in the reference frame of the passenger (Fig. 2.6a). ... time the pulse of light strikes the mirror M_1, it is registered by a click. The time interval between two successive clicks is ... passenger frame S' (Fig. 2.6b) a proper time interval since it is measured at the same position of the clock at rest in the S' frame.

Now suppose another observer fixed to the ground (S) measures the time interval between two successive clicks of the clock in the S' frame. The question is: what is the

time interval between two successive clicks in S' as measured by an observer in S? For the observer in S, the train, the passenger and the clock move to the right during this time interval (Fig.2.6c). So the time interval in S will be measured using two stationary clocks. One clock will record the time of the first click (A) and the other the time of the next click (C). In this manner, the observer in S compares the reading of one moving clock with that of two stationary clocks (Figs. 2.6b,c and d).

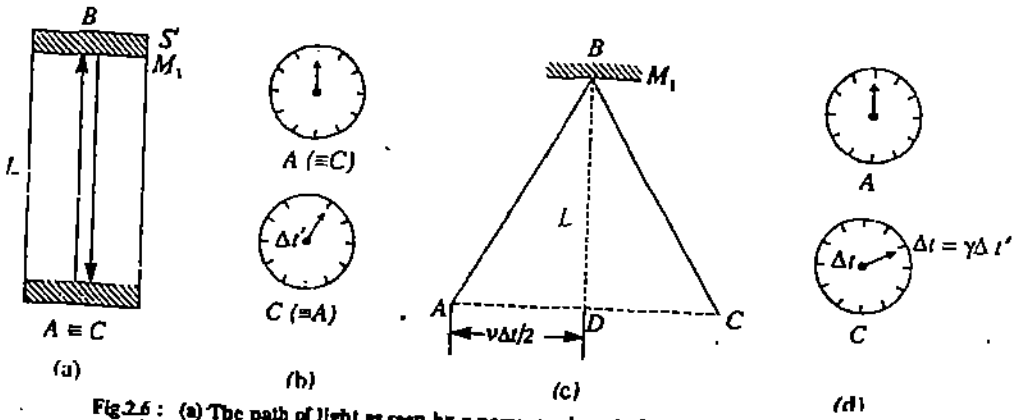


Fig.2.6 : **(a) The path of light as seen by a passenger in a clock at rest in the moving frame S' attached to the train. A click is registered each time the light pulse strikes the mirror M_1. In S', the points A corresponding to the first click and C to the next click are the same points; (b) the proper time interval ($\Delta t'$) between two successive clicks registered by the passenger on his (her) clock which is moving with respect to S; (c) the path of the light pulse in S' as observed by an observer in S. To S, A and C are located at different positions since the train has moved to the right at a speed v; (d) the improper time interval Δt registered by two stationary (synchronized) clocks located at the points A and C as measured in S fra... is longer, since $\Delta t = \gamma \Delta t'$.**

Since the train is moving to the right with respect to S, the position of the clock placed in S' changes relative to S. Hence, the path of the light pulse as observed from S is longer (see Fig. 2.6c). Thus, for the observer on the ground light travels greater distance than for the passenger on the train. Since the speed of light is the same in both the frames, the observer in S measures longer time to elapse between the two clicks than does the passenger in S' (see Figs. 2.6b and 2.6d). The observer in S concludes that the passenger's clock runs slow. Quantitatively, we can write the result as follows.

Refer to Fig. 2.6c. The time interval between two successive clicks measured by the clock in S' is

$$\Delta t' = \frac{2L}{c} \tag{2.17}$$

where L is the length of the tube. The time interval between two successive clicks as measured in S is

$$\Delta t = \frac{AB + BC}{c} = \frac{2AB}{c} \tag{2.18}$$

But from Pythagorean theorem

$$AB = [L^2 + (AC/2)^2]^{1/2}$$

Here $AC = v\,\Delta t$, since the train travels this distance in time Δt at speed v. Hence Eq. (2.18) becomes

$$\Delta t = \frac{2}{c}\left[L^2 + \left(\frac{v\,\Delta t}{2}\right)^2\right]^{1/2}$$

This is a quadratic equation in Δt. On solving it for Δt we get two roots:

$$\Delta t = \pm \frac{2L/c}{(1 - v^2/c^2)^{1/2}}$$

of which we neglect the negative root. Thus

$$\Delta t = \gamma \Delta t'$$

This is the same result as Eq. (2.15).

Let us compute some actual time intervals. Suppose S' is moving at a speed 7S_c relative to S. Then $\gamma = 1.5$ and $\Delta t = 1.5\, \Delta t'$. Thus, to an observer in S during the time clock in S' registers two clicks ($\Delta t' = 2$ units), the clock in S will register three clicks = 3 units). In other words, the moving clock ticks out time more slowly than a stationark.

Conversely, to S' the moving clok in S will tick out time more slowly led the clock in S measures the proper time interval. In this context we would like to out that statements like 'moving clocks rs slow' can be misunderstood. What tement actually means is that a clock moving at a constant velocity relative to tial frame S runs slow when timed by the clos in S. Remember, in all cases we aring to measurement of time and to noonfuse the term observer with one vs. An observer is one who measures phical quantities. We will now presentxperimental evidences for time dilation.

Experimental evidences for timilation

Time dilation, like length contrac, has been confirmed experimentalservations on muons (μ-mesons) by B. Rond D.B Hall in 1941. The mean lif muons is 2.2 μs. This is so small tat evemuons twelled at the speedof lighe top of the atmosphere where thy are cd, they ould not reach the surfacearth (a distance of about 10 kmsince fistance twelled would be only (2.) × ($3 \times 10^8 \, ms^{-1}$) = 660 m. et thesons are detected in laboratories on. This can be explained only by thfact thour frame ofeference the life timon is dilated by a factor γ . i. it is $\times 10^{-6}$ s). Thu if v = 0.998c, γ = e life time of a muon in our framoncreas65 μs enabling to travel such lon × 16 = 10560 m = 10.56 kmdistandis result is alconsistent with lenaction in the muon's frame – thistance0.5 km is contacted to 660 m in s frame – Indeed in equivalent erimen muons (at γ12) in the CERN in Europe in 1968, muovere d with dilatedtimes in accordanq. (2.15) to an accuracy of 1 %tomic flown around earth have also time dilation effects thus firming is a real phenon.

Time dilation effectsume gnificance for energy nuclear working at high-energy acceleors with of paicles h decay very fas have very small man lifes. For e. pis and s decay spntaner 100 times as fasas muWere it time ation would day a ear before theyid trad a few en at speight. Duo tin, their mean life is incr and they is led d Thus thean bd hundreds oket fme point acceler whey are placebre. they canbeed fether expes. In majme dilation becmater of everyday sucerhese phy

For furhercltrity, shound not outAC

SAQ 7

a) Why do we mesrve the of thisics ...

b) A free neutron at rest has a mean time ... of a free neutron as 2700 s, how ...

c) When a charged pion is at rest, an ... 10^{-8} s. If the ... time of charged pionsmeasured ... at a speed of 0.98c ... observer at a speed of 0.98c ... measure their mean life time to?

d) Light takes 2 years travel to the star ... an ... astronaut travelled fro the earth at ... speed of 0... how ... it take according to the astrons clock. Him/her reach there?

In the ... part (SAQ 7 you have rech at an interesting result. Suppose the astronaut

(say A) turns his/her spaceship around as soon as he reaches Alpha Centauri and returns to Earth travelling at the same speed. According to the astronaut, the round trip will take four years in all. But to the astronaut's twin (B) who stayed back on earth, the time elapsed would be $4.7 \times 2 = 9.4$ years. Thus, the twin who stayed back would age more than the twin who went on the space trip. This is a famous example in relativity known as the twin paradox. We now discuss this briefly.

The twin paradox

The so called **twin paradox** is this: Motion is relative. So cannot B say with equal right that it was (s)he who went on the trip and A remained where (s)he was? Consequently, should B not be younger when they meet?

The answer is no and we resolve the paradox as follows:

B has remained at rest in a single inertial frame. But A, in the simplest case of uniform to and fro motion (to the star and back) has to be accelerated briefly out of B's frame into another. Then A has to be decelerated again briefly to turn around and finally decelerated to stop at B's frame. These accelerations (positive and negative) are felt by A. Therefore, A can be under no illusion that it was (s)he who remained at rest. So there is a lack of symmetry in the situations of A and B – **they are not equivalent.** This gives rise to the age difference between them. Thus, the paradox is disposed of as soon as we can point out the asymmetry.

Apart from the effects of length contraction and time dilation there are some other consequences of special relativity theory. For instance, consider this problem: A spaceship, let us call it Enterprise, moves in one direction at a speed of $0.9c$ while another spaceship, named Endeavour moves in the opposite direction at a speed of $0.9c$. What is the relative speed of the ships? Classically, it would be $1.8c$ and Endeavour's crew would see the Enterprise moving at a speed faster than that of light. But according to special relativity the picture is quite different. The transformation of velocities under Lorentz transformation gives an entirely different result. Let us see what it is.

2.4 RELATIVISTIC TRANSFORMATION OF VELOCITY

Let the reference frame S' move with uniform velocity $V\hat{i}$ relative to the reference frame S. Suppose a particle moves with uniform velocity v relative to the S frame. What is the velocity v' of the particle relative to the S' frame?

Let us consider the x-component of the velocity v (see Fig. 2.7).

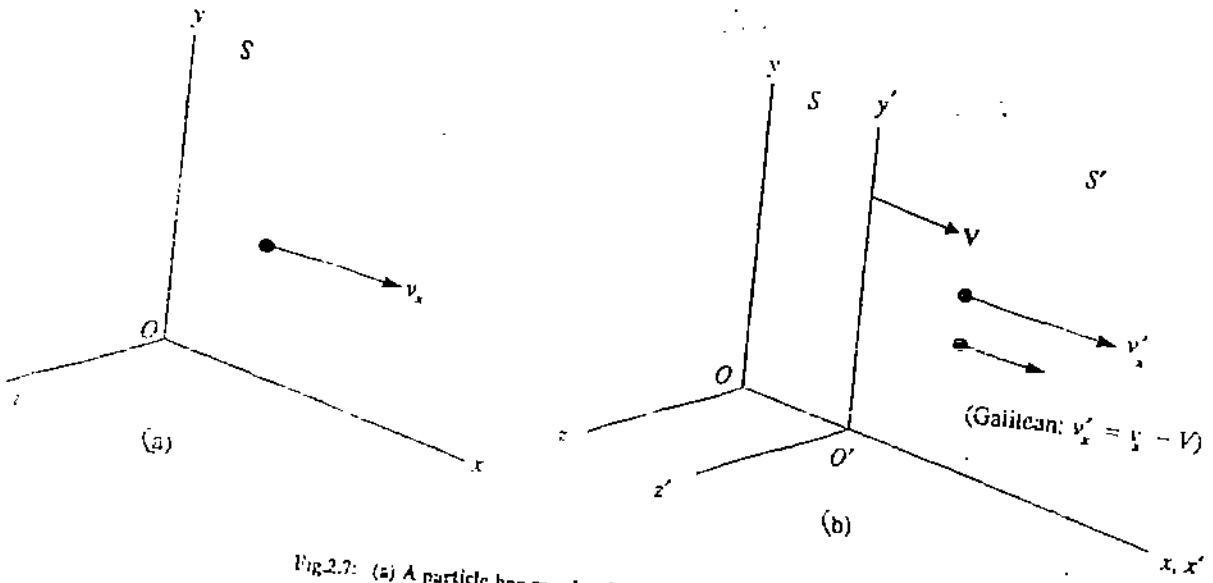


Fig.2.7: (a) A particle has speed v_x in the S frame; (b) according to the Lorentz transformation its speed v_x' in the S' frame moving with velocity $V\hat{i}$ with respect to S is $(v_x - V)/(1 - v_xV/c^2)$. Recall that the Galilean transformation predicts a value $v_x' = v_x - V$.

From Eq. (2.7a) we have

$$x' = \gamma(x - Vt); \qquad t' = \gamma\left(t - \frac{Vx}{c^2}\right)$$

whence

$$dx' = \gamma(dx - V dt), \qquad dt' = \gamma\left(dt - \frac{V dx}{c^2}\right)$$

The components of the velocity v′ are

$$v_x' = \frac{dx'}{dt'} = \frac{dx - V\,dt}{dt - V\,dx/c^2} = \frac{\frac{dx}{dt} - V}{1 - \frac{V}{c^2}\frac{dx}{dt}}$$

or $$v_x' = \frac{v_x - V}{1 - v_x V/c^2} = \frac{v_x - \beta c}{1 - v_x \beta/c}, \qquad \beta = \frac{V}{c} \tag{2.19 a}$$

Compare this result with that obtained from Galilean transformation: $v_x' = v_x - V$. What about the y and z components of v′? Since $y = y'$ and $z = z'$, we have

$$v_y' = \frac{dy'}{dt'} = \frac{dy}{\gamma(dt - V\,dx/c^2)} = \frac{\frac{dy}{dt}}{\gamma\left(1 - \frac{V}{c^2}\frac{dx}{dt}\right)}$$

(2.19 b)

or $$v_y' = \frac{v_y}{\gamma(1 - v_x V/c^2)}$$

(2.19 c)

and $$v_z' = \frac{v_z}{\gamma(1 - v_x V/c^2)}$$

You can obtain the inverse transformations from Eqs. (2.8) or by solving Eqs. (2.19a) to (2.19c) for the unprimed velocity components. Would you like to do an SAQ for this purpose?

(2.20 a)

(2.20 b)

$$v_y = \gamma(1 - \quad)$$

(2.20 c)

$$v_z = \gamma(1 - \quad)$$

Note that for $V << c$, these reduce to the Galilean transformation. Eqs. (2.19a to c) give the relativistic transformation of velocities. The inverse transformation formulae given by Eqs. (2.20a to c) are also referred to as relativistic velocity addition formulae. These equations can be regarded as giving the resultant of the two velocities $\mathbf{v}' = (v_x', v_y', v_z')$ and $\mathbf{V} = (V, 0, 0)$ to be $\mathbf{v} = (v_x, v_y, v_z)$.

We can apply the transformation equations (2.19 and 2.20) by reconsidering the example of the spaceships Enterprise and Endeavour. The speed of Enterprise relative to the earth is 0.9c, and the speed of Endeavour relative to earth is −0.9c. So, as observed in the S system, the spaceships are travelling with velocities ± 0.9c opposite to each other. Let the

So the second pulse is received at the origin of S' at a time $t' + \Delta t'$. Since the intial pulse arrives at the origin of S' at $t' = 0$, the total time elapsed between the reception of the two pulses at $x' = 0$ in S' is given by

$$t' + \Delta t' = \frac{\tau(1 + u/c)}{(1 - u^2/c^2)^{1/2}} = \tau \left[\frac{1 + u/c}{1 - u/c}\right]^{1/2} \tag{2.24}$$

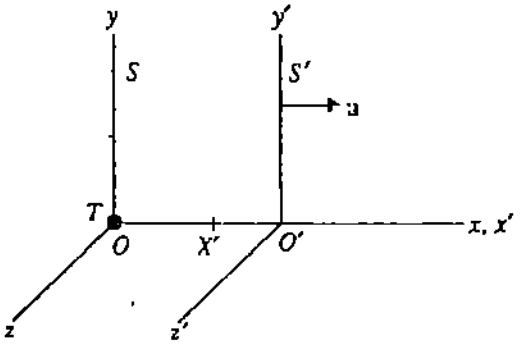


Fig.2.9: A transmitter T in S at $x = 0$ sends flashes of light with a period τ. The successive pulses of light are received in S', the initial pulse arriving at $x' = 0$ at $t' = 0$; $x' = X' = -u\tau(1 - u^2/c^2)$. is the point at which the second pulse is received in S' at $t' = \frac{\tau}{(1 - u^2/c^2)^{1/2}}$.

Do you realise that this is essentially a time dilation effect which occurs due to the relative motion of the transmitter and receiver? The time interval between two flashes as measured in S' is an improper time interval. The source of light in S acts as a clock which measures the proper time interval between the two flashes at the same point $x = 0$. To an observer in S' it acts as a moving clock which runs slow compared to the clocks at rest in S'. Hence an observer in S' measures a longer time interval between the two flashes [given by Eq. (2.24)] compared to the time interval normal measured by an observer in S.

The time between two successive flashes received in S' can equally well be interpreted as the time period of the light wave emitted from the transmitter in S, measured in S'. The frequency is the reciprocal of the period of the wave, so that

$$\nu' = \frac{1}{t' + \Delta t'} = \frac{1}{\tau}\left(\frac{1 - u/c}{1 + u/c}\right)^{1/2}$$

$$\text{or} \qquad \nu' = \nu\left(\frac{1 - u/c}{1 + u/c}\right)^{1/2} = \nu\left(\frac{1 - \beta}{1 + \beta}\right)^{1/2} \tag{2.25}$$

Here ν' is the frequency of the light wave as received in S', and ν $(= 1/\tau)$ is the frequency of the light wave transmitted in S. If the receiver is receding from the source, then u is positive and ν' is less than ν. You should note that even if the source were receding from the receiver, we would get the same result. This is unlike the Doppler effect in sound where the two results are different. If the receiver is approaching the source, we take u to be negative and ν' is greater than ν. Again the result would be the same if the source were approaching the receiver. In terms of wavelength, $\lambda = c/\nu$ and $\lambda' = c/\nu'$, so that

$$\lambda' = \lambda\left(\frac{1 + u/c}{1 - u/c}\right)^{1/2} = \lambda\left(\frac{1 + \beta}{1 - \beta}\right)^{1/2} \tag{2.26}$$

Eqs. (2.25 and 2.26) describe the relativistic longitudinal Doppler effect for light waves in vacuum. It expresses the effect when the relative motion of the source and observer is along the same axis.

Let us now sum up what you have studied in this unit.

2.6 SUMMARY

- The new coordinate transformation consistent with the postulates of special relativity is called the **Lorentz transformation**. It links the coordinates (x, y, z, t) assigned to an event in an inertial frame S, with the coordinates (x', y', z', t') assigned to the same event in another inertial frame S' moving at a velocity $\mathbf{v} = (v\hat{\mathbf{i}})$ with respect to S:

$$\begin{aligned} x' &= \frac{x - vt}{(1 - v^2/c^2)^{1/2}} = \gamma (x - \beta ct) \\ y' &= y \\ z' &= z \\ t' &= \frac{t - vx/c^2}{(1 - v^2/c^2)^{1/2}} = \gamma(t - \beta x/c) \end{aligned}$$

where $\gamma = 1/(1 - v^2/c^2)^{1/2}$, $\beta = v/c$. The **inverse Lorentz transformation** is

$$\begin{aligned} x &= \frac{x' + vt}{(1 - v^2/c^2)^{1/2}} = \gamma (x' + \beta ct) \\ y &= y' \\ z &= z' \\ t &= \frac{t' + vx/c^2}{(1 - v^2/c^2)^{1/2}} = \gamma(t' + \beta x/c) \end{aligned}$$

- The postulates of special relativity and the Lorentz transformation lead to **relativity of simultaneity, relativity of length and time- interval measurements.**

- **Simultaneity is relative.** Two events that occur simultaneously in one inertial frame of reference at two different positions are not necessarily simultaneous in another inertial frame of reference. And two events that occur at the same position but at two different times in one inertial frame, do not necessarily occur at the same positions in other inertial frames.

- **Length Contraction:** The length of an object depends upon the inertial frame in which it is measured. If L_0 is the proper length in the inertial frame in which the object is at rest, then in the inertial frame with respect to which the object is moving, its length is given by

$$L = \frac{L_0}{\gamma}$$

- **Time Dilation:** The duration of a time interval is a relative quantity which depends on the inertial frame in which it is measured. A clock moving uniformly with respect to an inertial frame of reference S goes slow by a factor $(1 - v^2/c^2)^{1/2}$ relative to the stationary clocks in S. The clocks in S record a longer time interval Δt for the events occurring at the same position in the inertial frame S' of the moving clock and timed by the moving clock as $\Delta t'$:

$$\Delta t = \gamma \Delta t'$$

- The **Lorentz transformation of velocity** in two frames S and S' in relative motion is :

$$v_x' = \frac{v_x - V}{1 - v_x V/c^2}$$

$$v_y' = \frac{v_y}{\gamma (1 - v_x V/c^2)}, \qquad \gamma = \frac{1}{(1 - V^2/c^2)^{1/2}}$$

$$v_z' = \frac{v_z}{\gamma (1 - v_x V/c^2)}$$

where (v_x, v_y, v_z) are the velocity components in S, (v_x', v_y', v_z') the velocity components in S' which is moving with a velocity $\mathbf{V} = V\hat{\mathbf{i}}$ relative to S.

- The inverse velocity transformation is also referred to as the **relativistic velocity addition formulae :**

$$v_x = \frac{v_x' + V}{(1 + v_x' V/c^2)}$$

$$v_y = \frac{v_y'}{\gamma(1 + v_x' V/c^2)}$$

$$v_z = \frac{v_z'}{\gamma(1 + v_x' V/c^2)}$$

- The application of special relativity to Doppler effect in optics leads to a correction. The **relativistic Doppler effect** relates the frequency ν of a light wave in a frame S in which the source is at rest to the frequency ν' as measured by an observer in S' moving relative to S at a velocity $\mathbf{v} = v\hat{\mathbf{i}}$:

$$\nu' = \nu\left(\frac{1-\beta}{1+\beta}\right)^{1/2}, \quad \beta = \frac{V}{c} \quad \text{(Source and observer receding)}$$

and

$$\nu' = \nu\left(\frac{1+\beta}{1-\beta}\right)^{1/2}, \quad \text{(Source and observer approaching)}$$

2.7 TERMINAL QUESTIONS

Spend 45 min

1. An observer in frame S, assigns the following coordinates to two events E_1 and E_2:

   E_1: $x_1 = 1.2 \times 10^9$m, $y_1 = 0$, $z_1 = 0$, $t_1 = 7$s

   E_2: $x_2 = 3.0 \times 10^9$m, $y_2 = 0$, $z_2 = 0$, $t_2 = 11$s

   Determine the coordinates of E_1 and E_2 as measured by an observer in S' moving relative to S at a speed $4c/5$. Let E_1 represent pushing the button of a detonator and E_2 an ensuing explosion. Does your result imply that in frame S' the explosion will occur even before the button is pushed? How will you resolve this paradox?

2. a) Show that two events simultaneous in S which are separated by Δx in space will be separated in S' in both space and time, such that

   $$\Delta x' = \gamma \Delta x, \qquad \Delta t' = -\frac{\beta}{c}\gamma \Delta x$$

   where S' is moving relative to S at a speed v in the x direction.

   b) Show that if L_0^3 is the rest volume of a cube, then the volume measured in a frame moving with a velocity v in a direction parallel to the edge of the cube is $L_0^3(1 - v^2/c^2)^{1/2}$.

   c) The mean life time of an Ω^--particle as measured by us is 7.4×10^{-10}s when it moves past us at such a speed that $\gamma = 9.0$. What is the proper mean life time of the particle?

3. We observe two galaxies receding in opposite directions at speeds 0.3c. What speed of recession would an observer in one of these galaxies observe for the other galaxy?

4. Protons are accelerated so that they attain a velocity 2×10^8cm s^{-1}. Afterwards they drift at a constant velocity through a region where they are neutralised to H atoms. In this process light is emitted and observed in a spectrometer. What is the Doppler shift of the wavelength in the light spectrum? The wavelength of light emitted when the atom is at rest is $\lambda = 4861.33\text{Å}$.

5. Show that, with $v'^2 = v_x'^2 + v_y'^2$ and $v^2 = v_x^2 + v_y^2$, we can write

$$c^2 - v^2 = \frac{c^2(c^2 - v'^2)(c^2 - V^2)}{(c^2 + v_x' V)^2}$$

where the symbols have their usual meanings. This gives a relation between the speed v of a particle in S with its speed v' in S'.

2.8 SOLUTIONS AND ANSWERS

SAQs (Self-Assessment Questions)

1. Substituting the Galilean coordinate transformations in Eq. (2.1b) we get

$$(x - vt)^2 + y^2 + z^2 = c^2 t^2$$

or $$x^2 - 2vtx + v^2t^2 + y^2 + z^2 = c^2t^2$$

which is not the same as Eq. (2.1a).

Hence, the Galilean coordinate transformations do not satisfy both Eqs. (2.1a) and (2.1b) at the same time.

2. $$2a_1a_2 = 2c^2a_1b_1 \quad (1)$$

$$a_1^2 - c^2b_1^2 = 1 \quad (2)$$

$$a_2^2 - a_1^2c^2 = -c^2 \quad (3)$$

We have to determine a_1, a_2 and b_1 from these three equations (1–3). To do so we also make use of Eq. (2.4a) whence $a_2 / a_1 = -v$. Thus, from (3) we get

$$a_1^2v^2 - a_1^2c^2 = -c^2, \quad \text{i.e.,} \quad a_1^2 = \frac{c^2}{c^2 - v^2} = \frac{1}{1 - v^2/c^2}$$

or $$a_1 = \frac{1}{(1 - v^2/c^2)^{1/2}}$$

$$\therefore \quad a_2 = -\frac{v}{(1 - v^2/c^2)^{1/2}} \quad \text{from (3)}$$

and (1) yields

$$b_1 = \frac{a_2}{c^2} = -\frac{v}{c^2(1 - v^2/c^2)^{1/2}}$$

Thus Eqs. (2.2a to d) become

$$x' = \frac{x}{(1 - v^2/c^2)^{1/2}} - \frac{vt}{(1 - v^2/c^2)^{1/2}} = \frac{1}{(1 - v^2/c^2)^{1/2}}(x - vt)$$

$$y' = y$$

$$z' = z$$

$$t' = -\frac{vx}{c^2(1 - v^2/c^2)^{1/2}} + \frac{t}{(1 - v^2/c^2)^{1/2}} = \frac{1}{(1 - v^2/c^2)^{1/2}}\left(t - \frac{vx}{c^2}\right)$$

since $$b_2 = a_1.$$

3. Since, $v/c \ll 1$, we can neglect v^2/c^2 in comparison with 1 and

$$x' = \frac{x - vt}{(1 - v^2/c^2)^{1/2}} \cong x - vt$$

$$y' = y$$

$$z' = z$$

In the last equation, let us consider the motion of the origin O' given by $x = vt$. Then

$$t' = \frac{t - v^2 t/c^2}{(1 - v^2/c^2)^{1/2}} = t\,(1 - v^2/c^2)^{1/2}$$

When $v/c << 1$, this equation yields

$$t' = t$$

Thus, the Lorentz transformation reduces to the *Galilean transformation* for speeds much smaller than the speed of light.

4. Substituting for γt from Eq. (2.7d) in (2.7a) we get

$$x' = \gamma x - \beta c\left(t' + \frac{\gamma \beta x}{c}\right)$$

$$= \gamma x\,(1 - \beta^2) - \beta c t'$$

$$= \frac{x}{\gamma} - \beta c t'$$

or $$x = \gamma\,(x' + \beta c t')$$

which is Eq. (2.8a).

Substituting for γx from Eq. (2.7a) in Eq. (2.7d) we get

$$t' = \gamma t - \frac{\beta}{c}(x' + \gamma \beta c t)$$

$$= \gamma t\,(1 - \beta^2) - \frac{\beta}{c} x'$$

$$= \frac{t}{\gamma} - \frac{\beta}{c} x'$$

or $$t = \gamma\,(t' + \beta x'/c)$$

which is Eq. (2.8d).

5. a) From Eq. (2.7a)

$$x'_2 - x'_1 = \gamma\,(x_2 - x_1) - \beta c\,(t_2 - t_1)$$

Since $x_2 = x_1$, but $t_2 \neq t_1$, we have

$$x'_2 - x'_1 = \beta c\,(t_1 - t_2) \neq 0 \text{ as long as } \beta \neq 0.$$

Hence $x'_2 \neq x'_1$. Thus events occurring at the same position in S but at different times, need not occur at the same position in S'.

b) At $x = 0$, $t = 0$ implies $t' = 0$ from Eq. (2.7d). However, at all other points ($x \neq 0$), Eq. (2.7d) gives

$$t' = \gamma\,(t - vx/c^2)$$

Hence, if $x \neq 0$, $t' \neq t$. Thus, events occurring at all other points in space do not occur at the same time in S and S'. Similarly, Eq. (2.7a) gives

$$x' = \gamma\,(x - vt) \neq x \text{ except at } x = 0 \text{ and } t = 0.$$

Thus events occurring at all other points in space are not simultaneous in S and S'.

6. a) No. Since the scale is at rest in the frame of the spaceship, its length to an observer in the spaceship will remain the same. From Example 1, you can see that to an observer on the earth, the scale is moving as well as rotating. When the scale is parallel to the direction of motion, its length will be measured to be the shortest. As it is rotated towards a direction perpendicular to the direction of motion (i.e., as $\theta_0 \to 90°$), its length will be measured to increase. Finally when the scale is perpendicular to the direction of motion its length will be measured to be the same in the spaceship and on the earth.

(b) Here $L = 50\text{cm} = 0.5\text{m}$ and $L_0 = 1.0\text{m}$

Thus $$\gamma = \frac{L_0}{L} = \frac{1.0\,\text{m}}{0.5\,\text{m}} = 2$$

or $$\frac{1}{(1 - v^2/c^2)^{1/2}} = 2$$

$$\therefore \quad 1 - \frac{v^2}{c^2} = \frac{1}{4}$$

$$\frac{v^2}{c^2} = \frac{3}{4}$$

$$\therefore \quad v = \frac{\sqrt{3}}{2} c = 2.6 \times 10^8\ \text{m s}^{-1}$$

7. a) This is because the speeds we encounter in everyday life are much smaller than c and the factor γ is approximately equal to 1. Thus $t' = t$ and we do not observe the effect of time dilation.

b) The proper mean life time of the neutron at rest is 900s. For the neutron moving relative to us it is 2700s. Thus

$$2700\text{s} = \gamma \times 900\text{s}$$

or $$\gamma = 3$$

which yields $v = 2.8 \times 10^8\ \text{m s}^{-1}$.

c) The proper mean life time of the charged pion is 2.6×10^{-8}s. Here $v = 0.98c$ yielding

$$\gamma = \frac{1}{(1 - v^2/c^2)^{1/2}} = 5$$

The mean life time of moving pions is $\gamma \times$ proper mean life time

$$= 5 \times 2.6 \times 10^{-8}\text{s} = 1.3 \times 10^{-7}\text{s}$$

To an observer moving with the pion, the pion would be at rest and its mean life time would be 2.6×10^{-8}s.

d) The astronaut travelling at $v = 0.9c$ would require

$$t = \frac{4.2}{0.9}\ \text{years} = 4.7\ \text{years}$$

to complete the journey according to an observer on earth. Since the astronaut's clock is at rest with respect to him, it measures the proper time interval $\Delta t'$ and Δt is the improper time interval. These are related by

$$\Delta t = \gamma \Delta t'$$

whence $$\Delta t' = \frac{\Delta t}{\gamma}\ \text{years}$$

$$= \frac{4.7}{\gamma}$$

For $v = 0.9\,c$,

$$\gamma = \frac{1}{\left(1 - \frac{v^2}{c^2}\right)^{1/2}} = 2.3$$

$$\therefore \quad \Delta t' = \frac{4.7}{2.3} = 2\ \text{years}$$

Thus, in the astronaut's frame the interval of time elapsed is 2 years.

8. From Eqs. (2.8) we have

$$x = \gamma(x' + Vt'), \qquad t = \gamma\left(t' + \frac{Vx'}{c^2}\right)$$

$$\text{whence } dx = \gamma(dx' + Vdt'), \quad dt = \gamma\left(dt' + \frac{Vdx'}{c^2}\right)$$

$$\text{Thus, } v_x = \frac{dx}{dt} = \frac{(dx' + Vdt')}{dt' + \frac{Vdx'}{c^2}} = \frac{\frac{dx'}{dt'} + V}{1 + \frac{V}{c^2}\frac{dx'}{dt'}} = \frac{v_x' + V}{1 + Vv_x'/c^2}$$

which is Eq. (2.20 a).

$$v_y = \frac{dy}{dt} = \frac{dy'}{\gamma\left(dt' + \frac{Vdx'}{c^2}\right)} = \frac{v_y'}{\gamma(1 + Vv_x'/c^2)} \qquad \text{(Eq. 2.20 b)}$$

Similarly,

$$v_z = \frac{v_z'}{\gamma(1 + Vv_x'/c^2)} \qquad \text{(Eq. 2.20 c)}$$

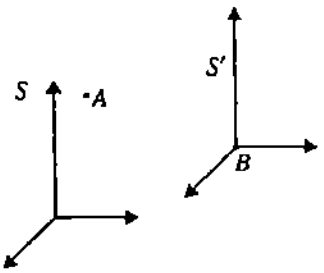


Fig. 2.10

9. a) Here we can apply the relativistic velocity transformation equations. Let S be a frame fixed on the earth and let S' be the frame attached to the spacecraft B (Fig. 2.10). Thus $v_x = 0.9c$, is the relative speed of A with respect to the earth. The relative speed of B with respect to A is given to be $0.5c$ and we have to determine its speed with respect to S, i.e., V. Since the relative speed of S' with respect to A is 0.5c, the speed of A in S', i.e., $v_x' = -0.5c$. Hence from Eq. (2.19a) we have

$$v_x' = \frac{v_x - V}{1 - v_x V/c^2}$$

$$\text{or} \quad -0.5\,c = \frac{0.9c - V}{1 - 0.9\,V/c}$$

$$\text{or} \quad -0.5c + 0.45\,V = 0.9c - V$$

$$\text{or} \quad 1.45\,V = 1.4\,c$$

$$V = \frac{1.4}{1.45}\,c = 0.97\,c$$

Thus, the relative velocity of B with respect to the earth is $0.97c$.

(b) Here we use the velocity addition formulae Eqs. (2.20). Here the velocity of the rocket (S') with respect to the earth(S) is $V = 0.9c$. It is given an additional velocity $v_x' = 0.4\,c$ in the frame S' attached to it. Hence the final velocity relative to the earth is

$$v_x = \frac{v_x' + V}{1 + v_x' V/c^2} = \frac{0.4\,c + 0.9\,c}{1 + 0.4 \times 0.9}$$

$$= \frac{1.3}{1.36}\,c = 0.96\,c$$

Terminal Questions

1. The space-time coordinates of E_1 and E_2 in the S' frame are obtained from Eqs. (2.7a to d). Here

$$\gamma = \left(1 - \frac{v^2}{c^2}\right)^{-1/2} = \left(1 - \frac{16}{25}\right)^{-1/2} = \frac{5}{3}$$

For E_1: $x_1' = \frac{5}{3} \times [1.2 \times 10^9 \text{m} - 4c/5 \times 7\text{s}]$

$= 2.0 \times 10^9 \text{ m} - 2.8 \times 10^9 \text{ m}$

or $x_1' = -8 \times 10^8 \text{ m}$

$y_1' = y_1 = 0$

$z_1' = z_1 = 0$

$t_1' = \frac{5}{3}\left[7\text{s} - \frac{1.2 \times 10^9 \text{ m} \times 4c}{5c^2}\right]$

$= \frac{35}{3}\text{s} - \frac{16}{3}\text{s}$

or $t_1' = 6.3 \text{ s}$

For E_2: $x_2' = \frac{5}{3} \times \left[3.0 \times 10^9 \text{ m} - \frac{4c}{5} \times 11 \text{ s}\right]$

$= 5.0 \times 10^9 \text{ m} - 4.4 \times 10^9 \text{ m}$

or $x_2' = 6 \times 10^8 \text{ m}$

$y_2' = y_2 = 0$

$z_2' = z_2 = 0$

$t_2' = \frac{5}{3}\left[11\text{s} - \frac{3.0 \times 10^9 \text{ m} \times 4}{5 \times 3 \times 10^8 \text{ ms}^{-1}}\right] = 5\text{s}$

In frame S, E_1 occurs before E_2 but in frame S', E_1 occurs after E_2, i.e., there seems to be a reversal in the order of events. We now have to find whether in such a case events E_1 and E_2 are related to each other. In other words, can a cause-effect relationship exist between E_1 and E_2? Can E_1 cause E_2 or vice-versa? What is the distance between E_1 and E_2 in S? It is $(x_2 - x_1) = 1.8 \times 10^9$m. The time interval between them is 4s. The distance light can travel in 4s is 1.2×10^9m, which is less than the distance between the two events. Again in S' the distance $(x_2' - x_1') = 1.4 \times 10^9$m and it is greater than the distance $c(t_1' - t_2') = 4.0 \times 10^8$m.

Thus, $(x_2 - x_1) > c(t_2 - t_1)$

and $(x_2' - x_1') > c(t_1' - t_2')$

Since nothing travels faster than light, in either frame, the two events cannot be connected by any kind of signal travelling from E_1 to E_2. Thus, E_1 cannot be the cause of E_2 in frame S'. Thus, there is no paradox in the reordering of the two events as they are independent of each other. These two cannot, therefore, represent the pushing of the button and the explosion.

2. a) It is given that $x_2 \neq x_1$ and $t_2 = t_1$ in S. We use the Lorentz transformation Eqs. (2.7a and d) since the events are simultaneous in S and obtain

$x_1' = \gamma(x_1 - vt_1)$

$x_2' = \gamma(x_2 - vt_2)$

whence $x_2' - x_1' = \gamma(x_2 - x_1) - \gamma v(t_2 - t_1)$

or $\Delta x' = \gamma \Delta x$, since $t_2 = t_1$

Similarly, $t_1' = \gamma(t_1 - \beta x_1/c)$

$$t_2' = \gamma (t_2 - \beta x_2/c)$$

whence $$t_2' - t_1' = \gamma (t_2 - t_1) - \frac{\gamma\beta}{c} (x_2 - x_1)$$

or $$\Delta t' = -\frac{\gamma\beta}{c} \Delta x, \text{ since } t_2 = t_1.$$

b) Since the moving frame moves along one edge of the cube, say the x-edge, the length of that edge only is contracted, it is

$$L = \frac{L_0}{\gamma}$$

The length of the other two edges remain the same, i.e. L_0, since those lengths are in directions perpendicular to the line of motion. Hence the volume of the cube as measured in S' will be

$$\frac{L_0}{\gamma} \times L_0 \times L_0 = L_0^3 (1 - v^2/c^2)^{1/2}$$

c) The proper mean life time $\Delta t'$ of the Ω^- particle in its frame is obtained from the relation $\Delta t = \gamma \Delta t'$. It is given that $\Delta t = 7.4 \times 10^{-10}$s and $\gamma = 9.0$. Therefore,

$$\Delta t' = \frac{\Delta t}{\gamma} = 8.2 \times 10^{-11}\text{s}$$

3. Let S be the frame attached to the earth where we are situated. The galaxy 1 has a velocity

$$v_x = +0.3c, \qquad v_y = 0, \qquad v_z = 0$$

Let the frame S' be attached to galaxy 2 which is moving in a direction opposite to 1. Then its velocity components relative to S are given as $V_x = -0.3c$, $V_y = 0$, $V_z = 0$. Therefore, the speed of galaxy 1 as seen from S', i.e., galaxy 2, is given by

$$v_x' = \frac{v_x - V}{1 - v_x V/c^2} = \frac{0.3c - (-0.3c)}{1 + (0.3)^2} = \frac{0.6c}{1.09}$$

$$= 1.65 \times 10^8 \text{ m s}^{-1}$$

4. The rest wavelength of the light is $\lambda = 4861.33$ Å (1 Å $= 10^{-10}$m). Therefore, the shifted wavelength is obtained from Eq. (2.26) with u to be negative since the protons are moving towards the region in which they are neutralised. Therefore,

$$\lambda' = \lambda \left(\frac{1 - u/c}{1 + u/c}\right)^{1/2}$$

Here $u = 2 \times 10^8$ cm s^{-1} $= 2 \times 10^6$ m s^{-1}

$$\therefore \qquad \lambda' = 4861.33 \left(\frac{1 - 2/300}{1 + 2/300}\right)^{1/2} \text{Å} = 4861.33 \left(\frac{298}{302}\right)^{1/2} \text{Å}$$

or $$\lambda' = 4829.03 \text{ Å}$$

5. Using Eqs. (2.20 a to c) we can write

$$v^2 = v_x^2 + v_y^2$$

$$= \frac{(v_x' + V)^2}{(1 + v_x' V/c^2)^2} + \frac{v_y'^2 (1 - V^2/c^2)}{(1 + v_x' V/c^2)^2}$$

$$= \frac{v_x'^2 + V^2 + 2v_x' V + v_y'^2 - v_y'^2 V^2/c^2}{(1 + v_x' V/c^2)^2}$$

$$= \frac{v'^2 + V^2 + 2v_x' V - v_y'^2 V^2/c^2}{(c^2 + v_x' V)^2} \cdot c^4 \text{ (since } v^2 = v_x'^2 + v_y'^2)$$

Hence

$$c^2 - v^2 = \frac{c^2 (c^2 + v_x' V)^2 - c^4 (v'^2 + V^2 + 2v_x' V - v_y'^2 V^2/c^2)}{(c^2 + v_x' V)^2}$$

$$= \frac{c^6 + c^2 v_x'^2 V^2 + 2v_x' V c^4 - c^4 v'^2 - c^4 V^2 - 2v_x' V c^4 + c^2 v_y'^2 V^2}{(c^2 + v_x' V)^2}$$

$$= \frac{c^6 + c^2 v'^2 V^2 - c^4 v'^2 - c^4 V^2}{(c^2 + v_x' V)^2}$$

$$= \frac{c^2 (c^2 - v'^2)(c^2 - V^2)}{(c^2 + v_x' V)^2}$$

UNIT 3 RELATIVISTIC DYNAMICS

Structure

3.1 Introduction
Objectives

3.2 Dynamics of a Single Particle
The Need to Redefine Linear Momentum
Relativistic Linear Momentum
Relativistic Force Law

3.3 Relativistic Energy
The Equivalence of Mass and Energy
Relativistic Energy and Momentum of a Free Particle

3.4 Summary

3.5 Terminal Questions

3.6 Solutions and Answers

3.1 INTRODUCTION

In Unit 1 you have studied that Newtonian mechanics is incompatible with the special theory of relativity. With the advent of particle accelerators it was experimentally established that Newton's laws do not apply to the motion of electrons moving at high speeds. In Unit 2 we obtained the new transformation equations (the Lorentz transformation) and examined their implications for length and time measurements. Now we must consider dynamic phenomena and suitably modify Newtonian mechanics so that it is consistent with the special theory of relativity.

We shall begin our study with the dynamics of a single particle and first examine the concepts of linear momentum and inertial mass (Sec. 3.2). You will learn that these basic entities need to be redefined. Having obtained the correct expressions for linear momentum and relativistic mass we shall rewrite Newton's force law and apply it to the motion of particles at high speeds.

In mechanics, we define force as the rate of change of linear momentum with time $\mathbf{F} = \frac{d\mathbf{p}}{dt}$. Recall that you have studied another way of defining a conservative force field—in terms of the space-rate of change of potential energy [$\mathbf{F} = -\nabla U$]. So, if the force law of mechanics is recast, the concept of energy would also need to be re-examined. In doing so we will arrive at the famous relation $E = mc^2$ demonstrating the equivalence of mass and energy. This single principle has, perhaps, transformed our world in unparalleled ways. You will discover its power in Sec. 3.3.

"One thing I have learned in a long life : that all our science, measured against reality, is primitive and childlike—and yet it is the most precious thing we have."

—Albert Einstein

Objectives

After studying this unit you should be able to

- derive the expression for relativistic linear momentum,
- apply relativistic force law to simple situations,
- interpret and apply the relativistic laws of conservation of linear momentum and energy,
- compute the mass, speed, momentum and energy of a relativistic particle.

3.2 DYNAMICS OF A SINGLE PARTICLE

In this section our aim is to reformulate Newton's laws of motion so that they are consistent with the special theory of relativity. You know that Newton's second law

defines force as the rate of change of linear momentum which is the product of mass and velocity of a body. Recall that one of the basic assumptions of Newtonian mechanics is that the mass of a body is independent of its state of motion with respect to the observer. Thus, equal forces acting on a body would produce equal accelerations, whatever be the instantaneous velocity of the body. So, if we applied a force indefinitely on a body, its velocity would go on increasing indefinitely, at a constant rate. *This stands in contradiction with our result* of Sec. 2.4 *that no material object can travel with a speed greater than the speed of light in vacuum.* Can we, therefore, intuitively conclude that the mass of a body should increase with its velocity and tend to infinity as the velocity approaches c? Then this modified definition of mass would yield a different linear momentum. This is one point of departure from Newton's laws.

Another difficulty with classical mechanics is that it requires action and reaction to be equal and opposite at all instants. If it applies to forces in contact (acting at the same point), even according to relativity we can say that they act at the same instant. However, for forces acting at a distance in relativity, the same instant differs from one inertial observer to another. Therefore, we cannot give meaning to action and reaction independent of the observer's frame of reference.

Thus, in modifying Newtonian mechanics we have to exclude the notion of forces 'acting-at-a-distance'. But we can include collision phenomena involving contact forces or field phenomena (e.g., force on charges in an electromagnetic field). Let us now consider the concrete example of collisions and understand further why we need to redefine linear momentum.

3.2.1 The Need to Redefine Linear Momentum

From Newton's second law we know that if the net external force on a system is zero, its linear momentum remains conserved. We would like to retain the law of conservation of linear momentum in special relativity as well. The reason is that conservation laws arise basically due to symmetry properties of space-time. These properties, in turn follow from the homogeneity and isotropic nature of space. Recall that in formulating Lorentz transformation in Sec. 2.2, we have assumed space to be homogeneous and isotropic. Hence, it is logical to assume that the law of conservation of linear momentum should be valid in the modified Newtonian mechanics as well.

The next question is: Can we use the classical expression ($\mathbf{p} = m\mathbf{v}$) for linear momentum? Let us consider the conservation of linear momentum in collisions and see if its classical definition allows the law to be invariant under Lorentz transformation.

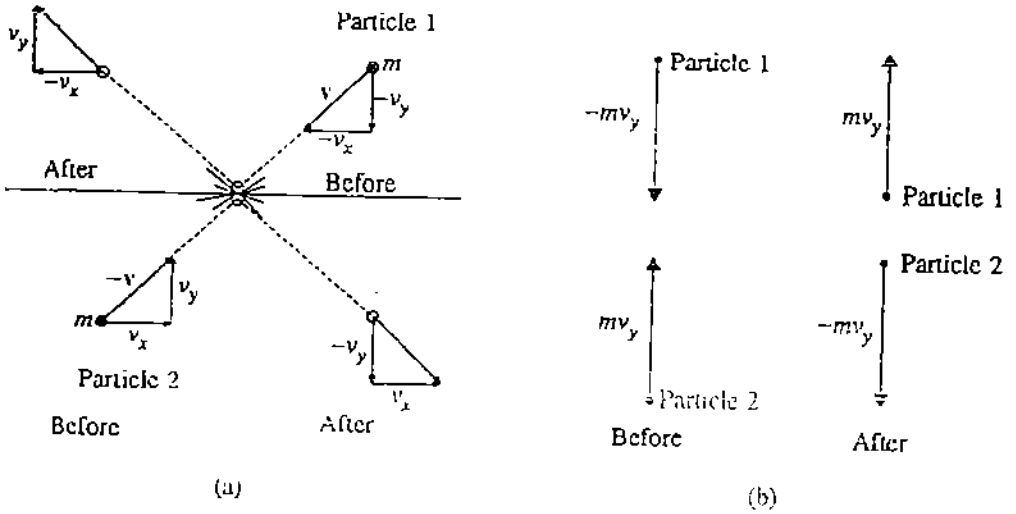


Fig. 3.1: (a) A collision between two particles of equal mass m in the xy plane. The velocities in x and y direction before and after collision are shown; (b) the non-relativistic momenta of particle 1 and 2 in the y-direction.

Consider a collision between two point particles of equal mass m (Fig. 3.1a). Let us choose a frame of reference S such that the particles approach each other with equal and opposite velocities. In this reference frame the centre of mass is at rest. The law of conservation of linear momentum, is given as

$$m_1\mathbf{v}_{1B} + m_2\mathbf{v}_{2B} = m_1\mathbf{v}_{1A} + m_2\mathbf{v}_{2A}$$

where m_1 and m_2 are the masses of the particles; $\mathbf{v}_{1B}$, $\mathbf{v}_{2B}$ and $\mathbf{v}_{1A}$, $\mathbf{v}_{2A}$ their respective velocities before and after collision. In this case $m_1 = m_2$, $\mathbf{v}_{1B} = -\mathbf{v}_{2B} = \mathbf{v}$ implying $\mathbf{v}_{1A} = -\mathbf{v}_{2A}$. Thus the total linear momentum before and after collision is zero. You can verify that linear momentum is conserved in this case.

The x components of each particle's velocity and hence its momentum remains the same before and after collision (Fig. 3.1a). The y velocity components of particles 1 and 2 (Fig. 3.1b) are as follows:

Particle	Before the collision	After the collision
1	$-v_y$	v_y
2	v_y	$-v_y$

Change in the y-momentum component of particle 1 $= mv_y - (-mv_y) = 2mv_y$

Change in the y-momentum component of particle 2 $= -mv_y - mv_y = -2mv_y$

Thus, the total change in the y component of linear momentum in S is zero.

So there is no problem in S with the newtonian definition of momentum $\mathbf{p} = m\mathbf{v}$.

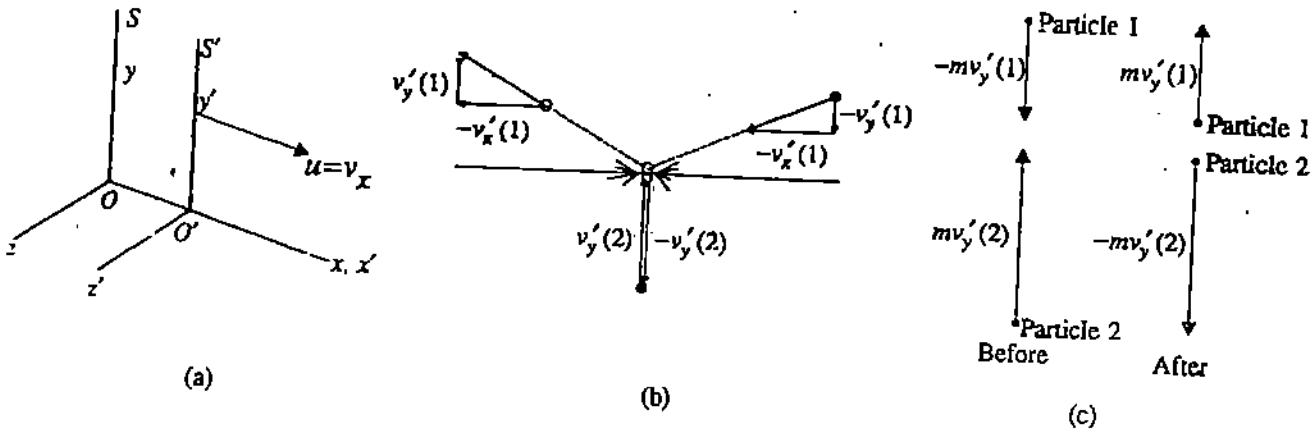


Fig.3.2 : (a) The frame S' has a velocity $u = v_x\hat{i}$ with respect to S; (b) viewed in S', the x velocity component of particle 2 is zero and $v_y'(1) = v_y'(2)$; (c) in S' the nonrelativistic momentum is not the same in the y' direction before and after collision.

Now consider a reference frame S' moving with velocity $\mathbf{u} = u\hat{i} = v_x\hat{i}$ with respect to S (Fig. 3.2a). Note that v_x is the x component of the velocity of particle 2 in S, and $-v_x$ is that of particle 1. From Eqs. (2.19a) to (2.19c) for relativistic velocity transformation, we obtain the velocity components in S' in terms of those in S as follows.

For particle 1:

$$-v_x'(1) = \frac{-v_x - u}{1 + v_x u/c^2} = -\frac{2v_x}{1 + v_x^2/c^2}$$

$$-v_y'(1) = \frac{v_y}{1 + v_x u/c^2}\left(1 - \frac{u^2}{c^2}\right)^{1/2} = \left(\frac{v_y}{1 + v_x^2/c^2}\right)\left(1 - \frac{v_x^2}{c^2}\right)^{1/2}$$

For particle 2:

$$v_x'(2) = \frac{v_x - u}{1 - v_x u/c^2} = 0$$

$$v_y'(2) = \frac{v_y}{1 - v_x u/c^2}\left(1 - \frac{u^2}{c^2}\right)^{1/2} = \frac{v_y}{(1 - v_x^2/c^2)^{1/2}}$$

Thus, the x velocity component of particle 2 is zero (Fig. 3.2b) and the magnitudes of $v_y'(1)$ and $v_y'(2)$ are not equal.

$$v'_y(1) \neq v'_y(2)$$

even though they were equal in S.

Therefore, the changes in nonrelativistic momenta [$-2mv'_y(2)$ and $2mv'_y(1)$] are not equal and opposite in the y' direction before and after collision (see Fig.3.2c). What do we conclude from this? We can say that the definition of momentum given by $\mathbf{p} = m\mathbf{v}$ does not ensure conservation of momentum in all inertial reference frames.

Do we now say that momentum conservation is inconsistent with Lorentz transformation? Or should we redefine momentum in such a way that **conservation of momentum holds in all reference frames moving at constant relative velocities?** Recall that in the beginning of this section we have agreed to uphold the law of conservation of linear momentum. Thus *we now have to look for a definition of momentum which is Lorentz invariant*, i.e., *it should remain unchanged under Lorentz transformation*. This essentially requires looking for a definition of velocity dependent mass as argued in the beginning of this section. But before we do that would you like to take a break and work out an exercise?

SAQ 1

Spend 10 min

Consider the collision of particles of masses m_{1B} and m_{2B} moving with velocities $\mathbf{v}_{1B}$ and $\mathbf{v}_{2B}$. Let their masses and velocities after collision be m_{1A}, m_{2A}, $\mathbf{v}_{1A}$ and $\mathbf{v}_{2A}$, respectively.

a) Write down the law of conservation of momentum for the collision of these particles.

b) In certain special cases when $\mathbf{v}'$ is parallel to $\mathbf{u}$, the velocity transformation law in vector form is given by

$$\mathbf{v} = \frac{\mathbf{v}' + \mathbf{u}}{1 + \mathbf{v}'\,\mathbf{u}/c^2}$$

where $\mathbf{v}$ is the particle's velocity in S frame, and $\mathbf{v}'$ its velocity measured in S' moving with a velocity $\mathbf{u}$ relative to S. Write the law of conservation of momentum in S'. Using this transformation and the fact that the masses of the particles remain the same in both frames, determine whether the conservation of momentum holds in S'.

You have found in SAQ 1(b) that momentum as defined in Newtonian mechanics is not universally conserved in all frames moving at a constant velocity relative to each other. We now have to look for a Lorentz–invariant definition of momentum. In the process we will arrive at a velocity-dependent definition of mass.

3.2.2 Relativistic Linear Momentum

Let us consider a system of N interacting particles P_i ($i = 1, 2, \ldots N$). Let m_i be the mass of the ith particle. At the moment we do not need any detailed information about the mode of interaction between these particles. Let the velocity of the i^{th} particle be $\mathbf{v}_i$ in the inertial frame S and $\mathbf{v}'_i$ in the inertial frame S'. Let the velocity of S' relative to S be $\mathbf{u} = u\hat{i}$. These particles are moving subject to the condition that the total mass and total linear momentum of the system is conserved in frame S, i.e.,

$$\sum_i m_i = \text{constant} \qquad (3.1a)$$

and

$$\sum_i m_i \mathbf{v}_i = \text{constant} \qquad (3.1b)$$

We have to find conditions so that the total mass and linear momentum is also conserved in S' frame, i.e.,

$$\sum_i m'_i = \text{constant} \qquad (3.2a)$$

and

$$\sum_i m'_i \mathbf{v}'_i = \text{constant} \qquad (3.2b)$$

where m_i' is the mass of the ith particle in S' frame.

We will now show that Eqs. (3.1) and (3.2) are compatible only if mass is redefined. For this purpose we will engage in a little algebraic manipulation which you need not memorise. We will use the following relations :

$$\gamma_i = \frac{1}{(1 - v_i^2/c^2)^{1/2}}, \quad \gamma_i' = \frac{1}{(1 - v_i'^2/c^2)^{1/2}}, \quad \gamma = \frac{1}{(1 - u^2/c^2)^{1/2}} \tag{3.3}$$

For the sake of simplicity we shall assume that the particles are in one-dimensional motion along the x-axis. Since v_i and v_i' are in the x-direction, from the velocity transformation laws (Eqs. 2.19 and 2.20) we have

$$v_i' = \frac{v_i - u}{1 - v_i u/c^2} \tag{3.4a}$$

and

$$v_i = \frac{v_i' + u}{1 + v_i' u/c^2} \tag{3.4b}$$

We will now determine a relation between $\gamma_i' v_i'$ and $\gamma_i v_i$. Using Eq. (3.4a) we can write

$$\gamma_i' v_i' = \frac{v_i - u}{1 - v_i u/c^2} \cdot \frac{1}{(1 - v_i'^2/c^2)^{1/2}}$$

$$= \frac{v_i - u}{[(1 - v_i'^2/c^2)(1 - v_i u/c^2)^2]^{1/2}} \tag{3.5}$$

From Eq. (3.4a) we can also show that

$$(1 - v_i'^2/c^2)\left(1 - \frac{v_i u}{c^2}\right)^2 = \left(1 - \frac{v_i^2}{c^2}\right)\left(1 - \frac{u^2}{c^2}\right) \tag{3.6}$$

You can work this out as an exercise.

Substituting Eq. (3.6) in (3.5) we get

$$\gamma_i' v_i' = \frac{v_i - u}{[(1 - v_i^2/c^2)(1 - u^2/c^2)]^{1/2}}$$

or

$$\gamma_i' v_i' = \gamma_i \gamma (v_i - u) \tag{3.7}$$

Now since γ and u are constants, we can rewrite Eqs. (3.1a and 3.1b) as follows without changing them

$$\sum_i m_i \gamma u = \text{constant}$$

$$\sum_i m_i v_i \gamma = \text{constant}$$

Subtracting one equation from the other we get

$$\sum_i m_i \gamma (v_i - u) = \text{constant} \tag{3.8}$$

But from Eq. (3.7), $(v_i - u) = \frac{\gamma_i' v_i'}{\gamma_i \gamma}$. Therefore, Eq. (3.8) yields

$$\sum m_i \frac{\gamma_i' v_i'}{\gamma_i} = \text{constant} \tag{3.9}$$

Comparing Eq. (3.2b) with Eq. (3.9) we obtain

$$m_i' = \frac{m_i \gamma_i'}{\gamma_i} \tag{3.10}$$

Thus, for the conservation of mass and momentum to hold in the S' frame, the mass m_i' should satisfy Eq. (3.10). What are the implications of Eq. (3.10)? For Eqs. (3.2a) and (3.2b) to have an unambiguous meaning, m_i must be independent of any measurement done in the S system, such as that of the particle velocities $\mathbf{v}_i$ or relative velocity $\mathbf{u}$ of the systems. Similarly, m_i' should be independent of any measurements in S' system such as $\mathbf{v}_i'$ and $\mathbf{u}$. Thus, we must have

$$\frac{m_i'}{\gamma_i'} = \frac{m_i}{\gamma_i} = \text{absolute constant} = m_{i0} \text{ (say)} \tag{3.11}$$

We can now express the result of Eq. (3.11) generally as follows : For any particle of mass m and for any inertial frame of reference, if $\mathbf{v}$ is the velocity of the particle relative to that frame, then

$$m = m_0 \gamma = \frac{m_0}{(1 - v^2/c^2)^{1/2}} \tag{3.12a}$$

For a particle at rest $v = 0$ and

$$m = m_0 \tag{3.12b}$$

Thus m_0 in Eq. (3.12b) or m_{i0} in Eq. (3.11) is the **rest mass** of the particle. It is also called the **proper mass** of the particle for it is the mass of the particle measured when it is at rest. Eqs. (3.12a) and (3.12b) tell us that the mass of a particle depends on its velocity and increases with an increase in its velocity. The definition of relativistic momentum then becomes

$$\mathbf{p} = m\mathbf{v} = \frac{m_0 \mathbf{v}}{(1 - v^2/c^2)^{1/2}} \tag{3.13a}$$

The components of momentum are

$$p_x = \frac{m_0 v_x}{(1 - v^2/c^2)^{1/2}}, \quad p_y = \frac{m_0 v_y}{(1 - v^2/c^2)^{1/2}}, \quad p_z = \frac{m_0 v_z}{(1 - v^2/c^2)^{1/2}} \tag{3.13b}$$

Notice that the denominator of each momentum component contains the magnitude v of the total velocity.

Thus, by redefining mass and momentum, we have been able to preserve the form of the momentum conservation law and have also satisfied the requirement that the law be Lorentz invariant, i.e., remain unchanged under Lorentz transformation. The difference is that instead of the mass of individual particles remaining constant, the total mass of the system is conserved (Eqs. 3.1a and 3.2a). This relativistic momentum conservation law has been proved to be true by experiments.

We have arrived at these definitions of mass and momentum considering one-dimensional motion only. The same results are obtained for three-dimensional motion where we have to use the three-dimensional velocity transformation laws. However, this is beyond the scope of this course. You may now like to try an SAQ.

(a) Using the Definitions of mass and linear momentum and show that the linear momentum is conserved in S' frame for the collision process described in SAQ 1.

(b) Find the rest mass of a negatively charged particle which has a momentum of magnitude 1.2×10^{-19} kg m s^{-1} and a speed of $0.6c$.

Spend 15 min

We can now write down the equation of motion of a particle in relativistic mechanics.

3.2.3 Relativistic Force Law

Recall that according to Newton's second law of motion the rate of change of momentum of a body is equal to the net force experienced by the body and is directed in the line of this force. Of course this is true for a suitable choice of units. This law takes the following generalised form which is consistent with the special theory of relativity.

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = \frac{d}{dt}(m\mathbf{v}) \tag{3.14a}$$

or

$$\mathbf{F} = m\frac{d\mathbf{v}}{dt} + \mathbf{v}\frac{dm}{dt} \tag{3.14b}$$

Substituting for m from Eq. (3.12a), we can also write this as

$$\mathbf{F} = \frac{m_0}{(1 - v^2/c^2)^{1/2}}\frac{d\mathbf{v}}{dt} + \mathbf{v}\frac{d}{dt}\frac{m_0}{(1 - v^2/c^2)^{1/2}} \tag{3.14c}$$

However, there is one fundamental difference between the nonrelativistic and relativistic cases. In the nonrelativistic case it is possible to reduce force on a body to a product of its mass and its acceleration. But the relativistic force law [Eq.3.14b] shows that the force acting upon a body and its acceleration may not even be parallel.

You can see that Eq. (3.14b) is not equivalent to writing

$$\mathbf{F} = \left[\frac{m_0}{(1 - v^2/c^2)^{1/2}}\right]\frac{d\mathbf{v}}{dt} = m\mathbf{a}$$

As defined in Eq. (3.14a) the relativistic force law automatically leads to conservation of relativistic linear momentum.

If F is zero, $\mathbf{p} = m_0\,\mathbf{v}/(1-v^2/c^2)^{1/2}$ must be a constant.

Hence, **in the absence of external forces, relativistic momentum is conserved.** Now suppose F as defined by Eq. (3.14a) is not zero, and acts on a body or a system of particles for some time. Then it changes the relativistic momentum of the body or the system by an amount $\Delta\mathbf{P}$ which is equal to the total impulse given to the system

$$\Delta\mathbf{P} = \mathbf{J} = \int \mathbf{F}\,dt \tag{3.15}$$

In the limit $v/c << 1$, Eq. (3.14a) reduces to

$$\mathbf{F} = m_0\frac{d\mathbf{v}}{dt} \tag{3.16}$$

which is the well known Newton's second law. Thus, Newtonian mechanics and relativistic mechanics overlap in a large domain of applications in which the motions of bodies are slow compared with the speed of light, i.e., when $\gamma \equiv 1$. When γ exceeds unity for a particle in motion, Newtonian mechanics is in slight error. For γ factors as high as 10^{11} observed for cosmic ray protons, Newtonian mechanics is totally wrong. But within its acceptable slow motion domain Newtonian mechanics continues to remain valid and useful.

Let us now consider an application of Eqs.(3.14a and b) to relativistic motion of a single particle.

Example 1 : Charged particle motion in a magnetic field

A charged particle moves perpendicular to a uniform magnetic field at relativistic speeds. Determine the radius of the particle's orbit.

Solution

The relativistic equation of motion is

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = \frac{d}{dt}(m\mathbf{v}) = m\frac{d\mathbf{v}}{dt} + \frac{dm}{dt}\mathbf{v}$$

where $m = \frac{m_0}{(1 - v^2/c^2)^{1/2}}$ is the particle's mass and v its velocity. Here **F** is the force on the particle due to the magnetic field and is given by

$$\mathbf{F} = q\mathbf{v} \times \mathbf{B}$$

Since $\mathbf{v} \perp \mathbf{B}$, $F = qvB\hat{\mathbf{n}}$, where $\hat{\mathbf{n}}$ is perpendicular to both **v** and **B**. Recall from Unit 3 of the course PHE-01 (Elementary Mechanics), that a force acting perpendicular to the velocity is a no-work force. So the particle's speed does not change under the action of this force. Only the direction of its velocity changes. Hence the particle's mass remains constant. Therefore, $\frac{dm}{dt} = 0$ and the equation of motion reduces to

$$\mathbf{F} = m\frac{d\mathbf{v}}{dt} = \gamma m_0 \frac{d\mathbf{v}}{dt}$$

Substituting **F** with the magnetic force we obtain

$$qvB\hat{\mathbf{n}} = \gamma m_0 \mathbf{a}$$

Since the particle moves in a circle, the acceleration **a** is simply the centripetal acceleration and we get

$$qvB = \frac{\gamma m_0 v^2}{r}$$

whence the radius of the particle's orbit, r, is

$$r = \frac{\gamma m_0 v}{qB} = \frac{mv}{qB}$$

This is of the same form as the classical expression for the radius (m_0v/qB) except that it is γ times larger. Further, the particle's mass increases with an increase in its speed. The increase in radius by a factor γ explains why many of today's high-energy particle accelerators are so huge.

An exercise at this stage will help you consolidate these concepts.

SAQ 4

The rest mass of proton is 1.67 × 10- kg. Determine the magnitude of force needed to give it an acceleration of 1.0 ×15 ms-2 in the direction of motion when v = 0.90c.

We now come to a crucial step further, taken by Einstein, which greatly extended the revolutionary impact of his special theory of relativity. While developing the notion of relativistic energy Einstein arrived at the now famous formula $E = mc^2$ which demonstrates the equivalence of mass and energy. Let us see how this was done.

3.3 RELATIVISTIC ENERGY

In this section we shall use the relativistic force law to find the expression for relativistic energy. Recall the definition of kinetic energy, from Unit 3 of PHE-01, which is the work done by the net external force in increasing the speed of the particle from some value v_A to v_B :

$$T_B - T_A = \int_A^B \mathbf{F} \cdot d\mathbf{l} \qquad (3.17)$$

where

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = \frac{d}{dt}\left[\frac{m_0 \mathbf{v}}{(1 - v^2/c^2)^{1/2}}\right]$$

Thus

$$T_B - T_A = \int_A^B \frac{d\mathbf{p}}{dt} \cdot d\mathbf{l}$$

$$= \int_A^B \frac{d}{dt}\left[\frac{m_0 \mathbf{v}}{(1 - v^2/c^2)^{1/2}}\right] \cdot d\mathbf{l}$$

Now we can write $d\mathbf{l} = \frac{d\mathbf{l}}{dt}\,dt = \mathbf{v}\,dt$ so that

$$T_B - T_A = \int_A^B \frac{d}{dt}\left[\frac{m_0\,\mathbf{v}}{(1 - v^2/c^2)^{1/2}}\right] \cdot \mathbf{v}\,dt$$

$$= \int_A^B \mathbf{v} \cdot \frac{d}{dt}\left[\frac{m_0\,\mathbf{v}}{(1 - v^2/c^2)^{1/2}}\right] dt$$

$$= \int_A^B \mathbf{v} \cdot d\left[\frac{m_0\,\mathbf{v}}{(1 - v^2/c^2)^{1/2}}\right]$$

Since for a function $f(t)$, $df = \frac{df}{dt}\,dt$

We can rewrite the integrand making use of the following relation

$$d(\mathbf{v} \cdot \mathbf{p}) = \mathbf{v}.\,d\mathbf{p} + d\mathbf{v} \cdot \mathbf{p}$$

Thus
$$T_B - T_A = \int_A^B [d(\mathbf{v} \cdot \mathbf{p}) - \mathbf{p} \cdot d\mathbf{v}]$$

$$= \mathbf{v} \cdot \mathbf{p}\,\Big|_A^B - \int_A^B \mathbf{p} \cdot d\mathbf{v}$$

$$= \frac{m_0 v^2}{(1 - v^2/c^2)^{1/2}}\,\Big|_A^B - \int_A^B \frac{m_0\,\mathbf{v} \cdot d\mathbf{v}}{(1 - v^2/c^2)^{1/2}}$$

Now
$$d\left(\frac{\mathbf{v} \cdot \mathbf{v}}{2}\right) = \mathbf{v} \cdot \frac{d\mathbf{v}}{2} + \frac{d\mathbf{v} \cdot \mathbf{v}}{2} = \mathbf{v} \cdot d\mathbf{v}$$

Using this relation we get

$$T_B - T_A = \frac{m_0 v^2}{(1 - v^2/c^2)^{1/2}}\,\Big|_A^B - \int_A^B \frac{m_0\,v\,dv}{(1 - v^2/c^2)^{1/2}}$$

$$= \frac{m_0 v^2}{(1 - v^2/c^2)^{1/2}}\,\Big|_A^B + m_0 c^2\,(1 - v^2/c^2)^{1/2}\,\Big|_A^B$$

where the second integral has been solved with a change of variable $\left(1 - \frac{v^2}{c^2}\right) = t$. Let us assume that the particle is at rest at the point A, so that $v_A = 0$ and B is any arbitrary point, at which $v_B = v$. Then, the relativistic kinetic energy of the particle starting from rest is

$$T = \frac{m_0 v^2}{(1 - v^2/c^2)^{1/2}} + m_0 c^2\,(1 - v^2/c^2)^{1/2} - m_0 c^2$$

$$= \frac{m_0 v^2 + m_0 c^2 (1 - v^2/c^2)}{(1 - v^2/c^2)^{1/2}} - m_0 c^2$$

$$= \frac{m_0 c^2}{(1 - v^2/c^2)^{1/2}} - m_0 c^2$$

or

$$T = mc^2 - m_0 c^2 \qquad (3.18a)$$

You can see that in the limit $\frac{v}{c} << 1$, this tends to the classical expression of kinetic energy: $T = \frac{1}{2} m_0 v^2$. Why don't you work it out quickly?

We will now discuss the interpretation of Eq.(3.18a) given by Einstein.

3.3.1 The Equivalence of Mass and Energy

The kinetic energy T (Eq. 3.18a) of a particle arises from the work done on it to bring it to speed v from rest. Suppose we rewrite (Eq. 3.18a) as

$$mc^2 = T + m_0c^2 \tag{3.18b}$$

= work done on the particle to change its speed + $m_0 c^2$

Einstein proposed a bold interpretation for Eq. (3.18b) as follows:

mc^2 is the total energy E of the particle.

The first term on the RHS in Eq. (3.18b) arises from (external) work done on the particle by the net external force. The second term, m_0c^2, represents the "rest" energy the particle possessed by virtue of its mass. Thus,

$$E = mc^2 \tag{3.19}$$

Eq. (3.19) does not state a mathematical relation between energy and mass. It states that energy and 'mass' are equivalent concepts. Energy is mass and mass is energy–except for the constant factor c^2. In fact, nowadays, it is common to express the mass of a particle in energy units.

Let us further understand the meaning of Eq. (3.19).

Eq. (3.19) states that mass and energy are equivalent – mass and energy are different names for the same quantity which we can call *mass-energy*. Anything that has a mass m has an energy E (= mc^2) and anything that has an energy E has a mass (= E/c^2). If energy ΔE is added to or taken away from a body, its mass will change by $\Delta m = \Delta E/c^2$, irrespective of the form of energy:

$$\Delta m = \Delta E/c^2 \tag{3.20}$$

ΔE could represent mechanical work, heat energy, absorption of light, or any other form of energy. Thus, Einstein's generalisation of the conservation of energy goes far beyond the classical conservation law of mechanical energy.

One of the most significant consequences of mass-energy equivalence is that the law of conservation of total mass of a system (discussed in Sec. 3.2) automatically implies the law of conservation of total energy of the system. These two classical laws then merge into one **relativistic law of conservation of mass-energy given by Eq. (3.19).**

This single principle embodying the equivalence of mass and energy has had a profound influence on our understanding of the universe. Perhaps its most dramatic examples are provided by the processes of nuclear fission and fusion. While on the one hand these phenomena hold the promise of unlimited energy supply, on the other the spectre of destruction looms large over us.

The atomic bombs responsible for the destruction of Hiroshima and Nagasaki involved the annihilation of a few grams of matter. In this sense, special relativity is not just the concern of scientists alone; it may well control the destiny of all of us.

The experimental evidence for Eqs. (3.19) and (3.20) first came from the mass defect of atomic nuclei. This quantity represents the deficit in the mass of a nucleus (compared with the sum of the masses of its constituent nucleons). It corresponds to the binding energy of the nucleus. This much energy must be supplied to the nucleus before it can break up into individual nucleons; conversely when a nucleus is formed from individual nucleons an equivalent amount of energy must be released. Extending this idea to a nuclear reaction, the net energy released or absorbed would be equivalent to the difference between the total mass defect of the reactants and products. The first direct experimental verification of mass-energy equivalence was done for the following nuclear reaction

$$_3Li^7 + {}_1H^1 \longrightarrow 2\ {}_2He^4$$

It involved a mass difference of

$$(7.0166 + 1.0076)\ \text{amu} - 2 \times 4.0028\ \text{amu} = 0.0186\ \text{amu}$$

with an equivalent energy difference of 27.7×10^{-6}erg. Measurements of the difference between the total kinetic energy of the α-particles produced and that of the incident proton gave the value $(27.6 \pm 0.05) \times 10^{-6}$erg. This was in excellent agreement with the theoretical result. Many experiments have now fully confirmed Eq. (3.19).

We will now make use of Eqs. (3.13a) and (3.19) to express the total energy of a free particle in terms of its momentum and arrive at an interesting result about massless particles.

3.3.2 Relativistic Energy and Momentum of a Free Particle

Classically, the relation between energy and momentum of a free particle is

$$E = \frac{1}{2}mv^2 = \frac{p^2}{2m}$$

For a relativistic free particle, we combine Eqs. (3.13a) and (3.19) to arrive at a corresponding relation. Squaring Eq. (3.13a) we get

$$p^2 = \frac{m_0^2 v^2}{1 - v^2/c^2}$$

whence simple algebra yields

$$p^2\left(1 - \frac{v^2}{c^2}\right) = m_0^2 \frac{v^2}{c^2} \cdot c^2$$

or

$$\frac{v^2}{c^2}\left(m_0^2 c^2 + p^2\right) = p^2$$

Therefore,

$$\frac{v^2}{c^2} = \frac{p^2}{p^2 + m_0^2 c^2}$$

$$\gamma = \frac{1}{(1 - v^2/c^2)^{1/2}} = \left(1 + \frac{p^2}{m_0^2 c^2}\right)^{1/2}$$

Now from Eq. (3.19)

$$E = mc^2 = \gamma m_0 c^2$$

$$= m_0 c^2 \left(1 + \frac{p^2}{m_0^2 c^2}\right)^{1/2}$$

or

$$E^2 = p^2c^2 + (m_0 c^2)^2 \tag{3.21}$$

Eq. (3.21) is the **relativistic energy-momentum relation.** A consequence of this relation is the possibility of 'massless particles'–particles which possess energy and momentum but no rest mass. Let us discuss this briefly.

Massless Particles

If we take $m_0 = 0$ in Eq. (3.21) we get

$$E = pc \tag{3.22}$$

We take the positive root on the assumption that particles whose energy decreases with increasing momentum would be unstable. Now in order to have non-zero momentum we must have a finite value for

$$p = \frac{m_0 v}{(1 - v^2/c^2)^{1/2}}$$

in the limit $m_0 \to 0$. This is only possible if $v \to c$ as $m_0 \to 0$. Thus, massless particles must travel at the speed of light. A well known example of particles which travel at the speed of light is that of *photons*. Thus we conclude that photons have zero rest mass and possess energy given by Eq. (3.22). Hence the momentum of a massless particle like the photon is given by

$$p = \frac{E}{c} \tag{3.23}$$

Conversely, particles which travel at the speed of light, such as photons, have zero rest mass.

We end this discussion with an exercise for you.

SAQ 6

(a) The rest mass of a free proton is 938 MeV/c^2. A proton has kinetic energy 250 MeV. Calculate the total energy (in MeV), relativistic mass, momentum and speed of the proton. It is given that 1 MeV = 1.6×10^{-13} J.

(b) A gamma ray of energy E_γ strikes a proton at rest in the laboratory. What is the momentum of the gamma ray in the laboratory frame? Show that the speed V of the centre of mass in the centre of mass frame is given by

$$V = \frac{E_\gamma c}{E_\gamma + m_p c^2}$$

where m_p is the proton rest mass.

Let us now sum up what you have studied in this unit.

3.4 SUMMARY

- The **relativistic linear momentum** of a particle of rest mass m_0 moving with velocity v is defined by

$$\mathbf{p} = m\mathbf{v} = \frac{m_0 \mathbf{v}}{(1 - v^2/c^2)^{1/2}}$$

where m is the **relativistic mass** of the particle defined by

$$m = \frac{m_0}{(1 - v^2/c^2)^{1/2}}$$

- The **relativistic force law** takes the following form:

$$\mathbf{F} = \frac{d\mathbf{p}}{dt}$$

where p is the relativistic linear momentum.

- If the net external force acting on a body or a system of particles is zero, the linear momentum of the system is conserved.

- If a force F acts on the system for a given time, it produces a change (ΔP) in the momentum of the system, equal to its impulse.

$$\Delta P = J = \int F \, dt$$

- The relativistic total energy of a particle is the sum of its **relativistic kinetic energy** and **rest–mass energy**

$$E = T + m_0 c^2$$

where the relativistic kinetic energy is given by

$$T = mc^2 - m_0 c^2$$

This yields an **equivalence between mass and energy** represented by the equation

$$E = mc^2$$

- The energy and momentum of a **relativistic free particle** are related by

$$E^2 = p^2 c^2 + m_0^2 c^4$$

For a particle of zero rest mass, such as the photon, the relativistic energy-momentum relation becomes

$$E = pc$$

3.5 TERMINAL QUESTIONS

Spend 45 min

1. In an inelastic collision two identical point particles A and B, each of rest mass m_0, collide and stick together forming a single body C of rest mass M_0. Each body has kinetic energy K as seen by an inertial observer. The situation before and after the collision is shown in the Fig. 3.3 below for a frame S'.

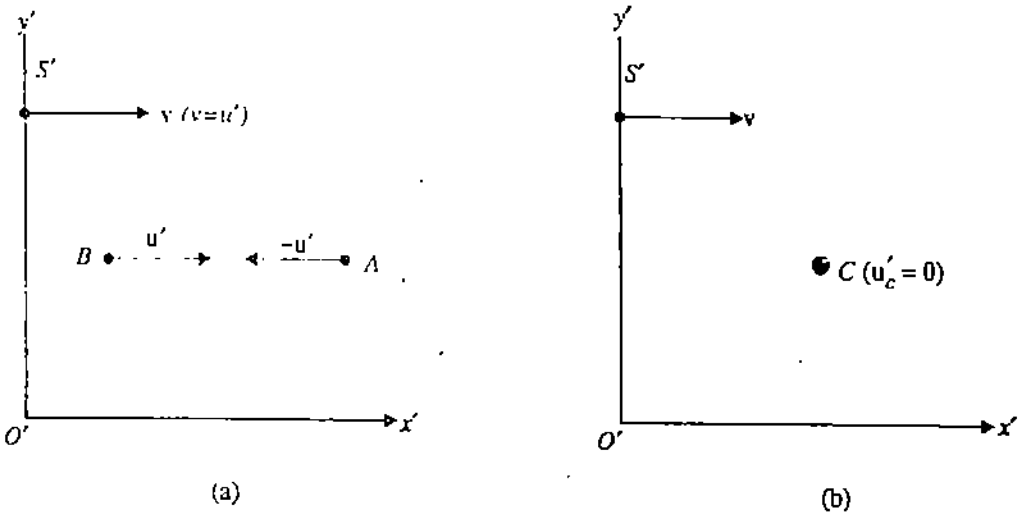


Fig.3.3 : Inelastic collision of two bodies of equal rest mass: (a) before collision; (b) after collision

Consider another reference frame S, with respect to which S' is moving with a velocity v ($v = u'$).

(a) Draw the same collision process as viewed by S before the collision and after the collision. Apply the law of conservation of momentum in frame S and show that the rest mass of the body C is given by

$$M_0 = \frac{2m_0}{(1 - u'^2/c^2)^{1/2}}$$

(b) In S' frame the total kinetic energy before the collision disappears after collision. Show that it equals the increase in the energy of the system which is equivalent to the increase in the rest mass of C.

(c) Show that in both frames S and S', the total energy and relativistic mass are conserved for this inelastic collision.

2. Use the relation $E = mc^2 = (T + m_0c^2)$ to express the relativistic force law given by Eq. (3.14b) as

$$\mathbf{F} = m\frac{d\mathbf{v}}{dt} + \frac{\mathbf{v}\,(\mathbf{F}.\mathbf{v})}{c^2}$$

Hence, obtain the force law for the cases when (i) F is parallel to v and (ii) F is perpendicular to v.

3. A particle of rest mass 5 kg has an initial speed of $2 \times 10^8 ms^{-1}$. It is acted upon by a constant relativistic force of magnitude 10^6N in the same direction as the initial relativistic momentum for 10^3s. Calculate

(a) the magnitude of the initial relativistic momentum
(b) the magnitude of the final relativistic momentum
(c) the final speed of the particle.

4. A proton has $\beta = 0.999$ in the laboratory. Calculate its energy and momentum in an inertial frame S' travelling in the same direction as the proton, with $\beta = 0.99$ with respect to the laboratory. Proton rest energy = 938MeV where 1 MeV = 1.6×10^{-13}J.

5. (a) Use the velocity addition formulae and the result of terminal question 5 of Unit 2 to obtain the following relativistic transformation equations for momentum and energy:

$$p_x' = \gamma(p_x - \beta E/c) \qquad p_x = \gamma(p_x' + \beta' E'/c)$$
$$p_y' = p_y \qquad p_y = p_y'$$
$$p_z' = p_z \qquad p_z = p_z'$$
$$E' = \gamma(E - \beta c p_x) \qquad E = \gamma(E' + \beta c p_x')$$

where the symbols have their usual meanings.

(b) Hence show that the quantity $(E^2 - c^2p^2)$ is Lorentz invariant.

3.6 SOLUTIONS AND ANSWERS

Self Assessment Questions (SAQs)

1. (a) Law of conservation of linear momentum for the collision of the particles

$$\mathbf{p}_{1B} + \mathbf{p}_{2B} = \mathbf{p}_{1A} + \mathbf{p}_{2A}$$

or

$$m_{1B}\ \mathbf{v}_{1B} + m_{2B}\ \mathbf{v}_{2B} = m_{1A}\ \mathbf{v}_{1A} + m_{2A}\ \mathbf{v}_{2A} \quad (\text{since } \mathbf{p} = m\mathbf{v}) \qquad (1)$$

(b) The law of conservation of momentum in S' is

$$\mathbf{p}'_{1B} + \mathbf{p}'_{2B} = \mathbf{p}'_{1A} + \mathbf{p}'_{2A}$$

or

$$m'_{1B}\ \mathbf{v}'_{1B} + m'_{2B}\ \mathbf{v}'_{2B} = m'_{1A}\ \mathbf{v}'_{1A} + m'_{2A}\ \mathbf{v}'_{2A} \qquad (2)$$

where m'_{1B}, m'_{2B} are moving with velocities $\mathbf{v}'_{1B}$ and $\mathbf{v}'_{2B}$ before collision and m'_{1A}, m'_{2A} are moving with velocities $\mathbf{v}'_{1A}$ and $\mathbf{v}'_{2A}$ after collision. It is given that $m_{1B} = m'_{1B}, m_{2B} = m'_{2B}, m_{1A} = m'_{1A}, m_{2A} = m'_{2A}$. Using the special transformation law

$$\mathbf{v} = \frac{\mathbf{v}' + \mathbf{u}}{1 + \mathbf{v}'\mathbf{u}/c^2}$$

We can write Eq. (1) as

$$m_{1B}\left[\frac{\mathbf{v}'_{1B} + \mathbf{u}}{1 + \mathbf{v}'_{1B}\,\mathbf{u}/c^2}\right] + m_{2B}\left[\frac{\mathbf{v}'_{2B} + \mathbf{u}}{1 + \mathbf{v}'_{2B}\,\mathbf{u}/c^2}\right] = m_{1A}\left[\frac{\mathbf{v}'_{1A} + \mathbf{u}}{1 + \mathbf{v}'_{1A}\,\mathbf{u}/c^2}\right] + m_{2A}\left[\frac{\mathbf{v}'_{2A} + \mathbf{u}}{1 + \mathbf{v}'_{2A}\,\mathbf{u}/c^2}\right] \qquad (3)$$

Clearly this equation cannot reduce to Eq. (2) and hence conservation of momentum does not hold in S'.

2. We have to show that

$$\left(1 - v_i'^2/c^2\right)\left(1 - \frac{v_i u}{c^2}\right)^2 = \left(1 - \frac{v_i^2}{c^2}\right)\left(1 - \frac{u^2}{c^2}\right)$$

From Eq. (3.4a)

$$1 - \frac{v_i'^2}{c^2} = 1 - \frac{1}{c^2}\left(\frac{v_i - u}{1 - v_i u/c^2}\right)^2$$

$$= \frac{c^2(1 - v_i u/c^2)^2 - (v_i - u)^2}{c^2(1 - v_i u/c^2)^2}$$

$$(1 - v_i'^2/c^2)\left(1 - \frac{v_i u}{c^2}\right)^2 = \left(1 - \frac{v_i u}{c^2}\right)^2 - \frac{1}{c^2}(v_i - u)^2$$

$$= 1 + \frac{v_i^2 u^2}{c^4} - \frac{2v_i u}{c^2} - \frac{v_i^2}{c^2} - \frac{u^2}{c^2} + \frac{2v_i u}{c^2}$$

$$= 1 - \frac{v_i^2}{c^2} - \frac{u^2}{c^2} + \frac{v_i^2 u^2}{c^4}$$

$$= 1 - \frac{v_i^2}{c^2}\left(1 - \frac{u^2}{c^2}\right) - \frac{u^2}{c^2}$$

$$= \left(1 - \frac{u^2}{c^2}\right)\left(1 - \frac{v_i^2}{c^2}\right)$$

3. (a) We start by working from Eq. (3) in the answer to SAQ 1(b) given as

$$m_{1B}\left[\frac{\mathbf{v}'_{1B} + \mathbf{u}}{1 + v'_{1B}\,u/c^2}\right] + m_{2B}\left[\frac{\mathbf{v}'_{2B} + \mathbf{u}}{1 + v'_{2B}\,u/c^2}\right] = m_{1A}\left[\frac{\mathbf{v}'_{1A} + \mathbf{u}}{1 + v'_{1A}\,u/c^2}\right] + m_{1B}\left[\frac{\mathbf{v}'_{2A} + \mathbf{u}}{1 + v'_{2A}\,u/c^2}\right] \quad (1)$$

Using Eq. (3.4b) we can show as in SAQ 2 that

$$\left(1 + \frac{v'_i\,u}{c^2}\right)^2\left(1 - \frac{v_i^2}{c^2}\right) = \left(1 - \frac{v_i'^2}{c^2}\right)\left(1 - \frac{u^2}{c^2}\right) \quad (2)$$

Applying Eq. (2) to this problem we can write

$$\left(1 + \frac{v'_{1A}\,u}{c^2}\right)^2 = \frac{(1 - v_{1A}'^2/c^2)}{(1 - v_{1A}^2/c^2)}\quad (1 - u^2/c^2)$$

Thus

$$\frac{m_{1A}}{1 + v'_{1A}\,u/c^2} = \frac{1}{(1 - u^2/c^2)^{1/2}}\left(\frac{1 - v_{1A}^2/c^2}{1 - v_{1A}'^2/c^2}\right)^{1/2} m_{1A}$$

$$= \left(\frac{1}{1 - u^2/c^2}\right)^{1/2} \frac{m_{1A}\,\gamma'_{1A}}{\gamma_{1A}}$$

$$= \frac{m'_{1A}}{(1 - u^2/c^2)^{1/2}} \quad \text{using Eq. (3.10)}$$

Similarly

$$\frac{m_{2A}}{(1 + v'_{2A}\,u/c^2)} = \frac{m'_{2A}}{(1 - u^2/c^2)^{1/2}},$$

$$\frac{m_{1B}}{(1 + v'_{1B}\,u/c^2)} = \frac{m'_{1B}}{(1 - u^2/c^2)^{1/2}}$$

$$\frac{m_{2B}}{(1 + v'_{2B}\,u/c^2)} = \frac{m'_{2B}}{(1 - u^2/c^2)^{1/2}}$$

Thus Eq. (1) becomes

$$\frac{m'_{1B}\,(\mathbf{v}'_{1B} + \mathbf{u})}{(1 - u^2/c^2)^{1/2}} + \frac{m'_{2B}\,(\mathbf{v}'_{2B} + \mathbf{u})}{(1 - u^2/c^2)^{1/2}} = \frac{m'_{1A}\,(\mathbf{v}'_{1A} + \mathbf{u})}{(1 - u^2/c^2)^{1/2}} + \frac{m'_{2A}\,(\mathbf{v}'_{2A} + \mathbf{u})}{(1 - u^2/c^2)^{1/2}}$$

or

$$\frac{m'_{1B}\,v'_{1B}}{(1-u^2/c^2)^{1/2}} + \frac{m'_{2B}\,v'_{2B}}{(1-u^2/c^2)^{1/2}} + \frac{(m'_{1B}+m'_{2B})\,u}{(1-u^2/c^2)^{1/2}}$$

$$= \frac{m'_{1A}\,v'_{1A}}{(1-u^2/c^2)^{1/2}} + \frac{m'_{2A}\,v'_{2A}}{(1-u^2/c^2)^{1/2}} + \frac{(m'_{1A}+m'_{2A})\,u}{(1-u^2/c^2)^{1/2}} \qquad (3)$$

Since the total mass of the system is conserved in the collision process,

$$m'_{1A} + m'_{2A} = m'_{1B} + m'_{2B} \qquad (4)$$

And using Eq. (3.13a) alongwith Eq. (4), we can write Eq. (3) as

$$\mathbf{p}'_{1B} + \mathbf{p}'_{2B} = \mathbf{p}'_{1A} + \mathbf{p}'_{2A}$$

Thus linear momentum is conserved in S' frame.

b) From Eq. (3.13a) we have

$$\mathbf{p} = \frac{m_0\,\mathbf{v}}{(1-v^2/c^2)^{1/2}}$$

or the magnitude $p = \dfrac{m_0\,v}{(1-v^2/c^2)^{1/2}}$ where $v = 0.99c$ and $p = 1.92 \times 10^{-21}$kg m s^{-1}.

Thus $$m_0 = \frac{1.92 \times 10^{-21}\ \text{kg m s}^{-1}}{0.99 \times 3.0 \times 10^{8}\ \text{m s}^{-1}}\ [1-(0.99)^2]^{1/2}$$

$$= 9.1 \times 10^{-31}\text{kg}.$$

4. It is given that the acceleration is parallel to the velocity of the proton. Now

$$\mathbf{F} = m\mathbf{a} + \mathbf{v}\frac{dm}{dt}$$

where

$$\frac{dm}{dt} = m_0 \frac{d}{dt}\left[\frac{1}{(1-v^2/c^2)^{1/2}}\right] = m_0 \frac{v}{c^2(1-v^2/c^2)^{3/2}}\frac{dv}{dt}$$

Thus, $$\mathbf{F} = m\mathbf{a} + \frac{m_0}{(1-v^2/c^2)^{3/2}}\frac{v^2}{c^2}\hat{\mathbf{v}}\frac{dv}{dt}$$

$$= m\mathbf{a} + \frac{m_0}{(1-v^2/c^2)^{3/2}}\frac{v^2}{c^2}\mathbf{a} \quad [\text{Since } \mathbf{a} \parallel \mathbf{v},\ \mathbf{a} = \hat{\mathbf{v}}\frac{dv}{dt}]$$

$$= m\left[1 + \frac{v^2/c^2}{1-v^2/c^2}\right]\mathbf{a} \quad [\text{since } m = \frac{m_0}{(1-v^2/c^2)^{1/2}}]$$

$$= \frac{m_0}{(1-v^2/c^2)^{3/2}}\mathbf{a} = \gamma^3 m_0\,\mathbf{a}$$

It is given that $m_0 = 1.67 \times 10^{-27}$kg, $a = 1.0 \times 10^{15}$ ms^{-2}.

Since $v = 0.9c$, $\gamma = 2.3$

Thus $F = \gamma^3 m_0 a = 2.03 \times 10^{-11}$N.

5. In the classical limit $v/c << 1$ we make the approximation $\dfrac{1}{(1-v^2/c^2)^{1/2}} \approx 1 + \dfrac{1}{2}v^2/c^2$. Then

$$T = \frac{m_0c^2}{(1-v^2/c^2)^{1/2}} - m_0c^2 = m_0c^2\left(1 + \frac{1}{2}\frac{v^2}{c^2} - 1\right) = \frac{1}{2}m_0v^2$$

6. a) For the proton

$$m_0c^2 = 938\ \text{MeV}$$

and

$$T = 200 \text{ MeV}$$

$\therefore$ Total energy $E = T + m_0c^2 = (200 + 938)$ MeV $= 1138$ MeV

$$\text{Rest mass} \quad m_0 = \frac{938 \times 1.6 \times 10^{-13} \text{ J}}{9 \times 10^{16} \text{ m}^2 \text{ s}^{-2}}$$

$$= 1.67 \times 10^{-27} \text{ kg}$$

$$\text{Relativistic mass } m = \frac{E}{c^2}$$

$$= \frac{1138 \times 1.6 \times 10^{-13} \text{ J}}{9 \times 10^{16} \text{ m}^2 \text{ s}^{-2}} = 2.02 \times 10^{-27} \text{ kg}$$

Momentum of the proton is obtained from Eq. (3.21)

$$E^2 = p^2c^2 + (m_0c^2)^2$$

$$m^2c^4 = p^2c^2 + m_0^2c^4$$

$$p^2 = m^2c^2 - m_0^2c^2$$

or

$$p^2 = m.(mc^2) - m_0(m_0c^2)$$

$$= (2.02 \times 10^{-27}\text{kg} \times 1138 - 938 \times 1.67 \times 10^{-27}\text{kg}) \times 1.6 \times 10^{-13}\text{J}$$

$$= (2.02 \times 1138 - 1.67 \times 938) \times 1.6 \times 10^{-40}\text{J kg}$$

$$= (2298 - 1566) \times 1.6 \times 10^{-40}\text{J kg}$$

$$= 1171 \times 10^{-40}\text{ J kg}$$

or

$$p = 3.4 \times 10^{-19}\text{kg m s}^{-1}$$

The speed of the proton

$$v^2 = \frac{p^2c^2}{p^2 + m_0c^2} = \frac{1171 \times 10^{-40} \times 9 \times 10^{16}}{1171 \times 10^{-40} + 2506 \times 10^{-40}} \text{ m}^2 \text{ s}^{-2}$$

or

$$v^2 = \frac{10539 \times 10^{16}}{3677} \text{ m}^2 \text{ s}^{-2}$$

$$= 2.87 \times 10^{16} \text{ m}^2 \text{ s}^{-2}$$

or $$v = 1.69 \times 10^8 \text{ m s}^{-1}$$

b) The momentum of the γ-ray in the laboratory frame is

$$p_\gamma = \frac{E_\gamma}{c}$$

The velocity of the centre-of-mass in the laboratory frame is given by

$$V = \frac{M_p v_p + M_\gamma v_\gamma}{M_p + M_\gamma}$$

Since the proton is at rest, $v_p = 0$

$$\therefore \quad V = \frac{M_\gamma v_\gamma}{M_p + M_\gamma}$$

or $$V = \frac{M_\gamma v_\gamma}{M_p + M_\gamma}$$

Now $M_\gamma v_\gamma = p_\gamma = \frac{E_\gamma}{c}$. Since gamma rays travel at the speed of light

$$v_\gamma = c \text{ and } M_\gamma = \frac{E_\gamma}{c^2}$$

Thus

$$V = \frac{E_\gamma / c}{M_p + E_\gamma / c^2} = \frac{E_\gamma c}{M_p c^2 + E_\gamma}$$

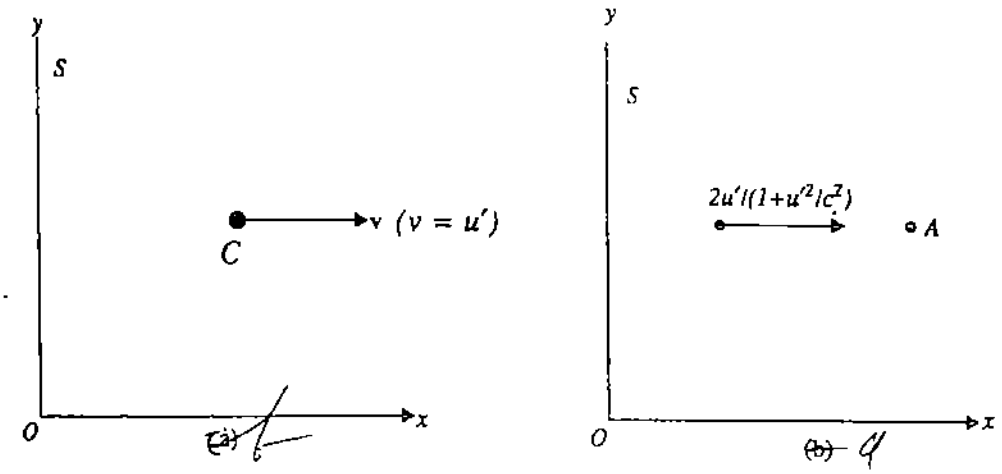
Terminal Questions

1. Using the velocity addition formula (Eqs. 2.20a to d) we can write the velocity of A and B in S as:

$$u_A = \frac{-u' + v}{1 - u' v/c^2} = 0$$

$$u_B = \frac{u' + v}{1 + u' v/c^2} = \frac{u' + u'}{1 + u'^2/c^2} = \frac{2u'}{1 + u'^2/c^2}$$

Since C is at rest in S', in S it will travel with the same velocity v ($v = u'$) as that of S'. Thus this collision process can be shown pictorially as follows.



The relativistic mass of B in S frame is

$$m_B = \frac{m_0}{(1 - u_B^2/c^2)^{1/2}}$$

$$= \frac{m_0}{\left[1 - \frac{4u'^2/c^2}{(1 + u'^2/c^2)^2}\right]^{1/2}}$$

$$= \frac{m_0 (1 + u'^2/c^2)}{(1 - u'^2/c^2)}$$

Now let us apply the law of conservation of momentum in frame S to its x-component (the y-component is automatically conserved):

$$p_{before} = p_{after}$$

$$m_B u_B + m_A u_A = M_C v$$

or

$$\frac{m_0 (1 + u'^2/c^2)}{(1 - u'^2/c^2)} \quad \frac{2u'}{(1 + u'^2/c^2)} + 0 = \frac{M_0 u'}{(1 - u'^2/c^2)^{1/2}}$$

since

$$M_C = \frac{M_0}{(1 - u'^2/c^2)^{1/2}}$$

where M_0 is the rest mass of C and $v = u'$.

or $$M_0 = \frac{2m_0}{(1 - u'^2/c^2)^{1/2}}$$

(b) Kinetic energy in the S' frame before collision is given by Eq. (3.18a).

$$K_A' = m_A' c^2 - m_0 c^2 = m_0 c^2 \left[\frac{1}{(1 - u'^2/c^2)^{1/2}} - 1\right]$$

$$K_B' = m_B' c^2 - m_0 c^2 = m_0 c^2 \left[\frac{1}{(1 - u'^2/c^2)^{1/2}} - 1\right]$$

$$\text{Total K.E.} = K_A' + K_B' = 2m_0 c^2 \left[\frac{1}{(1 - u'^2/c^2)^{1/2}} - 1\right] \quad (1)$$

Now the rest mass of C is greater than the sum of the rest masses of A and B by an amount $\Delta m = M_0 - 2m_0 = 2m_0\left[\frac{1}{(1 - u'^2/c^2)^{1/2}} - 1\right]$. The increase in energy equivalent to the increase in mass $= \Delta m c^2$

$$= 2m_0 c^2 \left[\frac{1}{\left(1 - \frac{u'^2}{c^2}\right)^{1/2}} - 1\right]$$

which is the same as (1). Therefore, the disappeared K.E. appears as an increase in the rest mass of the system. And the total mass-energy of the system is conserved.

(c) Consider the S frame:
The total energy (Eq. 3.18b) is

(i) Before collision: $E_A + E_B = m_0 c^2 + (m_0 c^2 + K_B)$ since $K_A = 0$

$$= 2m_0 c^2 + m_0 c^2 \left[\frac{1}{(1 - u_B^2/c^2)^{1/2}} - 1\right] \quad [\because K_B = m_B c^2 - m_0 c^2]$$

$$= 2m_0 c^2 + m_0 c^2 \left[\frac{1 + u'^2/c^2}{1 - u'^2/c^2} - 1\right]$$

$$= 2m_0 c^2 + m_0 c^2 \left[\frac{2u'^2/c^2}{1 - u'^2/c^2}\right]$$

$$= \frac{2m_0 c^2}{1 - u'^2/c^2}$$

(ii) After collision : $M_C c^2 = \frac{M_0 c^2}{(1 - u'^2/c^2)^{1/2}} = \frac{2m_0 c^2}{1 - u'^2/c^2}$

where we have substituted the value of M_0.

Thus the total energy in S frame is conserved. Now consider the S' frame. The total energy using (1) is

(i) Before collision: $m_0 c^2 + K_A' + m_0 c^2 + K_B' = 2m_0 c^2 + 2m_0 c^2 \left[\frac{1}{(1 - u'^2/c^2)^{1/2}} - 1\right]$

$$= \frac{2m_0 c^2}{(1 - u'^2/c^2)^{1/2}}$$

(ii) After collision: $M_0 c^2 = \frac{2m_0 c^2}{(1 - u'^2/c^2)^{1/2}}$

Hence the total energy in S' frame is also conserved.

S frame

(i) Before collision: $m_A + m_B = m_0 + \dfrac{m_0}{(1 - u_B^2/c^2)^{1/2}} = m_0 + \dfrac{m_0(1 + u'^2/c^2)}{(1 - u'^2/c^2)}$

$$= \frac{2m_0}{(1 - u'^2/c^2)}$$

(ii) After collision: $M_C = \dfrac{M_0}{(1 - u'^2/c^2)^{1/2}} = \dfrac{2m_0}{(1 - u'^2/c^2)}$

S' frame

(i) Before collision : $m_A' + m_B' = \dfrac{m_0}{(1 - u'^2/c^2)^{1/2}} + \dfrac{m_0}{(1 - u'^2/c^2)^{1/2}} = \dfrac{2m_0}{(1 - u'^2/c^2)^{1/2}}$

(ii) After collision: $M_0 = \dfrac{2m_0}{(1 - u'^2/c^2)^{1/2}}$

Hence the relativistic mass is also conserved in both S and S' frames.

2. The relativistic force law is

$$\mathbf{F} = m\frac{d\mathbf{v}}{dt} + \mathbf{v}\frac{dm}{dt}$$

Now $m = \dfrac{E}{c^2}$ so that

$$\frac{dm}{dt} = \frac{1}{c^2}\frac{dE}{dt} = \frac{1}{c^2}\frac{d}{dt}(T + m_0c^2) = \frac{1}{c^2}\frac{dT}{dt}$$

But $\dfrac{dT}{dt} = \dfrac{\mathbf{F} \cdot d\mathbf{l}}{dt} = \mathbf{F} \cdot \dfrac{d\mathbf{l}}{dt} = \mathbf{F}.\mathbf{v}$

Thus $\dfrac{dm}{dt} = \dfrac{1}{c^2}\mathbf{F}.\mathbf{v}$

With this the force law becomes

$$\mathbf{F} = m\frac{d\mathbf{v}}{dt} + \mathbf{v}\frac{(\mathbf{F} \cdot \mathbf{v})}{c^2}$$

(i) When F is parallel to v,

$$\mathbf{F} = m\mathbf{a} + \hat{\mathbf{v}}\frac{Fv^2}{c^2} \quad \text{where } \mathbf{a} = \frac{d\mathbf{v}}{dt} \text{ and } \mathbf{v} = v\hat{\mathbf{v}}$$

or $$\mathbf{F}\left(1 - \frac{v^2}{c^2}\right) = m\mathbf{a} \qquad [\because \mathbf{v} \parallel \mathbf{F}]$$

or $$\mathbf{F} = \frac{m}{(1 - v^2/c^2)}\mathbf{a}$$

$$\mathbf{F} = \frac{m_0}{(1 - v^2/c^2)^{3/2}}\mathbf{a}$$

Thus it turns out that both F and a are parallel to v. The quantity $\dfrac{m_0}{(1 - v^2/c^2)^{3/2}}$ is sometimes called "longitudinal mass".

(ii) When F is perpendicular to v, $\mathbf{F} \cdot \mathbf{v} = 0$ and

$$\mathbf{F} = m\frac{d\mathbf{v}}{dt}$$

$$= \frac{m_0}{(1 - v^2/c^2)^{1/2}} \text{ a}$$

The quantity $\frac{m_0}{(1-v^2/c^2)^{1/2}}$ is also called the "transverse mass".

3. a) The magnitude of the initial relativistic momentum is

$$p_i = \frac{m_0 v_i}{(1 - v_i^2/c^2)^{1/2}} = \frac{5\text{ kg} \times 2 \times 10^8\text{ m s}^{-1}}{(1 - 4/9)^{1/2}} = = 1.34 \times 10^9 \text{kg m s}^{-1}$$

(b) The change in the magnitude of relativistic momentum is given by $\Delta \mathbf{P} = \int \mathbf{F}\, dt = \mathbf{F} \int dt$ as the force is constant

$$\therefore \Delta P = F\ \Delta\ t = 10^6 \text{N} \times 10^3 \text{s} = 10^9 \text{kg m s}^{-1}$$

Hence the magnitude of the final relativistic momentum is

$$p_f = (1.34 \times 10^9 + 10^9)\text{ kg m s}^{-1}$$
$$= 2.34 \times 10^9\text{ kg m s}^{-1}$$

(c) The final speed can be obtained from

$$p_f = \frac{m_0\ v_f}{(1 - v_f^2/c^2)^{1/2}}$$

or
$$p_f^2 = \frac{m_0^2\ v_f^2}{1 - v_f^2/c^2}$$

or
$$p_f^2 \left(1 - \frac{v_f^2}{c^2}\right) = m_0^2\ v_f^2$$

or
$$v_f^2 \left(m_0^2 + \frac{p_f^2}{c^2}\right) = p_f^2$$

$$v_f = \frac{p_f}{(m_0^2 + p_f^2/c^2)^{1/2}}$$

$$= \frac{2.34 \times 10^9\text{ kg m s}^{-1}}{(25\text{ kg}^2 + 60.8\text{ kg}^2)^{1/2}}$$

$$= 2.53 \times 10^8\text{ m s}^{-1}$$

4. We first calculate the velocity v'' of the proton in the S'' frame. Using the velocity transformation relation

$$v'' = \frac{v - v'}{1 - vv'/c^2}$$

or

$$\frac{v''}{c} = \frac{\frac{v}{c} - \frac{v'}{c}}{1 - vv'/c^2}$$

or

$$\beta'' = \frac{\beta - \beta'}{1 - \beta\beta'}$$

$$= \frac{0.999 - 0.990}{1 - (0.999)\ 0.990)} = \frac{0.009}{0.011} = 0.82$$

Hence $E = \frac{m_0 c^2}{(1 - \beta''^2)^{1/2}} = \frac{938\text{ MeV}}{(1 - (0.82)^2)^{1/2}} = \frac{938}{0.572} = 1640\text{ MeV}$

$$p^2c^2 = E^2 - m_0^2 c^4$$

or $$p = \frac{1}{c}(E^2 - m_0^2 c^4)^{1/2}$$

$$= \frac{1.6 \times 10^{-13}}{3 \times 10^8 \text{ m s}^{-1}} \times [(1640)^2 - (938)^2]^{1/2} \text{ J}$$

$$= \frac{1.6 \times 10^{-13} \times 1345}{3 \times 10^8} \text{ kg m s}^{-1}$$

$$= 7.17 \times 10^{-19} \text{ kg m s}^{-1}$$

5. a) In frame S, the momentum and energy equations are

$$p_x = \frac{m_0 v_x}{(1 - v^2/c^2)^{1/2}}, \quad p_y = \frac{m_0 v_y}{(1 - v^2/c^2)^{1/2}}, \quad p_z = \frac{m_0 v_z}{(1 - v^2/c^2)^{1/2}},$$

$$E = \frac{m_0 c^2}{(1 - v^2/c^2)^{1/2}}$$

In frame S', the corresponding quantities are by definition

$$p'_x = \frac{m_0 v'_x}{(1 - v'^2/c^2)^{1/2}}, \quad p'_y = \frac{m_0 v'_y}{(1 - v'^2/c^2)^{1/2}}, \quad p'_z = \frac{m_0 v'_z}{(1 - v'^2/c^2)^{1/2}}$$

$$E' = \frac{m_0 c^2}{(1 - v'^2/c^2)^{1/2}}$$

The velocity addition equations (2.20a to c) and the result obtained in Terminal Question 5 of Unit 2 are

$$v_x = \frac{v'_x + V}{1 + v'_x V/c^2}, \tag{1a}$$

$$v_y = \frac{v'_y}{\gamma(1 + v'_x V/c^2)}, \quad \gamma = \frac{1}{(1 - V^2/c^2)^{1/2}} \tag{1b}$$

$$v_z = \frac{v'_z}{\gamma(1 + v'_x V/c^2)} \tag{1c}$$

$$(c^2 - v^2) = \frac{c^2 (c^2 - v'^2)(c^2 - V^2)}{(c^2 + v'_x V)^2}$$

$$= \frac{(c^2 - v'^2)(c^2 - V^2)}{c^2 (1 + v'_x V/c^2)^2}$$

If we divide throughout by c^2, invert and take the square root we get

$$\frac{1}{(1 - v^2/c^2)^{1/2}} = \frac{1 + v'_x V/c^2}{(1 - v'^2/c^2)^{1/2}(1 - V^2/c^2)^{1/2}} \tag{2}$$

Now

$$p_x = \frac{m_0 v_x}{(1 - v^2/c^2)^{1/2}} = \frac{m_0 (v'_x + V)}{(1 - v^2/c^2)^{1/2}(1 + v'_x V/c^2)}$$

$$= \frac{m_0 (v'_x + V)}{(1 + v'_x V/c^2)} \cdot \frac{1 + v'_x V/c^2}{(1 - v'^2/c^2)^{1/2}(1 - V^2/c^2)^{1/2}}$$

$$= \frac{m_0 (v'_x + V)}{(1 - v'^2/c^2)^{1/2}(1 - V^2/c^2)^{1/2}}$$

or $$p_x = \frac{p'_x}{(1 - V^2/c^2)^{1/2}} + \frac{m_0 (Vc^2/c^2)}{(1 - v'^2/c^2)^{1/2}(1 - V^2/c^2)^{1/2}}$$

or $$p_x = \frac{p_x' + E' V/c^2}{(1 - V^2/c^2)^{1/2}} = \gamma(p_x' + \beta E'/c)$$

$$p_y = \frac{m_0 v_y}{(1 - v^2/c^2)^{1/2}} = \frac{m_0 v_y' (1 - V^2/c^2)^{1/2}}{(1 - v^2/c^2)^{1/2} (1 + v_x' V/c^2)}$$

$$= \frac{m_0 v_y'}{(1 - v'^2/c^2)^{1/2}} \qquad \text{using Eq. (2)}$$

or $$p_y = p_y'$$

$$p_z = \frac{m_0 v_z}{(1 - v^2/c^2)^{1/2}} = \frac{m_0 v_z' (1 - V^2/c^2)^{1/2}}{(1 - v^2/c^2)^{1/2} (1 + v_x' V/c^2)^{1/2}} = \frac{m_0 v_z'}{(1 - v'^2/c^2)^{1/2}} = p_z'$$

$$E = \frac{m_0 c^2}{(1 - v^2/c^2)^{1/2}} = m_0 c^2 \left[\frac{1 + v_x' V/c^2}{(1 - v'^2/c^2)^{1/2} (1 - V^2/c^2)^{1/2}}\right]$$

$$= \frac{m_0 c^2}{(1 - v'^2/c^2)^{1/2} (1 - V^2/c^2)^{1/2}} + \frac{m_0 v_x' V}{(1 - v'^2/c^2)^{1/2} (1 - V^2/c^2)^{1/2}}$$

$$= \frac{E'}{(1 - V^2/c^2)^{1/2}} + \frac{V p_x'}{(1 - V^2/c^2)^{1/2}}$$

or $$E = \frac{E' + V p_x'}{(1 - V^2/c^2)^{1/2}} = \gamma(E' + \beta c p_x') \qquad (3)$$

The inverse relations are obtained by replacing $p_x \leftrightarrow p_x'$, $p_y \leftrightarrow p_y'$, $p_z \leftrightarrow p_z'$, $E \leftrightarrow E'$ and V by $-V$ in these equations. We tabulate the results below.

Relativistic Transformation for momentum and energy

$p_x' = \frac{p_x - EV/c^2}{(1 - V^2/c^2)^{1/2}}$	$p_x = \frac{p_x' + E'V/c^2}{(1 - V^2/c^2)^{1/2}}$
$p_y' = p_y$	$p_y = p_y'$
$p_z' = p_z$	$p_z = p_z'$
$E' = \frac{E - Vp_x}{(1 - V^2/c^2)^{1/2}}$	$E = \frac{E' + Vp_x'}{(1 - V^2/c^2)^{1/2}}$

(b) To show that $E^2 - c^2p^2$ is Lorentz invariant we have to prove that its value in the inertial frame S' remains unchanged, i.e.,

$$E'^2 - c^2p'^2 = E^2 - c^2p^2$$

Now $E^2 - c^2p^2 = E^2 - c^2p_x^2 - c^2p_y^2 - c^2p_z^2$

Using the result in (3) we have

$$E^2 - c^2p^2 = \gamma^2 (E' + \beta c\ p_x')^2 - c^2\ \gamma^2 (p_x' + \beta E'/c)^2 - c^2 (p_y'^2 + p_z'^2)$$

$$= \gamma^2 (E'^2 + \beta^2 c^2 p_x'^2 + 2E' \beta c\ p_x')$$

$$- c^2 \gamma^2 \left(p_x'^2 + \frac{\beta^2 E'^2}{c^2} + \frac{2\beta}{c} p_x E'\right) - c^2 \left(p_y'^2 + p_z'^2\right)$$

$$= \gamma^2 [E'^2 (1 - \beta^2) - p_x'^2 c^2 (1 - \beta^2)] - c^2 p_y'^2 - c^2 p_z'^2$$

$$= \gamma^2 \left[\frac{E'^2}{\gamma^2} - \frac{p_x'^2 c^2}{\gamma^2}\right] - c^2 p_y'^2 - c^2 p_z'^2$$

$$= E'^2 - p_x'^2c^2 - c^2\ p_y'^2 - c^2\ p_z'^2$$

$$= E'^2 - c^2\ (p_x'^2 + p_y'^2 + p_z'^2)$$

$$= E'^2 - c^2p'^2$$

Thus, the quantity $E^2 - p^2c^2$ is Lorentz invariant. This is nothing but the square of the rest energy.

FURTHER READING

1. Introduction to Special Relativity, R. Resnick, Wiley Eastern Limited, 1989.
2. Mechanics, Berkeley Physics Course – Volume I; C. Kittel, W.D. Knight, M.A. Ruderman, A.C. Helmholz, B.J. Moyer; Asian Student Edition, McGraw-Hill International Book Company, 1981.
3. Uncle Tompkins in Paperback, George Gamow, Cambridge University Press, 1965.

CONCLUSION

How did you like coming to grips with Einstein's special theory of relativity? Did you find your mental abilities being stretched to their limits by its demands? We hope that studying this theory has enriched your understanding of the physical world; and it has helped you throw away your prejudices and rely on your powers of reasoning. Perhaps you have even enjoyed the experience! If so, trust us when we say that it has been extremely stimulating and satisfying to explain its nuances and intricacies to you in this block.

Einstein's special theory of relativity is one of the greatest achievements of modern physics. Its intellectual appeal lies in its logic and internal consistency. Once we accept the two postulates, all else follows inevitably – the changed perception of space and time as well as its implications for the laws of physics.

The Special Theory of Relativity also has a practical value. The principle of relativity and the Lorentz transformation together give us a framework for 'testing' any physical theory even without recourse to experiment. All current and would-be physical theories are acceptable only if they fit into this framework. This makes the theory of relativity one of the most useful tools at the command of theoretical physicists.

And what about the person behind all this? We all know of Einstein's contributions as a scientist. He was no doubt one of the greatest scientific thinkers of all times. But there is more to the man than his contributions to science.

All of humanity takes great pride in him – not only for his scientific discoveries but also for his personal qualities. His courage, modesty, perseverance and sense of fun, and above all his immense contribution to international peace and understanding is an everlasting source of inspiration for all of us. You may like to go through an account of his life and work. Here we have tried to briefly recount his life story through the important events that occurred in his life.

1879 March 14: Albert Einstein is born in Ulm, Germany.

1884 Einstein is given a pocket compass; his first experience of scientific discovery.

1891 At the age of twelve, Einstein is given a geometry book; his second experience of awe and excitement at the "mysteries of nature".

1895 Einstein takes and fails the entrance exam for the Institute of Technology in Zurich, Switzerland. He goes, instead, to Aarau in Switzerland to study.

1896 Aged seventeen, Einstein resigns his German citizenship. He passes the entrance exam to the Institute of Technology in Zurich and begins study.

1900 Einstein graduates from the Institute of Technology and starts to look for a job. His first scientific paper is published.

1905 Einstein completes his Ph.D thesis and publishes several scientific papers, including two on Special Relativity. One contains the famous equation $E = mc^2$.

1906 He writes the first-ever paper on Quantum Mechanics. For almost the next twenty years, he produces a steady stream of new, original scientific ideas, which are published in scientific journals.

1909 Einstein resigns from the Patent Office and is appointed assistant Professor at the University of Zurich.

1911 Einstein is appointed Professor at the University of Prague, Czechoslovakia. He predicts that light will be seen to bend during an eclipse of the sun.

1912 Einstein, now thirty-three, returns to the Institute of Technology at Zurich where he is appointed full professor. Together with Marcel Grossman he works on the Theory of General Relativity.

1915 Einstein signs the "Manifesto to Europe" and calls for a League of Europe to bring about peace.

1919 Astronomical observations confirm Einstein's predictions about how light might appear to bend in space and suddenly he becomes world-famous.

1922 Einstein is appointed a member of the League of Nations Committee on International Co-operation. He is also awarded the Nobel Prize for Physics for 1921.

1925 Einstein undertakes further international lectures and publications. He signs, with Mahatma Gandhi and others, a manifesto calling for the abolition of military conscription.

1929 The Planck medal – one of the highest awards for physics – is given to Einstein.

1930 He signs a manifesto calling for world disarmament.

1932 Einstein is made professor at the Institute for Advanced Study, Princeton in the United States. He leaves Germany for ever.

1940 Einstein becomes a citizen of the United States although he keeps his Swiss nationality, as well.

1946 Einstein urges the United Nations to form a World Government to prevent future wars.

1952 He is offered the Presidency of the State of Israel but refuses.

1955 Despite being very ill, Einstein continues to campaign for the abolition of nuclear weapons and to work on scientific papers.
April 18: Albert Einstein, aged seventy-six, passes away.

NOTES

UTTAR PRADESH
RAJARSHI TANDON OPEN UNIVERSITY

UGPHS-08
Modern Physics

Block

2

AN INTRODUCTION TO QUANTUM MECHANICS

UNIT 4
Wave-Particle Duality 5

UNIT 5
Matter Waves and Uncertainty Principle 21

UNIT 6
Schrödinger Equation 36

UNIT 7
Observables and Operators 52

AN INTRODUCTION TO QUANTUM MECHANICS

In Block 1, you have studied the theory of special relativity, formulated by Einstein in 1905. You have learnt how it generalises classical physics to include the domain of high velocities and, in the process, radically alters our notions of space and time. In the theory of special relativity, the velocity of light (c) plays a fundamental role; it is the upper limit of the velocity of any material particle. Newtonian mechanics is an accurate approximation to relativistic mechanics in the domain where all relevant particle velocities are much smaller than c. You should note, however, that Einstein's theory of relativity does not modify the clear distinction between matter and radiation which is at the root of classical physics. Indeed all prequantum physics, relativistic or non-relativistic, is now referred to as classical physics.

During the late nineteenth century and the first quarter of the twentieth century, experimental evidence accumulated about the behaviour of atomic objects (electrons, protons, neutrons, ons and so on) which required new concepts radically different from those of cla physics. Initially new ideas were introduced in an adhoc fashion. But between the years 1925 and 1927, Schrödinger, Heisenberg and Born evolved a consistent description of the behaviour of matter on a microscopic scale which led to the birth of a new theory known as **quantum mechanics**.

What is quantum mechanics? To quote Richard Feynman, an outstanding physicist of this century: "Quantum mechanics is the description of the behaviour of matter in all its details and, in particular, of the happenings on an atomic scale". Quantum mechanics is a totally new way of interpreting data and predicting the behaviour of microscopic particles based on the idea of an essential discontinuity, the *quantum*, in the affairs of the world. Compared to the world-view of classical physics, quantum mechanics gives us a radically new view of the world.

The discovery of quantum mechanics proceeded along two important but separate tracks. The first track was based on the realisation that the allowed values of energy exchange between atomic objects are discrete — *they involve quantum jumps*. This track began with Max Planck' explanation of black-body radiation spectrum, got a big boost from Neils Bohr's theory of the hydrogen atom. It was ultimately carried to fruition by Werner Heisenberg who formulated the matrix mechanics version of quantum mechanics.

The second track began with Albert Einstein's explanation of the photo-electric effect and his discovery of the wave-particle duality of light. Louis de Broglie generalised the attribute of *wave-particle duality* to ma r. Erwin Schrödinger discovered the *wave equation for matter waves* and Ma : Born interpreted the de Broglie-Schrödinger waves as *waves of probability*. This was he wave mechanics version of quantum mechanics. Finally, Paul Dirac showed that the two versions, matrix mechanics and wave mechanics, are entirely equivalent.

The basic concepts of quantum mechanics (wave-particle duality, uncertainty principle, quantisation of physical quantities, etc.) are directly related to the existence of a universal constant, called *Planck's constant h*. Just as the velocity of light plays a central role in relativity, so does Planck's constant in quantum mechanics. Because Planck's constant is very small, quantum physics essentially deals with phenomena at the atomic and subatomic levels. Indeed, if like Mr. Tompkins you were to visit, in your dreams, a world wherein Planck's constant had a large value, you would have the weirdest of experiences! For instance, if you went tiger hunting in a quantum jungle, you would find tigers spread out in space and you might have to return empty handed! For other such experiences, we refer you to the book 'Mr. Tompkins in Paperback' by George Gamow.

In this block we will introduce the basic concepts of quantum mechanics. We will begin our discussion by briefly presenting, in Unit 4, two of the key experimental results (of black-body radiation and photoelectric effect) along with the quantum hypothesis of Planck, Einstein's explanation of the photoelectric effect and Bohr's works which led to the birth of quantum physics. Then we switch over to the de Broglie-Schrödinger track

of the development of quantum mechanics because it is based on mathematics more familiar to you. We introduce **wave-particle duality** in Unit 4 itself. Unit 5 deals with the concepts of **matter waves** and the **uncertainty principle**. The wave equation of matter waves, called the **Schrödinger equation** is the subject matter of Unit 6. Finally in Unit 7 we discuss the basic concepts of the matrix version of quantum mechanics without going into detailed matrix algebra and briefly present the basic features of the unified version of wave mechanics and matrix mechanics formulated by Dirac. Contentwise, the units are more or less evenly balanced and will take about the same time (5 to 6 hours each) to study.

Now there are two aspects of learning quantum mechanics. The first and foremost is, as Richard Feynman used to say, to learn to calculate. However, the quantum mechanical way of calculating is quite different from the classical ways; you will find that you have to get used to a radically new way of thinking. So you will also have to learn to think quantum mechanically. This will involve a certain effort towards exploring the meaning of quantum mechanics, but it will certainly be worthwhile. And if the exploration of the meaning shocks you at times, do not worry. You can take consolation from a comment made by Neils Bohr, "Those who are not shocked when they first come across quantum theory cannot possibly have understood it."

We hope you enjoy studying the block and we wish you success.

Acknowledgement

We are thankful to Shri Gopal Krishan Arora for his invaluable contribution in word-processing of the entire course, and for secretarial assistance.

UNIT 4 WAVE-PARTICLE DUALITY

Structure

4.1 Introduction
Objectives
4.2 The Birth of Quantum Physics
4.3 The de Broglie Hypothesis
Experimental Evidence for the Existence of Matter Waves
Wave-Particle Duality
4.4 Summary
4.5 Terminal Questions
4.6 Solutions and Answers

4.1 INTRODUCTION

You have already taken several courses in physics, from elementary mechanics to electric and magnetic phenomena which constitute classical physics. However, that is not enough if you wish to go deep into physics. You may ask: Why? The answer is that classical physics on its own cannot explain many a natural/observed phenomena. Hence, there seems to be a need of a new physics. In this unit we will dwell briefly on some phenomena and experimental results which defy classical analysis. We will introduce you to the *quantum postulate of Planck*, which was given to explain the experimental results of black body radiation.

It was further extended by Einstein and Bohr to explain the phenomena of photoelectric emission and line spectra of atoms, respectively. The sequence of events is chosen to give you an idea of how quantum physics came into being. Then we discuss one of the basic concepts which laid down the foundations of quantum mechanics, namely, the **wave-particle duality**.

Objectives

After studying this unit you should be able to

- discuss how quantum physics emerged,
- calculate the de Broglie wavelength of a particle in motion,
- explain the concept of wave-particle duality.

4.2 THE BIRTH OF QUANTUM PHYSICS

You already know that a black body absorbs all radiations which fall on it. (Since it does not reflect light and appears black, hence the name — black body.) Usually in the laboratory experiments, a hollow body (cavity) with blackened walls and having a small hole, as shown in Fig. 4.1, acts as a black body. The radiation contained in the body and emitted from the hole produces a **black body spectrum**. In the last century, a number of experiments were carried out to measure the energy per unit volume contained by a black body, denoted by $\rho_T(\lambda)$, at different temperatures. Some of the representative curves showing the variation of $\rho_T(\lambda)$ as a function of λ (black body spectrum) at different temperatures are given in Fig. 4.2. Various investigators tried to explain the nature of these curves using well established laws of classical physics, including thermodynamics. You have studied in Block 4 of the course PHE-06 on Thermodynamics and Statistical Mechanics that it was **Planck** who came up with a theoretical explanation of the black body radiation curve. You know that till 1900, most of the measurements of the energy spectrum of black body radiation were made at smaller wavelengths. These could be satisfactorily explained by Wein's formula given by

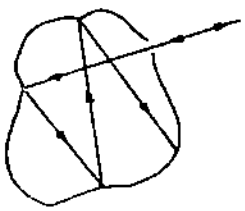


Fig. 4.1 : A black body.

$$\rho_T(\lambda)\, d\lambda = a\lambda^{-5} \exp(-b/\lambda kT)\, d\lambda \qquad (4.1)$$

where a and b were adjustable parameters.

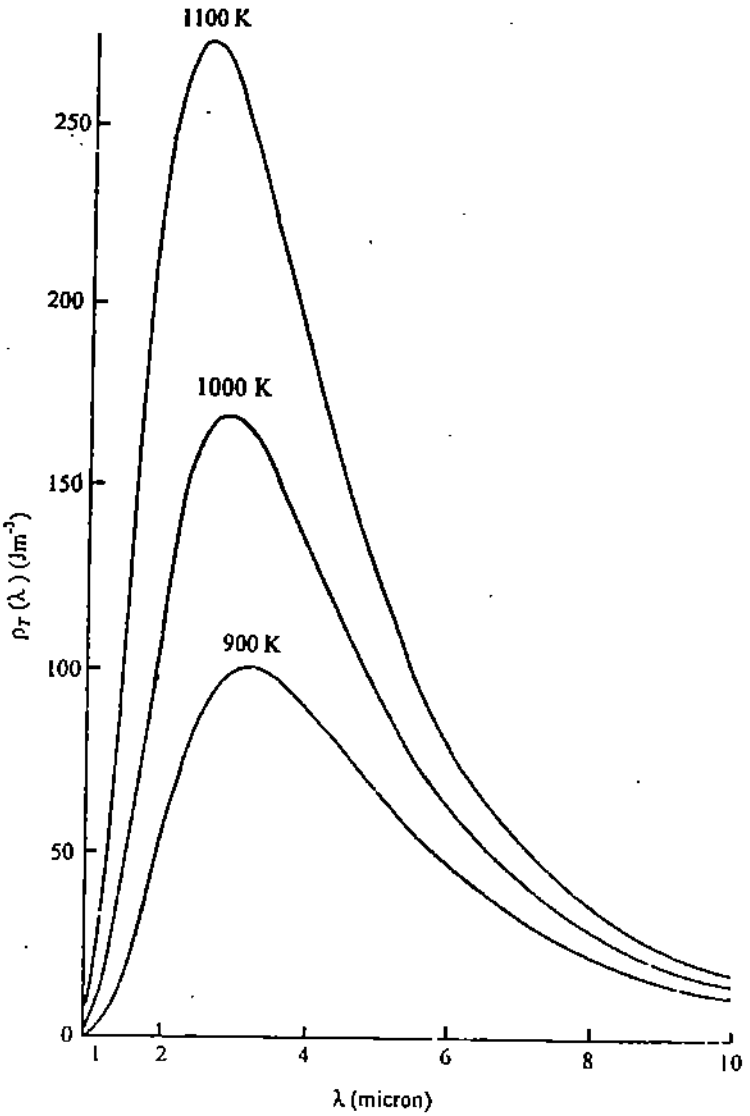


Fig. 4.2 : Black body radiation.

Planck found empirically that by the replacement of $\exp\left(-\frac{b}{\lambda kt}\right)$ by $\exp\left(\frac{b}{\lambda kt}-1\right)^{-1}$, a set of parameters a and b could be obtained which would fit Eq. (4.1) beautifully with the experimental data over all values of λ. However, he found it extremely difficult to give a theoretical justification for the above mentioned replacement. Ultimately, out of desperation, on 18th December, 1900, Planck declared that the only way to derive the correct black body radiation formula was to postulate that the **exchange of energy between matter (walls) and radiation (cavity) could take place only in bundles of a certain size.**

To realise the significance of the above postulate let us consider the following simple example. Suppose two litres of milk is to be distributed between two persons. In how many ways can you distribute the milk? Milk is an infinitely divisible quantity. Hence, you can divide milk between two persons in an infinite number of ways. Now suppose you are restricted to distribute milk only in the units of a litre. Now the number of distribution reduces just to 3 (both persons can receive only 0, 1 or 2 litres of milk). What a drastic change! The number of ways will increase to 5 if you reduce the unit to half a litre.

In classical physics, energy is regarded as an infinitely divisible quantity. Hence the exchange of energy between the walls and the cavity can take place in an infinite number of ways. However, through his postulate Planck reduced the number of ways to be finite. In his model, the exchange must take place in the units of U_0. Thus he introduced the idea of discreteness in the division of energy (which was thought to be

infinitely divisible). If the energy $U(\lambda)$ is to be exchanged, $U(\lambda)/U_0$ must be an integer. If it is not, Planck suggested that it should be an integer close to $U(\lambda)/U_0$.

Planck further postulated that the unit or the *quantum* of energy U_0 is directly proportional to its frequency, i.e.,

$$U_0 = h\nu \tag{4.2}$$

The constant of proportionality h is now known as **Planck's constant** in his honour. Its value is 6.62618×10^{-34} Js. Planck was awarded the Nobel prize for physics in 1918 for his work on black body radiation. [You should note that greater the value of ν, higher will be the value of the quantum of energy U_0 and consequently, lesser will be the number of ways in which energy U can be exchanged.] This new concept of Planck's gave birth to a new physics, known as *quantum physics*. Hence, it is appropriate to take 18th December, 1900 as the date of birth of quantum physics which later on developed into quantum mechanics.

In a further development, Einstein used Planck's quantum hypothesis to successfully explain **the photoelectric effect**.

The Photoelectric Effect

In 1887, Hertz, while working on electromagnetic waves, discovered that the air in a spark gap became a better conductor when it was illuminated by ultraviolet rays. Further investigations by him showed that zinc acquired a positive charge when it was irradiated with ultraviolet rays, i.e., it lost negative charges. In 1900, Leonard showed that the ejected particles were electrons. A series of such experiments revealed that electrons are emitted from a metal surface when light of sufficiently high frequency falls upon it. This phenomenon is known as the **photoelectric effect**.

Fig. 4.3 shows a schematic diagram of the apparatus that was employed in some of these experiments. An evacuated tube contains two electrodes connected to an external circuit like that shown schematically. The anode is made up of the metal plate whose surface is to be irradiated. Some of the photoelectrons that emerge from the irradiated surface have sufficient energy to reach the cathode despite its negative polarity, and they constitute the current that is measured by the ammeter in the circuit.

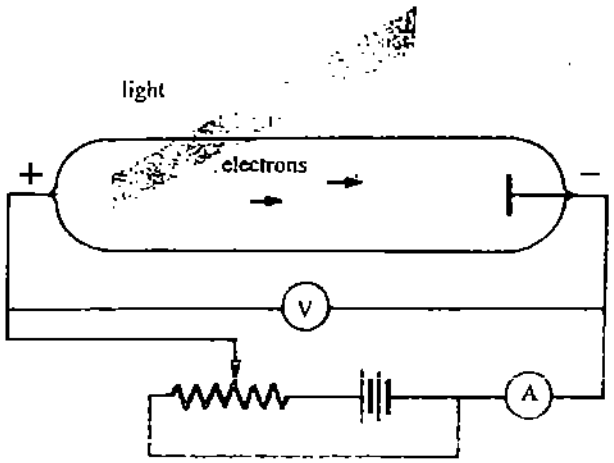


Fig. 4.3 : Schematic diagram of the apparatus for photoelectric effect.

As the collecting voltage V, which retards the electrons, is increased, fewer and fewer electrons get to the cathode and the current drops. Ultimately, when V equals or exceeds a certain value V_0, of the order of a few volts, no further electrons strike the cathode and the current ceases. Figs. 4.4a and b show the experimental curves corresponding to this effect when the intensity of light and collecting voltage V are kept constant.

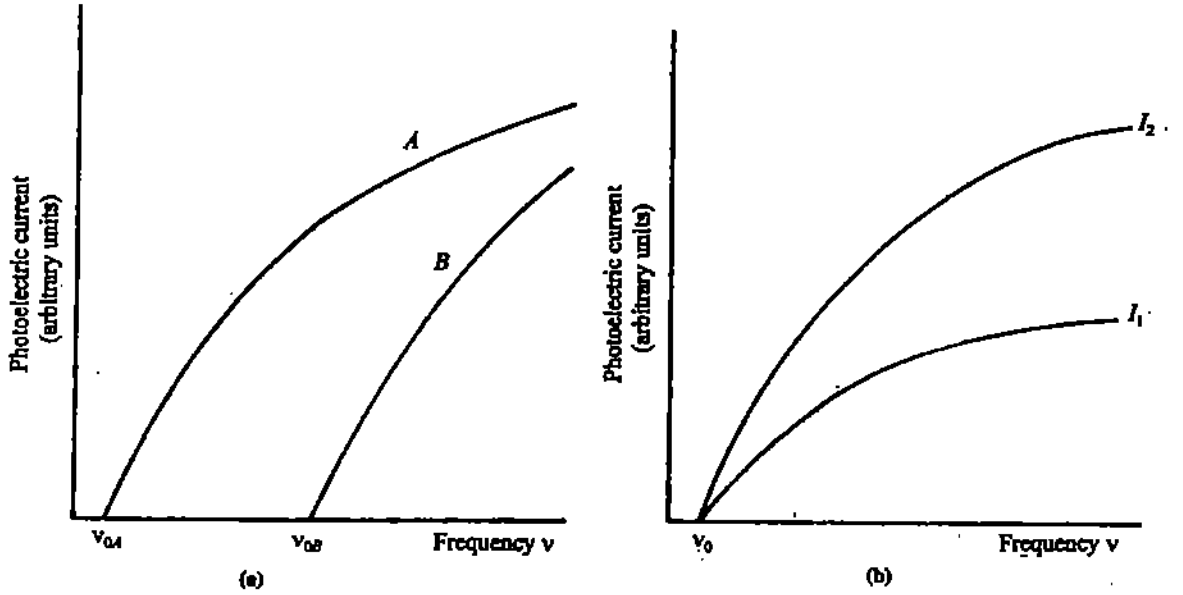


Fig. 4.4 : (a) Variation of the photoelectric current (in arbitrary units) as a function of frequency for two materials *A* and *B*. The intensity of light and the collecting voltage are kept constant; (b) variation of the photoelectric current (in arbitrary units) as a function of frequency for a single material at two values of the intensity; of the intensities, I_2 is greater than I_1.

Note that photoelectron emission occurs only when the frequency of the falling radiation is higher than some threshold frequency ν_0. It was found that for $\nu < \nu_0$, no emission takes place, no matter how intense the radiation is, The value of ν_0 depends upon the material of the surface irradiated. It was also determined that for a given frequency ν ($> \nu_0$), the kinetic energy of the emitted photoelectrons has values between zero and a definite maximum value E_{max}. For any given metal, E_{max} is proportional to $(\nu-\nu_0)$ and is independent of the intensity of the falling light. Further, when electromagnetic waves fall on the material, emission of photoelectrons starts instantaneously (within 10^{-9}s), no matter how weak or strong the falling light is. All these features of the photoelectric effect could not be explained by the classical electromagnetic theory of light on the basis of its wave nature.

You are familiar with the dual nature of light about which you have studied in Unit 1 of the physics elective PHE-09 entitled Optics.

In 1905, Einstein proposed a simple but revolutionary explanation for the photoelectric effect. Einstein extended Planck's postulate of the quanta of energy to the quanta of energy of the electromagnetic field. He viewed the photoelectric phenomenon as a collision between a photon (a quantum of the energy of an electromagnetic field) and a bound electron. In the collision, the photon is completely absorbed and the energy of the bound electron is increased by $h\nu$. Since the electrons are bound in the metal their initial energy E is negative and the largest value of E is $-W$, where W is the work function of the metal. Hence, to escape from the metal, the electron has to use at least an energy equal to W. Thus, the maximum kinetic energy of the photoelectrons will be

$$E_{max} = \frac{1}{2} m v^2_{max} = h\nu - W \qquad (4.3a)$$

If we take $W = h\nu_0$, Eq. (4.3a) may be written as

$$\frac{1}{2} m v^2_{max} = h(\nu-\nu_0) \qquad (4.3b)$$

The implications of Eqs. (4.3 a and b) are:

(1) Since v_{max} has to be positive, no emission can take place for $\nu < \nu_0$.

(2) E_{max} is proportional to $(\nu-\nu_0)$.

(3) An increase in the radiation intensity of frequency ν corresponds to an increase in the number of photons. Since each one of them has the same energy $h\nu$, there is no increase in the energy of the photoelectrons. Only the number of emitted electrons and hence the photoelectric current increases (see Fig. 4.5).

(4) Since the effect is produced by mechanical collisions between electrons and photons, the energy transfer from photons to the electrons is instantaneous. Consequently the time lag is very small.

(5) Since work function $W(=h\nu_0)$ is a characteristic property of the emitting surface, ν_0 is independent of the intensity of incident radiation.

You thus see that Einstein's quantum theory explained each and every aspect of the photoelectric effect with brilliant success, and so the absorption of light in the form of packets or quanta was firmly established.

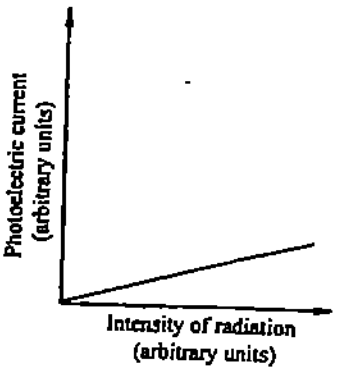


Fig. 4.5 : Variation of the photoelectric current with the intensity of radiation of frequency ν ($> \nu_0$) at a constant collecting voltage.

The next important step in the development of quantum physics was the explanation by Bohr of the stability of the atom as well as the line spectrum emitted by hydrogen atoms. You must be familiar with the Bohr model of the hydrogen atom from your +2 physics courses. However, we have included the details here for completeness.

Bohr's postulates for atomic model

The classical crisis with the model of the atom was not dissimilar to the case of black-body radiation. Ernest Rutherford had proposed the nuclear model of an atom based on his discovery of the atomic nucleus. The electrons, in this model, were supposed to revolve around the nucleus. But then they must radiate and lose energy and eventually spiral into the nucleus. A classical nuclear atom turned out to be unstable!

In 1913, Neils Bohr proposed an atomic model which accounted for the stability of the atom, by injecting quantum ideas into Rutherford's theory. The model also proved highly successful for explaining the spectrum of the hydrogen atom. Bohr's atomic model was based on the following four postulates, three of which were radically different from the earlier models.

(1) Electron in an atom moves in circular orbits about the nucleus with the centripetal force supplied by the Coulomb attraction between the electron and the nucleus.

(2) Of the infinite number of possible circular orbits, only those are allowed for which the value of the orbital angular momentum $|\mathbf{L}|$ of the electron is an integral multiple of $h/2\pi$.

You should note that Bohr preferred quantization of angular momentum instead of energy as was done by Planck, in order to introduce h (the quantum of action) in his theory.

(3) An electron moving in an allowed orbit does not radiate any energy. These states of constant energies are called **stationary states**.

Note that the electron is not stationary in a stationary state.

(4) Energy is emitted (or absorbed) from an atom only when its electron jumps from one allowed orbit of energy E_i to another allowed orbit of energy E_f. The frequency of the emitted (or absorbed) radiation is given by Einstein's frequency condition

$$h\nu = E_i - E_f \qquad \text{(emission } E_i > E_f\text{)}$$

$$= E_f - E_i \qquad \text{(absorption } E_i < E_f\text{)}$$

You should appreciate that the above four postulates are a hybrid of classical and non-classical physics. For example, the first postulate is in accordance with classical physics while the fourth postulate uses quantum ideas. The postulates of the quantization of angular momentum and stationary states are also non-classical.

The first postulate yields the following result for the nth allowed orbit:

$$\frac{mv_n^2}{r_n} = \frac{Ze^2}{4\pi\varepsilon_0 r_n^2} \qquad (4.4)$$

where m is the mass of the electron, Ze is the charge of the nucleus and the electron is moving with a speed v_n in the nth allowed orbit of radius r_n. The second postulate yields

$$\mathbf{L}_n = \mathbf{r}_n \times \mathbf{p}_n = \frac{nh}{2\pi} \hat{\mathbf{L}} ; \qquad n = 1, 2, 3, \ldots \tag{4.5}$$

and

$$L_n = m v_n r_n \qquad \text{(since } \mathbf{v}_n \perp \mathbf{r}_n \text{ for a circular orbit)}$$

where $\hat{\mathbf{L}}$ is a unit vector perpendicular to the plane of the orbit. Thus, we have

$$m v_n r_n = n \hbar, \text{ where } \hbar = \frac{h}{2\pi}, \qquad n = 1, 2, \ldots.$$

or
$$v_n = \frac{Ze^2}{2\varepsilon_0} \frac{1}{nh} \tag{4.6}$$

and

$$r_n = \frac{n^2 h^2 \varepsilon_0}{Z e^2 m \pi} \tag{4.7}$$

The total energy of the electron is the sum of the kinetic energy T_n and the potential energy U_n. Hence for the nth stationary orbit

$$E_n = \frac{1}{2} m v_n^2 - \frac{Ze^2}{4\pi \varepsilon_0 r_n} \tag{4.8a}$$

Putting the values of v_n and r_n from Eqs. (4.6) and (4.7), we may write

$$E_n = -\frac{Z^2 e^4 m}{8 \varepsilon_0^2 h^2} \frac{1}{n^2} \tag{4.8b}$$

or
$$E_n = -\frac{R_\infty Z^2}{n^2} \tag{4.8c}$$

where
$$R_\infty = \frac{m e^4}{8 \varepsilon_0^2 h^2} \tag{4.8d}$$

Thus $E_n \propto n^{-2}$. The suffix ∞ on R appears because the mass of the proton has been assumed to be infinity. Putting the standard values of m, e, ε_0 and h in Eq. (4.8d), we obtain $R = 2.18 \times 10^{-18}$ J (or 13.6eV).

According to Bohr's fourth postulate, the frequency ν_{mn} of the emitted (absorbed) radiation when the electron jumps from the nth state to mth state is given by

$$\nu_{mn} = \frac{R_\infty Z^2}{h} \left(\frac{1}{m^2} - \frac{1}{n^2} \right) \tag{4.9}$$

This agrees remarkably well with the frequency spectrum of the hydrogen atom (see Fig. 4.6). Immediately after Bohr's atomic theory was published, Franck and Hertz performed experiments which demonstrated the existence of discrete energy states.

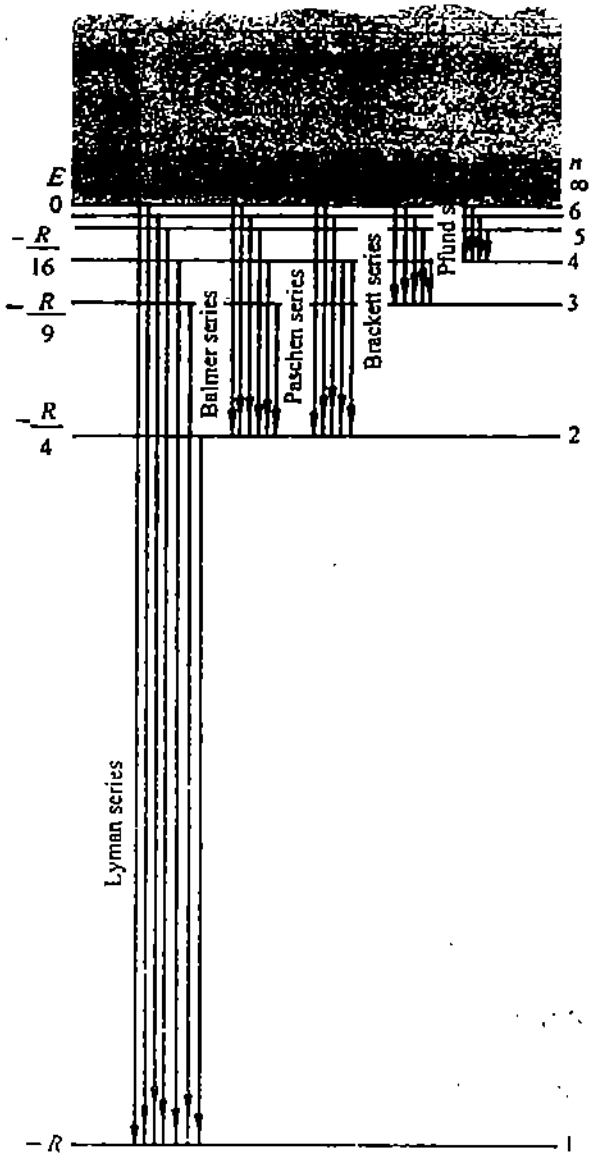


Fig. 4.6 : Energy levels of the hydrogen atom according to the Bohr model. Some of the transitions leading to Lyman series, Balmer series, Paschen series, Brackett series and Pfund series are also shown. All negative energy states represent bound states while positive energy states are continuous states.

However radical these ideas were, you can notice that there was still an attempt to retain a link with the classical physics. You know that at high temperatures or at low frequencies, the black body radiation formula given by Planck reduces to the classical Rayleigh-Jeans formula. Based on these ideas Bohr, in 1923, gave the **Correspondence Principle.**

Bohr's Correspondence Principle

The idea underlying this principle is as follows: *The principles of quantum physics must yield the same results as those of classical physics in situations for which classical physics is valid.* According to the correspondence principle

(a) Quantum theory should give the same results for the behaviour of any physical system as classical physics, in the limit in which the quantum numbers specifying the state of the system become very large.

(b) A selection rule holds true over the entire range of the quantum number concerned. Hence, at large quantum numbers if any selection rule is required to obtain correspondence between classical and quantum physics then the same selection rule holds at low quantum numbers also.

To sum up the discussion so far, recall that Planck introduced a non-classical postulate to explain the black body spectrum which was further extended by Einstein. According to this postulate, the energy states of a simple harmonic oscillator of frequency ν are discrete and the energy of the *n*th discrete state is equal to $nh\nu$ where n is a positive integer and h is a universal constant (called the **Planck constant**). Einstein regarded a quantum of energy as a particle. The quantum of electromagnetic wave is known as **photon**. As you know, this is a particle of zero rest mass which always travels with the velocity of light and carries a momentum $h\nu/c$. Einstein explained the photoelectric effect by regarding the phenomenon as a collision between a photon of energy $h\nu$ and the weakly bound metallic electron in which a photon is completely absorbed and gives its whole energy to the electron which may escape from the metal.

You have also read about Bohr's atomic model according to which electrons in an atom move around the nucleus in certain allowed orbits. In these orbits, energy is conserved and the angular momentum of the electron is an integral multiple of $h/2\pi$. An atom emits or absorbs radiations only when its electron jumps from one allowed orbit to another. His theory proved highly successful in explaining the discrete frequency spectrum of the hydrogen atom. Thus a new physics was born, which is now known as **"old quantum theory"**.

Fig.4.7 : Louis Victor, prince de Broglie (pronounced de Broy), 1892-1987, French theoretical physicist. For his discovery of wave properties of matter he was awarded the Nobel Prize in 1929, after his hypothesis was experimentally confirmed.

This was further developed by de Broglie, Heisenberg and Schrödinger into a new mechanics, now known as quantum mechanics.

Let us now study one of the basic concepts which forms the foundation of quantum mechanics, namely, the de Broglie hypothesis, which led to wave-particle duality.

4.3 THE DE BROGLIE HYPOTHESIS

Louis de Broglie (Fig. 4.7) must have been a lover of music, for he realised that Bohr's stationary orbits of electrons confined in atoms must have something in common with stationary waves on guitar strings. Could the discreteness of atomic orbits be due to the discreteness of electron waves in captivity? On a guitar string, stationary waves form a discrete pattern of harmonics just like the discrete Bohr orbits. De Broglie asked: Could atomic electrons be confined waves and therefore produce a discrete stationary wave pattern? For example, the lowest atomic orbit is one in which one electron wavelength fits the circumference of the orbit and the higher orbits fit two or more electron wavelengths (Fig. 4.8).

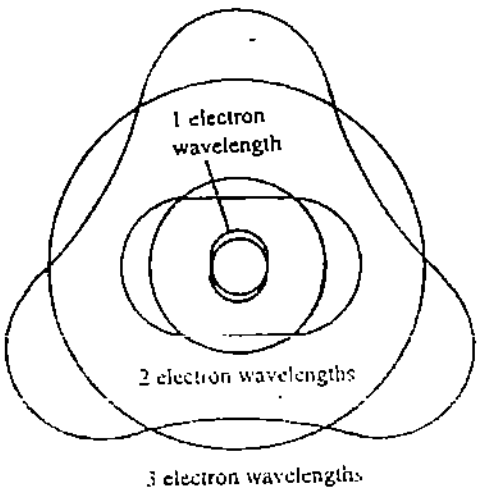


Fig. 4.8 : Stationary waves of electrons confined in an atom. The electron wave fits an integral number of wavelengths in each of the successive Bohr orbits.

It was established by then that electromagnetic waves exhibit both wave and particle properties. But electromagnetic waves and material particles are two important classes of entities which appear to be basic to the structure of matter. If electromagnetic waves have a dual wave-particle nature then why shouldn't electrons, and in fact, all matter, too have a dual nature?

As a young French graduate student, de Broglie, in 1924, argued with a great amount of insight that since nature loves symmetry and simplicity in physical phenomena, all material particles should exhibit both wave and particle nature. He further argued that the wave description of light in geometrical optics is an approximation of the more general wave analysis. Similarly, the description of particle motion in terms of line trajectories is an approximation of a more general description of the particle, containing its wave aspect. De Broglie further proposed that the wavelength and frequency of the matter waves should be determined by the momentum and energy of the particle in exactly the same way as for photons. Recall that for a photon, energy and momentum are related as follows:

$$E = pc$$

and from Eq. (4.2), $E = h\nu = \hbar\omega$, ($\omega = 2\pi\nu$). From wave theory, the angular frequency ω is related to the wave number k:

$$\omega = ck$$

where c is the wave speed.

Combining these equations, we get

$$\begin{aligned} pc &= \hbar\omega \\ &= \hbar\, c\, k \end{aligned}$$

or
$$p = \hbar k \tag{4.10}$$

Since $k = \frac{2\pi}{\lambda}$, Eq. (4.10) can also be rewritten to get an expression for the **de Broglie wavelength** of *matter waves associated with a particle having momentum* p:

$$\boxed{\text{de Broglie wavelength } \lambda = \frac{h}{p}} \tag{4.11}$$

Thus, the wavelength of the wave associated with a particle of matter in motion is inversely proportional to the particle's momentum, and the constant of proportionality is just the Planck constant. We can recast Eq. (4.11) into another form by using the relation $E = p^2/2m_0$ for a free particle:

$$\boxed{\lambda = \frac{h}{p} = \frac{h}{(2m_0E)^{1/2}}} \tag{4.12}$$

where m_0 is the rest mass of the particle, and E its energy.

Now the **phase velocity** v_p of a wave is given by $v_p = \nu\lambda$. Hence, we may write

$$v_p = \frac{E}{h}\frac{h}{p} = \frac{E}{p} \tag{4.13a}$$

Putting $E = mc^2$ and $p = mv$ in Eq. (4.13a), we get

$$v_p = c^2/v \tag{4.13b}$$

Since $v < c$, the phase velocity of waves associated with matter turns out to be greater than the velocity of light. Does this disturb you? Do not worry because no physical quantity like energy, signal or information etc. associated with the wave, travels with its phase velocity.

Let us now understand what Eq. (4.11) means.

Eq. (4.11) is a complete statement of the **wave-particle duality**. It clearly shows that a particle with a momentum p can exhibit wave-like properties and the wavelength of the associated matter waves is h/p. The converse is also true, i.e., a wave of wavelength λ can exhibit particle-like properties and the momentum of the wave-matter is h/λ. However, you should clearly understand the difference between electromagnetic waves and matter waves. For matter waves, phase velocity is always greater than the velocity of light but, since energy is carried by the particle, the velocity with which energy is transported by matter waves is equal to the velocity of the particle. On the other hand, the phase velocity of electromagnetic waves and the velocity with which energy is transported by them are both equal to the velocity of light.

By now, are you not wondering that if matter can exhibit wave-like properties, why don't macroscopic objects appear like waves to us?

To understand this, make the following simple calculation.

Special 5 min

SAQ 1

A ball of mass 10^{-3} kg moves with a velocity of 10^{2} m s^{-1}. What is the de Broglie wavelength of the ball?

Clearly, this wavelength is too small to be detected experimentally. Hence, we can say that matter waves are associated with macroscopic objects. However, their wave character is not observable. Thus, they can safely be described as particles under all circumstances. Now do the following exercise.

SAQ 2

You have calculated the wavelength to be about 1.2×10^{-10} m for the 100 eV electron and 2.9×10^{-14} m for the 1 MeV neutron, respectively. Clearly the wavelength of the matter wave associated with 100 eV electron lies in the *X*-ray region and is of the same order as the atomic spacing in a crystal. The electron waves, like *X*-rays, are, therefore, expected to undergo diffraction by crystals. On the other hand, the wavelength of a 1 MeV neutron is too small for observing the diffraction by a diffracting grating. However, low energy neutrons, say 100 eV neutrons, would have wavelength in the *X*-ray region and then their diffraction pattern can be obtained.

We thus see that for the motion of macroscopic objects, the de Broglie hypothesis does not change the classical description as developed by Newton. But for microscopic objects the wavelengths of matter waves are long enough to undergo observable diffraction. The diffraction of matter waves was observed experimentally as early as 1927.

4.3.1 Experimental Evidence for the Existence of Matter Waves

The first experimental demonstration of matter waves came three years after de Broglie advanced his hypothesis and was accidentally obtained by two American physicists Clinton J. Davisson and his assistant L.H. Germer. They were studying the scattering of electrons by crystals using the apparatus shown in Fig. 4.9.

Electrons from an electron gun were accelerated by a positive electrode maintained at V volts as compared to the filament. The accelerated beam of electrons was incident on a strip of nickel containing many crystals and the number of electrons scattered in various directions was then measured. A smooth variation in the intensity of the scattered electron with the angle was observed. Midway through the experiment, an accident occurred which permitted air to enter into the vacuum tube containing the strip.

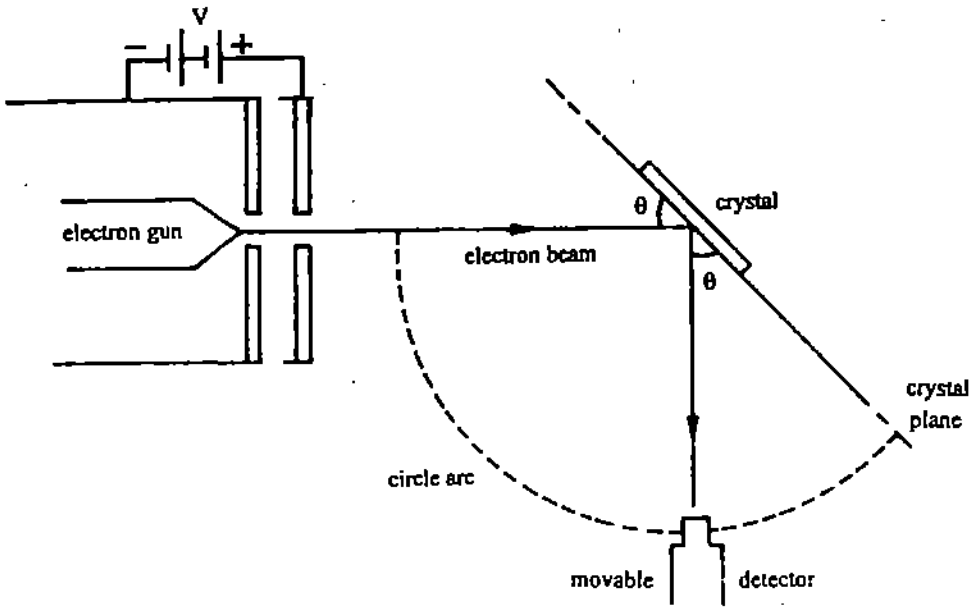


Fig. 4.9 : Schematic diagram of Davisson-Germer Experiment.

This resulted in the formation of an oxide film on the surface. Davisson and Germer were then forced to heat the strip to a very high temperature in order to reduce the oxide. This heating, and slow cooling had the effect of turning the polycrystalline nickel sample into a large single crystal. After this forced heat treatment of the sample, the experiment was repeated. The experimental results this time were quite different from those obtained before the accident. The intensity of scattered electrons showed some sharp maxima and minima at certain angles which were found to depend upon the electron energy and hence upon the accelerating voltage. This pattern was similar to a wave diffraction pattern. Some typical results from these experiments are shown in Fig. 4.10.

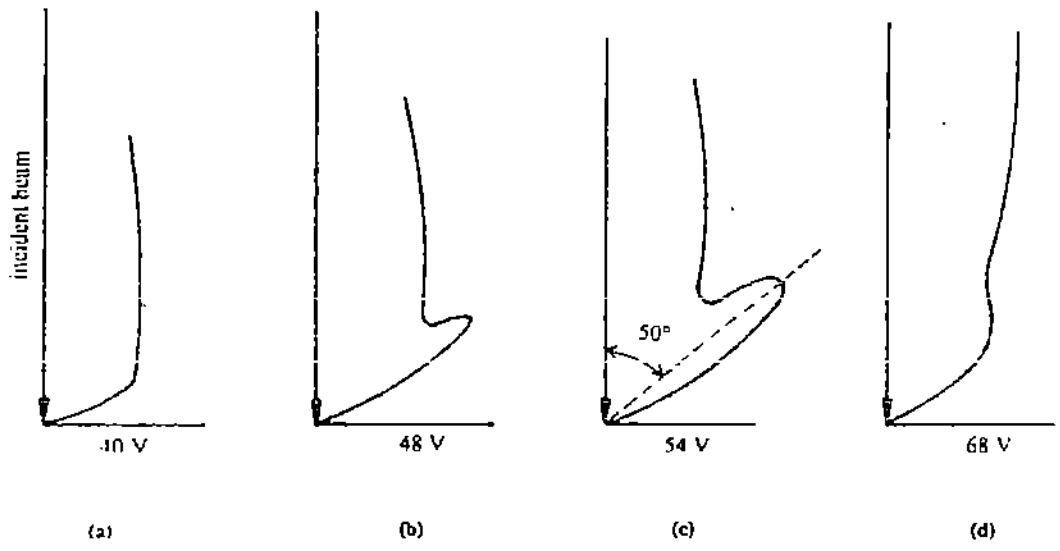


Fig. 4.10 : Polar plots of some typical results of Davisson-Germer experiment.

Davisson and Germer found that the angular peaks of scattered electrons could be explained as diffraction of electrons by atoms if the regular arrays of atoms in the nickel crystal are regarded as a diffraction grating.

Diffraction of Electrons

Using Bragg's analysis developed for *X*-ray scattering, and the angles of scattering at which the pronounced peaks are produced, they calculated the wavelength of electron waves. They finally found very good agreement between their values of the wavelength and those predicted by de Broglie hypothesis. Thus the validity of the de Broglie hypothesis was established.

In one particular experiment, the electron beam was accelerated to a potential of 54 volts and the maximum intensity was observed at an angle of 65° between a particular family of crystal planes and the incident (or the scattered) beam. The spacing between the crystal planes, as measured by *X*-ray diffraction technique was found to be 0.91 Å. Let us now calculate the wavelength of the electron-waves from these electron diffraction data.

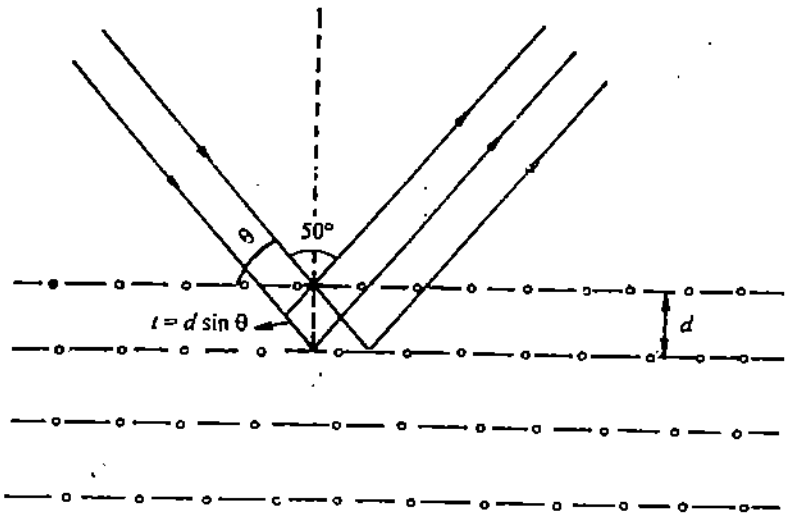


Fig. 4.11 : Bragg's analysis of scattering by crystal planes.

Fig. 4.11 clearly shows that the phase difference between waves coming from adjacent scattering planes of crystals is given by $(2\pi/\lambda) \times 2d \sin\theta$. For constructive interference we, therefore, have

$$(2\pi/\lambda)\ 2d \sin\theta = 2\pi n$$

i.e., $$\lambda = 2d \sin\theta/n, \qquad n = 1, 2, 3, \ldots \tag{4.14}$$

Putting the given data in Eq. (4.14) along with $n = 1$, we get

$$\lambda = 2\ (0.91\ \text{Å}) \sin 65° = 1.65\ \text{Å} \tag{4.15a}$$

Let us now compare this value of λ with the corresponding value predicted by de Broglie's hypothesis. The de Broglie wavelength of an electron accelerated by a potential of V volts is given by

$$\lambda = \frac{h}{p} = \frac{h}{(2em_0V)^{1/2}} = \frac{12.264}{(V(\text{volt}))^{1/2}}\ \text{Å} \tag{4.15b}$$

In deriving Eq. (4.15b) we have used Eq. (4.12). Hence the wavelength of a 54 eV electron is

$$\lambda = \frac{12.264}{(54)^{1/2}}\ \text{Å} = 1.67\ \text{Å} \tag{4.15c}$$

You will notice that the agreement between the two values of the wavelength of electron, given by Eqs. (4.15a) and (4.15c) is remarkable.

Within months of the discovery by Davisson and Germer, a British physicist G.P. Thomson also discovered diffraction effects with a beam of highly energetic electrons. In 1928, Thomson repeated his experiment by using platinum rather than the celluloid film. The diffraction rings observed with the polycrystalline metallic foil were found to be exactly similar to those observed for *X*-rays of the same wavelength (as that of electron). This experiment of Thomson had provided even more convincing evidence in support of the de Broglie hypothesis. De Broglie was rewarded with a Nobel Prize in Physics in 1929 and the same was awarded to Davisson and Thomson in 1937. It is of interest to note that Sir J.J. Thomson was awarded Nobel Physics Prize in 1906 for discovering the electron as a particle carrying negative electrical charge and in 1937 his son G.P. Thomson got the same prize for establishing the wave nature of electron. With suitably designed experiments, the wave-like behaviour of particles such as α-particles, protons, neutrons etc. has also been established.

You may now like to work out an SAQ.

SAQ 3

Spend 10 min

Electrons of 400 eV are diffracted through a crystal and a second order maximum is observed where the angle between the diffracted beam and incident beam is 30°. Calculate

(i) the wavelength of the electron matter wave.

(ii) the interplanar distance of those lattice planes which are responsible for this maximum.

Has the wave behaviour of electrons come as a surprise to you or shocked your sensibilities? After all, all along we have been accustomed to regarding an electron as a particle and now it exhibits wave properties. What is this duality in an electron's behaviour? What is wave-particle duality? Let us further explore the meaning of wave-particle duality.

4.3.2 Wave-Particle Duality

Let us go to the very basics. What do we mean by a **particle**? We define it as an entity possessing a **definite position, size, mass, velocity, momentum**, energy, etc. Its motion is described by Newton's laws of motion. It must remain in a state of uniform motion or a state of rest when no external force acts on it. It should accelerate under the influence of a force field and move along a particular trajectory with a well defined position and time relationship. This, in a nutshell, is a picture of a particle that we have acquired from our studies in physics so far. If there exists any entity which does not conform to this description of a particle, we should not call it a particle.

Now, what do we understand by a **wave**? A wave is characterised by its properties of **periodicity in space** and **periodicity in time**; it possesses a **wavelength, amplitude, frequency** and propagates at a certain **wave velocity. It can transport energy without transport of matter. It cannot be localized and extends in space.** These are the basic ideas associated with a wave. If we find anything which does not conform to all of these ideas, we should not call it a wave.

Having thus conceptualised a particle and a wave, the next step is fairly easy. If there exists something in nature which has neither purely particle properties nor purely wave properties but has properties of both, e.g., mass, momentum, wavelength, amplitude, frequency and is neither localized at a point nor extends to infinity, we should call it neither a particle nor a wave. For want of a better name, we simply call it a **wave-particle**.

The fact that there are no particles and no waves in this universe but only wave-particle dualities should not unduly bother you. If that's the way nature works, that's the way we accept it. The definitions of a wave and a particle are still very useful and serve as good approximations as was amply demonstrated in SAQ 1. To sum up, the concept of

wave-particle duality applies universally to all objects. However, since this duality involves Planck's constant, which has a very small value, the effect is appreciable only in the microscopic world. In the macroscopic world of our experience, objects obey the classical laws of motion.

You should also realise that wave-particle duality arises because of the finite value of Planck's constant. In classical physics it is assumed that $h = 0$, i.e., energy quanta do not exist. Hence it cannot explain wave-particle duality. However, by now you have seen that there are compelling reasons to take energy in the quantized form such that each quanta of energy is equal to $h\nu$ or its integer multiple.

One last word on the concept of wave-particle duality; its acceptance is not merely a question of belief or faith — it is a question of experimental observation and accepting a model for explaining them.

Let us now summarise what you have studied in this unit.

4.4 SUMMARY

- To explain black body radiation Planck proposed the idea of a **quantum of energy** for a harmonic oscillator. According to **Planck's quantum postulate**, a quantum of energy E for a wave is given by

  $$E = h\nu$$

  where ν is the frequency of the wave and h is a universal constant, known as **Planck's constant.**
- This idea was extended by Einstein to light to explain the **photoelectric effect.** Einstein regarded the quantum of light, the photon, to be a particle of energy $E = h\nu$.
- Bohr used the quantum postulate and certain other postulates in his **atomic model** and successfully explained the stability of atom and the line spectra of hydrogen atoms. He further gave the **correspondence principle** which establishes a correspondence between classical and quantum physics.
- Just as electromagnetic fields exhibit both wave-like and particle-like properties, de Broglie proposed that matter too had wave-like properties, giving rise to **wave-particle duality.** The **de Broglie wavelength** of a particle of momentum p is

  $$\lambda = \frac{h}{p}$$

  For a free particle of mass m and energy E

  $$\lambda = \frac{h}{(2mE)^{1/2}}$$

4.5 TERMINAL QUESTIONS

Spend 30 min

1. Derive Bohr's angular momentum quantization condition for the Bohr atom from de Broglie's relation.
2. High energy protons of 200 GeV (1 GeV = 10^9 eV) are diffracted by a hydrogen target at an angle θ given by

   $$p \sin\theta = \frac{1.2}{c} \text{ GeV}$$

   Note that the protons are moving at relativistic energies. Estimate the radius of the proton.
3. A 150 eV increase in an electron's energy changes its de Broglie wavelength by a factor of two. Calculate the initial de Broglie wavelength of the electron.

4. Calculate the de Broglie wavelength and the kinetic energy of electrons which undergo first-order Bragg diffraction by a nickel crystal at an angle of 30°. For nickel, $d = 2.15$ Å.

4.6 SOLUTIONS AND ANSWERS

Self-Assessment Questions

1. The de Broglie wavelength is

$$\lambda = \frac{h}{p} = \frac{6.626 \times 10^{-34}\ \text{Js}}{10^{-3}\ \text{kg} \times 10^{-2}\ \text{m s}^{-1}} = 6.626 \times 10^{-29}\ \text{m}$$

2. $$\lambda_e = \frac{h}{(2m_0 E)^{1/2}} = \frac{6.626 \times 10^{-34}\ \text{Js}}{(2 \times 9.109 \times 10^{-31}\ \text{kg} \times 100 \times 1.6 \times 10^{-19}\ \text{J})^{1/2}}$$

$$= 1.227 \times 10^{-10}\ \text{m}$$

$$\lambda_n = \frac{6.626 \times 10^{-34}\ \text{Js}}{(2 \times 1.675 \times 10^{-27}\ \text{kg} \times 10^{6} \times 1.6 \times 10^{-19}\ \text{J})^{1/2}}$$

$$= 2.862 \times 10^{-14}\ \text{m}$$

3. $$\lambda_e = \frac{h}{(2m_0 E)^{1/2}} = \frac{6.626 \times 10^{-34}\ \text{Js}}{(2 \times 9.109 \times 10^{-31}\ \text{kg} \times 400 \times 1.6 \times 10^{-19}\ \text{J})^{1/2}}$$

$$= 0.61 \times 10^{-10}\ \text{m} = 0.61\ \text{Å}$$

$$n\lambda_e = 2d \sin\theta$$

$$n = 2,\ \theta = 30°,\ \lambda_e = 0.61 \times 10^{-10}\ \text{m}$$

$$\therefore\ d = \frac{\lambda_e}{\sin\theta} = 1.22\ \text{Å}$$

Terminal Questions

1. De Broglie visualised that atomic electrons were confined waves and, therefore, produced a discrete stationary wave pattern. Then only those orbits would be allowed in which an integral number of electron wavelengths could fit the circumference (see Fig. 4.8). For example, one wavelength would fit the circumference of the lowest atomic orbit and two or more electron wavelengths would fit into higher orbits. Thus if de Broglie waves of wavelength λ fit a Bohr orbit of radius r to satisfy the stationary condition, we must have

$$2\pi r = n\lambda, \qquad n = 1, 2 \ldots\ldots$$

Since $\lambda = \frac{h}{p} = \frac{h}{mv}$, we get

$$\frac{2\pi r\, mv}{h} = n$$

or $$mvr = \frac{nh}{2\pi}$$

which is the Bohr angular momentum quantization.

2. $1\ \text{GeV} = 10^9\ \text{eV}$

Let R be the radius of the proton. The dimension of the 'slit' which scatters protons is $2R$. Therefore

$$\lambda = 2 \times 2R \sin\theta$$

or $$\sin\theta = \frac{\lambda}{4R}$$

For protons $p = \dfrac{200 \text{ GeV}}{c}$ $(\because E = pc)$

$$\therefore \sin\theta = \frac{1.2 \text{ GeV}}{pc}$$

$$= \frac{1.2}{200} = 0.006$$

From de Broglie relation

$$\lambda = \frac{h}{P} = \frac{6.626 \times 10^{-34} \text{ Js} \times 3 \times 10^{8} \text{ ms}^{-1}}{200 \times 10^{9} \times 1.6 \times 10^{-19} \text{ J}}$$

$$= 6.212 \times 10^{-18} \text{ m}$$

$$R = \frac{\lambda}{4 \sin\theta} = \frac{6.212 \times 10^{-18} \text{ m}}{4 \times 6 \times 10^{-3}}$$

$$= 2.5 \times 10^{-16} \text{ m}.$$

The radius of the proton is of the order of 10^{-16} m.

3. Since the energy of the electron increases, its wavelength will decrease. If λ be the initial electron wavelength, E its initial energy and ΔE, the increase in energy, we can write using Eq. (4.12):

$$\lambda = \frac{h}{(2mE)^{1/2}}$$

and $$\frac{\lambda}{2} = \frac{h}{[2m(E+\Delta E)]^{1/2}}$$

Simple algebra yields the relation

$$\lambda^2 = \frac{3h^2}{2m\Delta E}$$

or $$\lambda = h\left(\frac{3}{2m\Delta E}\right)^{1/2}$$

Substituting the values of h, m and $\Delta E = 150$ eV, we get

$$\lambda = 6.626 \times 10^{-34} \text{ Js}\left[\frac{1.5}{9.1 \times 10^{-31} \text{ kg} \times 150 \times 1.6 \times 10^{-19} \text{ J}}\right]^{1/2} = 1.73 \text{ Å}$$

4. For the first-order Bragg diffraction, $n = 1$ and

$$\lambda = 2d \sin\theta$$

$$= 2 \times 2.15 \text{ Å} \sin 30^\circ$$

$$= 2.15 \text{ Å}$$

The kinetic energy of the electron is obtained from Eq. (4.12):

$$E = \frac{h^2}{2m\lambda^2}$$

or $$E = \frac{(6.626 \times 10^{-34} \text{ Js})^2}{2 \times 9.1 \times 10^{-31} \text{ kg} \times (2.15 \times 10^{-10})^2 \text{ m}^2} = 52 \times 10^{-19} \text{ J} = 32.5 \text{ eV}.$$

UNIT 5 MATTER WAVES AND UNCERTAINTY PRINCIPLE

Structure

5.1 Introduction
Objectives
5.2 Matter Waves
5.3 The Uncertainty Principle
Some Thought Experiments
Some Applications of the Uncertainty Principle
5.4 Summary
5.5 Terminal Questions
5.6 Solutions and Answers

5.1 INTRODUCTION

In Unit 4, you have learnt how quantum physics emerged in order to explain certain experimental results and natural phenomena which could not be accounted for by classical physics. You have also studied the concepts of wave-particle duality and matter waves given by de Broglie. We will explore these concepts further in this unit. You know that waves are spread all over space whereas particles are localised. So a single wave would be inadequate for describing a real particle correctly. The question is: How do we represent matter waves (or wave-particles) in space? For this purpose, we have to introduce the concept of a **wave packet**.

A discussion of wave packets leads us to another fundamental principle of quantum physics, namely, **Heisenberg's uncertainty principle**. You will study some applications of the uncertainty principle, in particular for the microscopic world. The uncertainty principle was met with a great deal of opposition from stalwarts in physics, especially Albert Einstein. The debate amongst physicists, and in particular between Bohr and Einstein about the validity of this principle makes a very interesting reading in the history of quantum mechanics. In this unit we will give you a flavour of some ideal (thought) experiments which provided support for the uncertainty principle and ultimately established it as one of the fundamental principles of quantum mechanics. In the next unit, you will study the Schrödinger equation which is a major pillar of quantum mechanics.

Objectives

After studying this unit you should be able to

- explain the concept of a wave packet,
- derive the relation between phase velocity and particle velocity,
- apply Heisenberg's uncertainty principle to microscopic systems,
- discuss the γ-ray microscope, single slit and double slit interference experiments in support of the uncertainty principle.

5.2 MATTER WAVES

From the discussion in Sec. 4.3.2 of Unit 4, you know that classically, a particle can be localised at a single point but a wave cannot. Thus, at least for the microscopic particles for which the wave-particle duality is significant, we are forced to abandon the classical description of a particle. We have to look for a new description which should be consistent with the de Broglie hypothesis. What should this new description of matter waves associated with particles be like? Well, for one, matter waves should always be associated spatially with the particle in such a manner that the resultant amplitude is non-zero only in the neighbourhood of the particle.

Now the de Broglie equation (4.11) yields

$$\frac{\Delta\lambda}{\lambda} = -\frac{\Delta p}{p} \tag{5.1}$$

This equation shows that if $\Delta p = 0$ then $\Delta\lambda \neq 0$. That is, *we are required to represent a particle of definite momentum with a single wave of fixed wavelength.* However, you know that a wave of single wavelength and frequency is spread out in time and space. Thus, it cannot be localised and cannot represent a particle. Can we find an intermediate solution? If we associate some uncertainty with the momentum, i.e., if we take $\Delta p > 0$, then $\Delta\lambda$ is also finite. The following discussion shows that if a finite spread is allowed in the wavelength, we can fruitfully exploit this to represent a microscopic particle.

Consider a simple situation in which two sinusoidal waves with slightly different wavelengths are superimposed. You have studied in the course PHE-02 (Oscillations and Waves) that the character of the resultant wave is quite different from the individual waves. For example, suppose we consider two travelling waves represented by

$$\psi_1 = A \sin (kx - \omega t)$$

and $$\psi_2 = A \sin [(k + dk)x - (\omega + d\omega)t]$$

where A is their amplitude, k ($=2\pi/\lambda$) is the wave number and ω ($=2\pi\nu$) is the angular frequency. The superposition of these waves yields a resultant wave given by

$$\psi = \psi_1 + \psi_2$$

$$= 2A \left[\cos\left(\frac{dk}{2}\right)x - \left(\frac{d\omega}{2}\right)t\right] \sin\ (kx - \omega t)$$

Using the result

$\sin\theta + \sin\phi =$

$2\cos\left(\frac{\phi-\theta}{2}\right)\sin\left(\frac{\phi+\theta}{2}\right)$

we get the value of ψ.

where we have ignored $\frac{dk}{2}$ and $\frac{d\omega}{2}$ in the sine term as they are infinitesimal compared to k and ω. Fig. 5.1 shows a graph of ψ. You can see that ψ has an envelope equal to $2A\cdot\cos\left(\frac{dk}{2}x - \frac{d\omega}{2}t\right)$ modulating the sine wave given by $\sin(kx - \omega t)$.

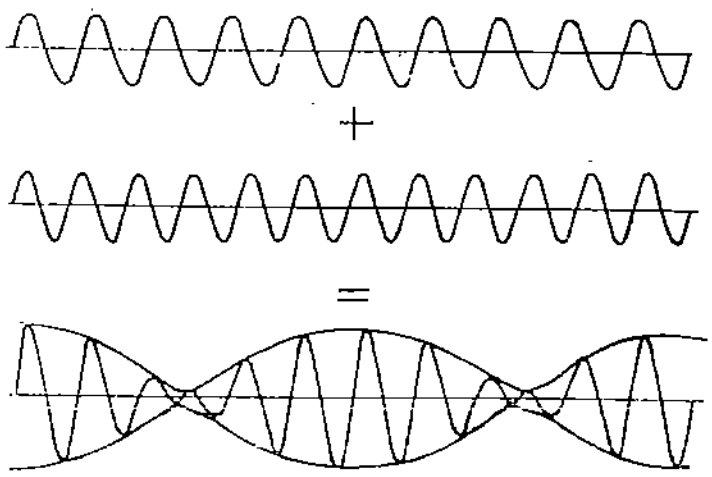


Fig. 5.1 : A sketch of the resultant of the superposition of two travelling waves.

Similarly, by superimposing a very large number of waves having wavelengths close to a central wavelength λ_0, a wave packet such as shown in Fig. 5.2 can be constructed. The superposition of these waves results in a variation in amplitude that defines the shape of the wave packet. The wave packet has regular spacing between successive maxima or minima, equal to the central wavelength λ_0. Thus the wavelength of the

wave packet is λ_0 but at any instant of time it is localized in a finite region of space. Clearly, such a wave packet exhibits both wave and particle aspects.

Thus, in this new representation, a microscopic particle may be represented by a wave packet. To sum up, we may define a wave packet as follows:

> **A wave packet** is a group of waves with slightly different wavelengths and frequencies interfering with one another in such a manner that the amplitude of the group (i.e., the envelope) is non-zero only in the neighbourhood of the particle.

The spread of a wavepacket in wavelength (and in frequency) depends on the required degree of localization in space (and time). You should note that the central wavelength λ_0 is given by the de Broglie equation (4.11).

Fig. 5.2 : A wave packet.

How do we determine the velocity of a wave packet? Clearly, if the velocities of the individual waves being superimposed are the same, the velocity with which the wave packet travels is the common wave velocity. However, in the case of de Broglie waves, the wave velocity varies with wavelength; the individual waves do not travel at the same velocity. Thus, the wave packet has a different velocity from the waves that compose it. Let us now determine the **phase velocity** v_p and the **group velocity** v_g of the wave packet.

You know that the phase velocity v_p of a wave is given by $v_p = \frac{\omega}{k}$. Hence from Eqs. (4.2 and 4.11) of Unit 4, the phase velocity of the wave packet is given by

$$v_p = \frac{\omega}{k} = \frac{E}{p}$$

Putting $E = mc^2$ and $p = mv$ in this equation, we obtain

$$\text{phase velocity} \quad v_p = \frac{c^2}{v} \tag{5.2}$$

Since $v < c$, it is clear that the phase velocity of a wave packet associated with a particle is greater than that of light. This should not disturb you because no physical quantity like energy, information or signals etc., associated with the wave, travels with the phase velocity. These entities move with the group velocity which is given by

$$v_g = \frac{d\omega}{dk} = \frac{dE}{dp} \tag{5.3}$$

Now from the special theory of relativity

$$E^2 = p^2 c^2 + m_0^2 c^4, \tag{5.4a}$$

$$E = mc^2, \tag{5.4b}$$

and

$$p = mv \tag{5.4c}$$

Hence using these three equations we obtain

$$v_g = \frac{dE}{dp} = \frac{pc^2}{E} = \frac{p}{m} = v$$

or

$$\text{group velocity} \quad v_g = v \tag{5.5}$$

Therefore, the group velocity of the wave packet is nothing but the particle's velocity.

Thus far you have learnt that a single wave is not enough to represent a particle. We need to superimpose a group of waves which yields a wave packet travelling at a group velocity equal to the particle's velocity. You have seen that there is an uncertainty Δp in the momentum of the wavepacket and a spread $\Delta\lambda$ in its wavelength. Before we proceed further to analyse the implications of this discussion, we would like you to calculate the phase velocity and group velocity of a wave packet.

Spend 5 min

SAQ 1

The energy of a free electron including its rest mass energy is 1 MeV. Calculate the group velocity and the phase velocity of the wave packet associated with the motion of the electron.

You have studied so far in this section that a moving particle must be regarded as a wave packet which satisfies the de Broglie relation. We construct a localised wave packet by superimposed waves which leads us to an uncertainty in its momentum and wavelength.

The fact that a moving particle must be represented by a wave packet rather than a localised entity suggests that there is a fundamental limit to the accuracy with which we can measure the particle's position and momentum. For example, the wider the wave packet is, the greater are the number of waves in it, and the better our chances are to determine the particle's wavelength and hence its momentum. But, because the particle can be anywhere in the wave packet, we cannot determine its position with precision. If, however, the wave packet is narrow, the particle's position is better defined, but now its wavelength (or its momentum) is difficult to determine. So the smaller is the uncertainty Δx in the particle's position, the larger becomes the uncertainty Δp in its momentum, and vice versa.

Fig. 5.3 : Werner Heisenberg, 1901-1976, was a German theoretical physicist. He was one of the founders of quantum mechanics, and received the Nobel Prize in 1932.

Thus we can say that *a direct consequence of the wave-particle duality is the appearance of uncertainties (spreads) in the momentum and the position of a particle.* If one of them becomes definite, the other becomes completely indefinite. This situation is in sharp contrast to that of classical mechanics according to which it is possible to determine precisely the position and the momentum of a particle at any time t. In 1927, Heisenberg (Fig. 5.3) advanced the above concept in the form of the **uncertainty principle.**

5.3 THE UNCERTAINTY PRINCIPLE

Heisenberg discovered that the product of the position and momentum uncertainties of a quantum object such as the wave packet is greater than or equal to Planck's constant h. Thus, according to Heisenberg's uncertainty principle

$$\Delta x\, \Delta p_x \geq \hbar \qquad (5.6a)$$

where Δx and Δp_x are the uncertainties in the x component of the position and momentum of a microscopic particle, respectively and $\hbar = h/2\pi$.

Similar relationships hold for the y and z components of the positions with their respective momenta of the object. However, you should note that the Heisenberg uncertainty principle does not impose restriction on the simultaneous and precise measurements of x, y and p_z or y, p_x and p_z etc. The restrictions are only on what are called as **conjugate variables**, i.e., x along with p_x, y along with p_y and z along with p_z. Thus, we have

$$\Delta y\, \Delta p_y \geq \hbar \qquad (5.6b)$$

$$\Delta z \Delta p_z \geq \hbar, \tag{5.6c}$$

and $$\Delta r \Delta p_r \geq \hbar \tag{5.6d}$$

A general statement of Heisenberg's uncertainty principle can be given as follows:

The Uncertainty Principle

The values of two (canonically conjugate) variables cannot be simultaneously measured with infinite accuracy (zero error) for a microscopic particle. The product of uncertainties in the simultaneous measurement of conjugate variables always has a value above a certain minimum (which is approximately equal to Planck's constant).

You should note that the uncertainty relation $\Delta x \Delta p_x \geq \hbar$ has been obtained purely as a mathematical property of a wave packet. Hence this relation is as fundamental as wave-particle duality. Like wave-particle duality, the uncertainty principle, though universally applicable is of significance only for microscopic systems.

According to Eq. (5.6a), for a microscopic system there cannot be a state in which p_x as well as x have definite values. We can never simultaneously ascertain values of both position and momentum with arbitrary accuracy. If the position of the microscopic particle is defined (measured) precisely then the uncertainty in its momentum will be infinite, i.e., we will not have any idea of what its momentum is. Similarly, if we are able to precisely measure the particle's momentum, we will have no knowledge of its position. Thus, for instance, in quantum mechanics, a microscopic particle's motion cannot be described by equations like $x = a \sin \omega t$ because it implies definite velocity at a definite position. In other words, the uncertainty principle does not allow the concept of a trajectory. *Thus, unlike classical physics, a definite path of a microscopic particle with definite velocity at every point on the path is not possible in quantum mechanics.* We will take up this point once again.

In several text books you will come across the following form of the uncertainty principle

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

You should keep in mind that the right hand side expresses the order of $\hbar$. The lower limit of $\hbar/2$ for $\Delta x \, \Delta p_x$ is rarely attained; Eq. (5.6a) holds more usually or even $\Delta x \Delta p_x \geq \hbar$ holds.

Another form of the uncertainty principle finds use in atomic processes. Sometimes we might wish to measure the energy emitted in an atomic process in a time interval Δt. Then we require an uncertainty relation between energy and time. For this we write Eq. (5.6a) as

$$\frac{m \Delta x}{p} \frac{p \Delta p}{m} \geq \hbar$$

The first factor is $\frac{\Delta x}{v}$ or Δt, and since $E = \frac{p^2}{2m}$, $\Delta E = \frac{p \Delta p}{m}$.

Thus, we get

$$\Delta E \Delta t \geq \hbar \tag{5.7a}$$

where Δt is the uncertainty in the time localizability of the wave packet and ΔE is the uncertainty in its energy. A more precise calculation based on the nature of wave packets changes this result to

$$\Delta E \Delta t \geq \frac{\hbar}{2} \tag{5.7b}$$

Eqs. (5.7a and b) tell us that *in order for a microscopic particle to have a well defined energy state, the state must last for a very long time — it must be stationary.* If the energy state is shortlived, e.g., the excited state of an atom, its energy is uncertain. This is revealed in the width of spectral lines. Suppose the excited atom having life-time Δt makes a transition to a lower state. Then according to Eq. (5.7a) the energy (or the frequency) of the radiation emitted by the atom is uncertain by an amount $\hbar/\Delta t$. Thus the radiation will not be monochromatic as $\Delta \nu = \Delta E/h$. It will contain frequencies

between $\nu + \Delta\nu$ and $\nu - \Delta\nu$. And the line width $\Delta\nu$ of the spectral line, also known as natural width, will be

$$\boxed{\Delta\nu = \frac{1}{2\pi\,\Delta t}} \qquad (5.7c)$$

Let us now consider an application of the uncertainty principle.

Example 1

Calculate the minimum uncertainty in the momentum of a ^{4}He atom confined to a 0.40 nm region.

Solution

We know only that the ^{4}He atom is somewhere in the 0.40 nm region; therefore $\Delta x = 0.40$ nm. Equation (5.7a) gives us $\Delta p_x \geq \hbar/\Delta x$. Using the equal sign to obtain the minimum, we have

$$(\Delta p_x)_{min} = \frac{\hbar}{\Delta x} = \frac{6.626 \times 10^{-34}\ \text{J s}}{2\pi \times 0.40 \times 10^{-9}\ \text{m}} = 2.64 \times 10^{-25}\ \text{kg m s}^{-1}.$$

This example gives us a reasonable picture of what happens to ^{4}He atoms at low temperatures if we try to make them stay in one region by solidifying helium. Even at temperatures approaching absolute zero, the ^{4}He atoms have considerable momentum. Since ^{4}He has a mass of 6.7×10^{-27} kg, a momentum spread of 2.64×10^{-25} kg m s^{-1} means that at some time the ^{4}He atom probably has a momentum of at least that much, or a speed of at least

$$v = \frac{\Delta p}{m} = \frac{2.64 \times 10^{-25}\ \text{kg m s}^{-1}}{6.7 \times 10^{-27}\ \text{kg}} = 394\ \text{m s}^{-1}$$

which is over 1400 km h^{-1}! So even as $T \to 0$ K, this large zero-point motion persists because of the Heisenberg uncertainty principle. The associated kinetic energy is so large that ^{4}He will not solidify even as $T \to 0$ K, unless more than 20 atm of external pressure are applied. This pressure pushes the atoms close enough together so that their attractive binding forces will be large enough to hold the solid crystal together.

You may now like to work out an SAQ.

Spend 10 min

SAQ 2

(a) The average life time of an excited atom is about 10^{-8}s. What is the order of the natural width ($\Delta\nu$) of the line emitted by the atoms?

(b) The radius of an atomic nucleus is typically 5×10^{-15}m. What is the lowest limit of the energy that an electron must have to be in the atomic nucleus?

You should understand that the (theoretical) limits set by the uncertainty principle have nothing to do with the accuracy of our measuring instruments. Even the most sophisticated instruments shall be limited by the uncertainty principle. This concept was found difficult to accept by many a leading scientist, including Albert Einstein. Hence, a number of thought (ideal) experiments were proposed and debated in the Solvay Congress held at Brussels in 1930 to disprove the above principle but without any success. The analysis of some of the thought experiments illustrates very well the physical implications of the principle. Therefore, we discuss them briefly in the next section.

5.3.1 Some Thought Experiments

We will describe here some thought experiments that help us understand the uncertainty principle better. All these attempts reflect a search for ways of violating this principle.

In this direction, they seek to determine the position and momentum of a microscopic particle to an arbitrary accuracy.

Measurement of the position of an electron: The gamma ray microscope

Let us consider a conceptual experiment first discussed by Heisenberg which attempts to measure the position of electron as accurately as possible. This experiment consists of an arrangement in which an electron is illuminated and its image is observed through a microscope (Fig. 5.4). Electrons travel in a given direction (the positive x-direction) in the form of a well defined monoenergetic beam, i.e., the velocity of the electrons is known exactly. The position of an electron can be located by observing the light (photons) scattered by the electron into the microscope. Clearly, the precision with which the position of an electron can be determined is equal to the resolving power of the microscope. Thus it is equal to the minimum distance by which the microscope can resolve two objects, i.e.,

$$\Delta x = \lambda/\sin\phi$$

where λ is the wavelength of the photon used to observe the electron and ϕ is the half angle subtended by the aperture of the microscope at the position of the electron. This result is a standard result from optics. Thus, to obtain as accurate a position measurement as possible, light of short wavelengths must be chosen, such as γ-rays.

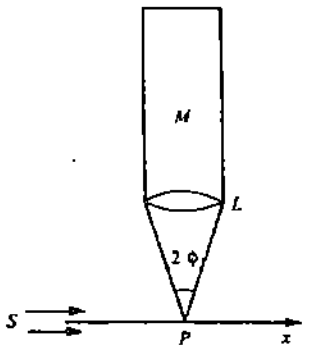


Fig. 5.4 : Position measurement of electron by Heisenberg's γ-ray microscope. Photons from a source S are scattered into a microscope M from an electron located at P.

Now in order that an electron be observed, it should scatter at least one photon into the microscope. In the process of scattering the photon would transfer momentum to the electron, causing it to recoil. For instance, if the photon is scattered by 90°, the momentum imparted to the recoiled electron along x-direction would be equal to that of the incident photon which is given by h/λ. But the photon can be scattered at any angle between 0° and ϕ. Hence, the x-component of its momentum after scattering can lie anywhere between 0 and $p \sin\phi$, where p is its total momentum. Since momentum is conserved, the magnitude of the electron's recoil momentum along x-direction is uncertain by the same or a greater amount, i.e.,

$$\Delta p_x \geq p \sin\phi = \frac{h}{\lambda}\sin\phi$$

Thus the product of the two uncertainties is

$$\Delta x\, \Delta p_x \geq h$$

which is consistent with Heisenberg's uncertainty relation. It is evident that by taking λ small enough (i.e., by using γ-rays) Δx may be made quite small. This, however, would increase Δp_x such that the product of Δx and Δp_x is always finite and given by the uncertainty relation.

To gain further insight into the uncertainty relation let us look at one of the most famous of these thought experiments : the *single slit diffraction experiment*.

Single slit diffraction experiment

Consider a highly collimated beam of photons moving along the x-direction, such that $p_x = p_0 = h/\lambda$ and $p_y = 0$. Let the beam be incident upon a single slit of width d (Fig. 5.5). The photons are diffracted by the slit and the diffraction pattern is shown in Fig. 5.5.

Since the slit is of finite width d, the position of the photons along the y-direction is uncertain by an amount d, i.e., $\Delta y = d$. What can we say about the component of their momentum in the y-direction?

All we know is that the photon will arrive at the screen somewhere within the diffraction pattern but we don't know where. Thus the uncertainty in momentum is given by the angular spread of the pattern. Since most of the photons hit the screen within the central maximum, we can obtain a rough estimate of the spread in p_y (i.e., Δp_y) by confining ourselves to the analysis of the central maximum. From Fig. 5.5, you can see that for the central maximum, p_y can take values ranging from $-p_0 \sin\theta$ to $p_0 \sin\theta$. Therefore,

$$\Delta p_y = 2p_0 \sin\theta$$

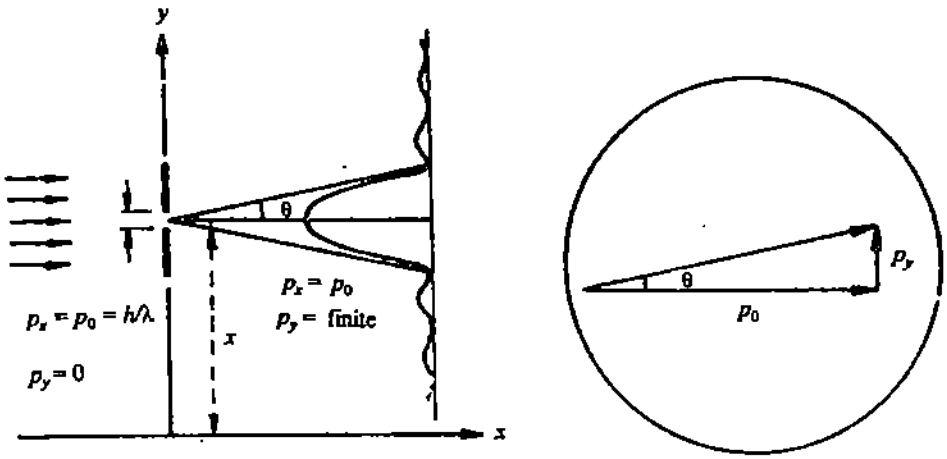


Fig. 5.5 : Single slit diffraction experiment.

Now we know from the diffraction theory that the angular spread of the pattern is inversely proportional to the width of the slit.

$$\sin\theta = \frac{\lambda}{d} = \frac{\lambda}{\Delta y}$$

where λ is the wavelength of incident light. Hence, we obtain

$$\Delta y\, \Delta p_y = d\,(2p_0 \sin\theta) = 2p_0\,\lambda$$

or $$\Delta y\, \Delta p_y = 2h \qquad \left(\because \lambda = \frac{h}{p_0}\right)$$

This is consistent with the uncertainty relation $\Delta y\, \Delta p_y \geq \hbar$. Trying to reduce the width of the slit (to reduce Δy) leads to a greater spread of diffraction pattern increasing the momentum uncertainty. Thus it is impossible to measure the position and momentum of a microscopic particle precisely at the same time.

Finally, we describe the **double slit experiment** which is another milestone in establishing the uncertainty principle.

The double slit experiment

In the double-slit experiment, a beam of monoenergetic microscopic particles (such as photons, electrons, protons etc.) are allowed to pass through two slits before falling on a fluoroscent screen placed nearby (see Fig. 5.6).

If after some time we plot the total number of particles arriving at the screen as a function of position, we observe an interference pattern. This is a characteristic of waves and can be explained as follows: The matter waves corresponding to the particle are split at the two slits and then interfere with one another. But beware of thinking of these matter waves as classical waves, because the particles do arrive at the fluoroscent screen in a particle like way: We get one localised flash everytime a particle strikes the screen. However, the totality of spots made by a large number of particles looks like the wave interference pattern. But then, is the wave-like behaviour seen only when we observe a group of particles? What happens when only one particle arrives at the slit?

To answer this question, suppose we make the particle beam very weak so that at any one instant only one particle arrives at the slit. Do we still get an interference pattern? Quantum mechanics says yes to this and experimental data seems to agree with this viewpoint. It is not easy to accept this picture. You may ask: Can a single particle split, pass through both slits and the two halves interfere with one another? Quantum mechanics says yes to all these questions. As Paul Dirac, one of the pioneers of quantum mechanics, put it, "Each photon [or a microscopic particle] interferes only with itself". Why don't we find out whether it is correct by making a measurement?

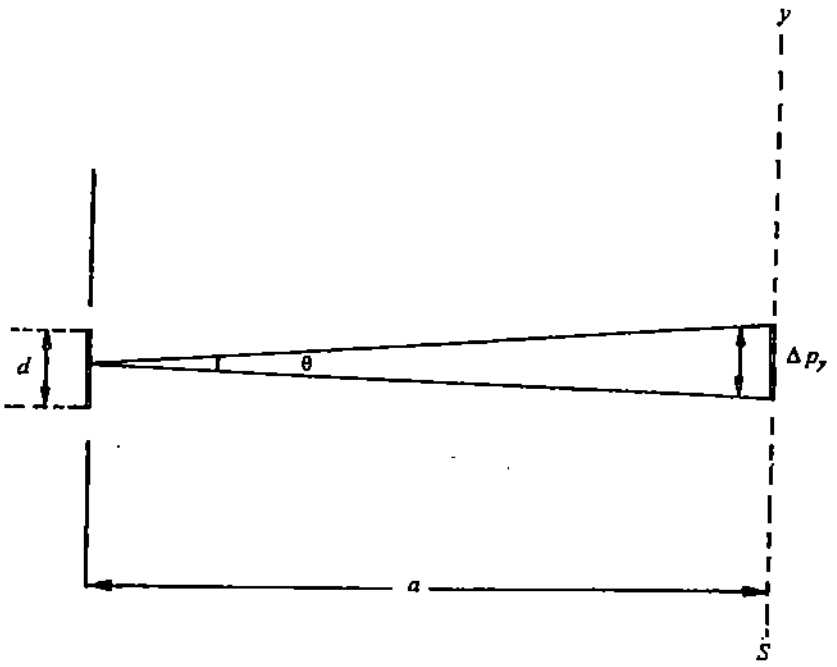


Fig. 5.6 : The double slit experiment.

The simplest way of doing this is by means of a thought experiment: Look with a flashlight! We can focus a flashlight on the slits to see which slit the particle is really passing through. What do we obtain? We find that the interference pattern is destroyed. How do we explain this effect? This effect can be explained by the uncertainty principle. As soon as we try to locate the particle and determine the slit (A or B) through which it passed, we lose information about its momentum. As we have seen in the γ-ray microscope experiment, the collision of the particle with the photon we are using to observe it, affects its momentum and introduces an uncertainty in it. Mathematically, to observe whether the particle goes through one of the slits, the photon wavelength must be smaller than at least half the distance d between the slits: ($\Delta y < d/2$) Therefore, its momentum ($=h/\lambda$) must be larger than $2h/d$ (as per the de Broglie relation). The interaction of this photon with the particle will make its momentum uncertain by an amount Δp_y given by the uncertainty relation.

This uncertainty in the particle momentum introduces an uncertainty in its position on the screen. As shown in Fig. 5.6 it is given by

$$\frac{\Delta y}{a} = \frac{\Delta p_y}{p_0} \geq \frac{2\hbar}{dp_0} = \frac{\lambda_p}{d\pi} \quad \left(\because \Delta y\, \Delta p_y \geq \hbar,\ \Delta y < d/2\right)$$

or

$$\Delta y \geq \frac{a\lambda_p}{d\pi}$$

where λ_p ($=h/p_0$) is the de Broglie wavelength of the particle. Now the condition for constructive interference is

$$d \sin\theta_n = n\lambda_p$$

so that the distance between two adjacent maxima is

$$y_m = a \sin\theta_{n+1} - a \sin\theta_n = \frac{a\lambda_p}{d}$$

Thus

$$\Delta y \geq \frac{y_m}{\pi}$$

In other words, the uncertainty in the position of the electron (produced as a result of attempting to detect it near the slit) is of the order of the distance between the two adjacent maxima. This uncertainty is enough to shift the interference pattern observed at the screen up and down in the y-direction, by a distance roughly equal to the distance between the two maxima. Such a random shift is just enough to smear out the interference pattern so that no interference is observed. So if we attempt to determine the slit through which the particle passes, the interference pattern is destroyed. The fact of the matter is that as soon as we lose information about the particle's momentum, we must also lose information about its wavelength (as per de Broglie's relation). But if there were interference fringes, from their spacing we would be able to measure the wavelength. Thus the fringe pattern cannot exist any more — the interference pattern is destroyed.

Complementarity Principle

The point is that the position and momentum measurements are really *complementary*, as Bohr first pointed out; they are mutually exclusive processes. This means that we can concentrate on the momentum and measure the wavelength of the particle from the interference pattern and hence its momentum. But then we cannot tell which slit the particle went through. Or we can concentrate on the position and lose information about the wavelength and momentum. You have seen that when we try to find out which slit the particle passes through, we lose the interference pattern and hence, the information about its wavelength and momentum. In his **complementarity principle**, introduced in 1928, Bohr described this situation by stating that the wave and particle aspects of a physical system are complementary — when we localize (find out which slit the particle goes through), we reveal the particle aspect; and when we don't localize (don't worry about which slit the particle goes through), we reveal the wave aspect. However, we cannot reveal both the aspects at the same time — they are complementary.

You may well ask: Is it that the microscopic particles are both wave and particle and we can see only one attribute with a particular experimental arrangement? That is, they possess well defined position and linear momentum at each instant but we are unable to measure them simultaneously? Or, is it that the particles just do not possess well defined position and momentum simultaneously? While Einstein was of the former view which he never gave up, Bohr and Heisenberg took the latter viewpoint. Their interpretation of quantum mechanics is also referred to as the Copenhagen interpretation. Thus, the uncertainty relation propagates the view that these uncertainties arise as a result of an inherent limitation of nature; these are intrinsic to the nature of the quantum world. However precise may be the measuring devices or the method of measurement, there is no escape from these uncertainties. The Heisenberg uncertainty principle is clearly a consequence of wave-particle duality. It reflects the fact that quantum mechanics, although a complete theory, provides a less detailed description of a physical system than does classical physics. This description is governed by the complementarity principle.

The uncertainty principle is a fundamental principle of quantum mechanics. You have noticed the role of Planck's constant — it is so small that the limitations imposed by the uncertainty principle are significant only in the domain of microscopic particles, namely atoms, molecules, subatomic, nuclear and subnuclear particles. On this scale, however, this principle is of great help in understanding many phenomena. Let us now study some interesting applications of this principle.

5.3.2 Some Applications of the Uncertainty Principle

(a) The path of an object

To define the path of a particle in an exact manner we must know its exact position and velocity simultaneously. Such a knowledge is not permitted by the uncertainty principle. Hence the path (or the orbit) of an object in quantum mechanics is not defined. This invalidates Bohr's theory of the hydrogen atom which assigns position and velocity simultaneously to the orbiting electron.

(b) The angular momentum of an object

The angular momentum L of an object is defined as a cross product of its position

r and momentum p. Since r and p are not known simultaneously, L is also uncertain. However, as you will learn later, L^2 (=L.L) can have well defined values.

(c) The size of an atom

You can use uncertainty principle even to determine the approximate size of an atom. Let us take the example of the hydrogen atom. A hydrogen atom has a proton and an electron. If we assume the size of the atom to be a then the uncertainty in the position of the electron is about a (the electron is inside the atom). Hence according to Heisenberg uncertainty principle, the uncertainty in the electron's momentum is given by $\Delta p = \hbar/a$. The total (non-relativistic) energy of the electron is equal to

$$E = \frac{p^2}{2m_0} - \frac{1}{4\pi\varepsilon_0}\frac{e^2}{a}$$

For a stable atom, E will be minimum. Hence we replace p by $\hbar/a$ and equate dE/da to zero. This yields

$$a = 4\pi\varepsilon_0 \frac{\hbar^2}{m_0 e^2} = 0.5 \text{ Å}$$

and the corresponding value of the energy E is

$$E = -\frac{1}{2}\frac{m_0 e^4}{(4\pi\varepsilon_0)^2 \hbar^2} = -13.6 \text{ eV}$$

The negative sign of the energy shows that the electron is bound to the proton. You will note that these values are in good agreement with the experimental data.

(d) The existence of electrons inside the nucleus

In SAQ 2(b) you have used the uncertainty principle to show that the electrons do not exist inside the nucleus. The size of a nucleus is of the order of one Fermi (10^{-15}m). Therefore, if electrons are present inside nucleus, then the maximum uncertainty in their position is $\Delta x = 10^{-15}$m. Hence Δp will be $\hbar/\Delta x = 10^{-19}$ Js m^{-1}. The total energy may be obtained from the relation $E^2 = p^2c^2 + m_0^2c^4$, or $E = pc$ as m_0c^2 is much smaller than pc. Thus we obtain

$$E = 3 \times 10^{-11} \text{ J} = (3/1.6) \times 10^8 \text{ eV} = 200 \text{ MeV}.$$

However, experimentally we find that during β-decay of a nucleus, electrons of energies between 2–3 MeV are ejected. Hence we conclude that electrons were not present in the nucleus before the decay.

(e) Zero point energy

According to kinetic theory, the kinetic energies of atoms oscillating about their positions in crystals are proportional to the absolute temperature. Hence, at absolute zero, the atoms, according to this theory, would stop oscillating and would remain fixed in their lattice position. But, according to uncertainty relation both position and the momentum cannot be specified at the same time with complete accuracy. This means that the atomic oscillators even at absolute zero would retain a certain amount of oscillatory motion enough to obey the uncertainty relation. The energy possessed by the atomic oscillator at absolute zero is termed as **zero point energy** Experimental studies of the motion of the atom at a temperature (0.001 K), quite close to absolute zero, have shown the reality of the zero-point energy.

You may like to end this section with an SAQ.

SAQ 3 *Spend 5 min*

A linear harmonic oscillator of mass m oscillates with a frequency $\nu = \frac{1}{2\pi}\sqrt{\frac{k}{m}}$, where k is its force constant. Use the uncertainty principle to show that the minimum energy of the oscillator is $h\nu/2$.

Let us now summarise what you have studied in this unit.

5.4 SUMMARY

- Wave-particle duality and the localization of the particles leads to the representation of a particle by a group of waves called a **wave packet. The group velocity** v_g of the wave packet is equal to the particle velocity v and the **phase velocity** v_p is given by c^2/v.
- The concept of a wave packet leads to **Heisenberg's uncertainty principle** according to which two canonically conjugate variables like x and p_x or E and t cannot be simultaneously determined with perfect accuracy. The product of the uncertainties associated with these variables, i.e., $\Delta x\, \Delta p_x$ and $\Delta E\, \Delta t$ is of the order of the Planck constant h:

  $$\Delta x\, \Delta p_x \geq \hbar,$$

  $$\Delta E\, \Delta t \geq \hbar$$

- Some of the notable consequences of the uncertainty principle are as follows:
  - The path of a particle is not defined in quantum physics.
  - Electrons do not exist inside the nucleus.
  - Atomic oscillators possess a certain amount of energy, known as the **zero-point energy**, even at absolute zero temperature.
- Several thought experiments, such as the **γ-ray microscope experiment, the single slit diffraction experiment** and the **double slit experiment** have helped in firmly establishing the validity of the uncertainty principle.

5.5 TERMINAL QUESTIONS

Spend 30 min

1. Show that the uncertainty principle can be expressed in the form $\Delta L\, \Delta \theta \geq \hbar$, where ΔL is the uncertainty in the angular momentum of the particle and $\Delta\theta$ is the uncertainty in its angular position.
2. The radius of a hydrogen atom is 5.3×10^{-11} m. Estimate the minimum kinetic energy of the electron in this atom using the uncertainty principle.
3. An atom remains in an excited state for 10^{-8} s. Calculate the uncertainty in its energy.
4. Consider that a microscopic object is moving along the x-axis and the uncertainities in its position are Δx_0 and Δx, respectively, at $t = 0$ and $t = t$. Show that Δx is directly proportional to t and inversely proportional to Δx_0. From this problem what do you learn about the spreading of the waves associated with the motion of an object?

5.6 SOLUTIONS AND ANSWERS

Self-Assessment Questions

1. Phase velocity $v_p = \dfrac{c^2}{v_g}$

   and

   $$v_g = \frac{p}{m}$$

   where $p = (m^2c^2 - m_0^2c^2)^{1/2} = (m^2 - m_0^2)^{1/2}\, c$

   and $m = \dfrac{E}{c^2} = \dfrac{10^6 \times 1.6 \times 10^{-19}\ \text{J}}{9 \times 10^{16}\ \text{m}^2\,\text{s}^{-2}} = 1.778 \times 10^{-30}$ kg.

Therefore, $p = [(17.8)^2 - (9.11)^2]^{1/2} \times 10^{-31} \times 3 \times 10^8$ kg m s^{-1}

$= 4.58 \times 10^{-22}$ kg m s^{-1}

$$\therefore \quad v_g = \frac{4.58 \times 10^{-22} \text{ kg m s}^{-1}}{1.778 \times 10^{-30} \text{ kg}}$$

$= 2.576 \times 10^8$ m s^{-1}.

and $$v_p = \frac{9}{2.576} \times 10^8 \text{ m s}^{-1} = 3.5 \times 10^8 \text{ m s}^{-1}$$

2. (a) The order of the natural line width is

$$\Delta \nu = \frac{1}{2\pi \Delta t} = \frac{10^8}{2\pi} \text{ Hz} = 1.6 \times 10^7 \text{ Hz}$$

(b) The uncertainty in the electron's position is

$\Delta x = 5 \times 10^{-15}$ m. Therefore,

$$\Delta p \geq \frac{\hbar}{\Delta x} \geq \frac{6.626 \times 10^{-34} \text{ Js}}{2\pi \times 5 \times 10^{-15} \text{ m}} \geq 2.11 \times 10^{-20} \text{ kg m s}^{-1}$$

The momentum would also be of the same order if this is the uncertainty in it. This suggests that the K.E. of the electron is far greater than its rest energy and we can write

K.E. = pc so that

K.E. = $pc \geq (2.11 \times 10^{-20}$ kg ms$^{-1}) \times (3 \times 10^8$ ms$^{-1})$

$\geq 6.33 \times 10^{-12}$ J

≥ 39 MeV

Thus the K.E. of an electron must exceed 39 MeV for it to be a nuclear constituent. Experiments indicate that electrons in an atom have only a fraction of this energy. Thus we can conclude that electrons are not present in atomic nuclei.

3. The energy of the linear harmonic oscillator is

$$E = \frac{p^2}{2m} + \frac{1}{2} kx^2$$

This is a constant of motion. We can represent the constant value of E by means of averages of the kinetic and potential energies over a cycle of motion by writing

$$E = \frac{\langle p^2 \rangle}{2m} + \frac{1}{2} k \langle x^2 \rangle$$

The average values of x and p should vanish for an oscillating particle. So we can identify $\langle p^2 \rangle$ and $\langle x^2 \rangle$ with the squares of the corresponding uncertainties:

$$\langle x^2 \rangle = \langle x \rangle^2 + (\Delta x)^2 \equiv (\Delta x)^2$$

and $$\langle p^2 \rangle = \langle p \rangle^2 + (\Delta p)^2 = (\Delta p)^2 = \left(\frac{\hbar}{2 \Delta x} \right)^2$$

Thus

$$E = \frac{(\Delta p)^2}{2m} + \frac{1}{2} k (\Delta x)^2 = \frac{\hbar^2}{8m (\Delta x)^2} + \frac{k}{2} (\Delta x)^2$$

since from the uncertainty principle $\Delta x \Delta p \geq \hbar/2$. To determine the minimum energy of the oscillator we put

$$\frac{dE}{d(\Delta x)} = 0$$

or $$-\frac{\hbar^2}{4m\,(\Delta x)^3} + k(\Delta x) = 0$$

or $$(\Delta x)^2 = \left(\frac{\hbar^2}{4mk}\right)^{1/2}$$

The minimum energy is

$$E_{min} = \frac{\hbar^2}{8m}\left(\frac{4mk}{\hbar^2}\right)^{1/2} + \frac{1}{2}k\left(\frac{\hbar^2}{4mk}\right)^{1/2}$$

$$= \frac{\hbar}{4}\left(\frac{k}{m}\right)^{1/2} + \frac{\hbar}{4}\left(\frac{k}{m}\right)^{1/2} = \frac{\hbar}{2}\left(\frac{k}{m}\right)^{1/2} = \frac{h}{2\,(2\pi)}\left(\frac{k}{m}\right)^{1/2}$$

or $$E_{min} = \frac{h\nu}{2}, \text{ since } \nu = \frac{1}{2\pi}\left(\frac{k}{m}\right)^{1/2}.$$

Terminal Questions

1. Consider a particle moving in a circle of radius r. If Δx is the arc length corresponding to angular position $\Delta\theta$, then we can rewrite Eq. (5.6) as

   $$r\Delta\theta\; m\Delta v \geq \hbar$$

   or $$\Delta\theta\; mr\Delta v \geq \hbar$$

   But $L = mvr$ for the particle and $\Delta L = m\Delta vr$, since m and r are constant. Hence we obtain

   $$\Delta L\; \Delta\theta \geq \hbar$$

2. The uncertainty in the electron's position is

   $$\Delta x = 5.3 \times 10^{-11}\ \text{m}$$

   and

   $$\Delta p \geq \frac{\hbar}{\Delta x} = \frac{1.054 \times 10^{-34}\ \text{Js}}{5.3 \times 10^{-11}\ \text{m}} = 1.99 \times 10^{-24}\ \text{kg m s}^{-1}$$

   An electron with such a low magnitude of momentum behaves almost like a classical particle [since $\lambda = \frac{h}{p} \simeq 10^{-10}$ m] and its kinetic energy is

   $$\text{K.E.} = \frac{p^2}{2m} \geq \frac{(1.99 \times 10^{-24})^2\ \text{kg}^2\text{m}^2\,\text{s}^{-2}}{2 \times (9.1 \times 10^{-31}\ \text{kg})} = 2.2 \times 10^{-18}\ \text{J} = 13.7\text{eV}$$

3. The energy of the atom is uncertain by an amount

   $$\Delta E \geq \frac{\hbar}{\Delta t} = \frac{1.054 \times 10^{-34}\ \text{Js}}{10^{-8}\text{s}} = 1.054 \times 10^{-26}\ \text{J}$$

4. If v_g is the group velocity of the wave packet associated with the microscopic particle then at time t

   $$\Delta x = v_g t = \frac{p_0}{m}\,t = \frac{h}{\lambda_0 m}\,t$$

where λ_0 is the initial wavelength of the wave packet at time $t = 0$. This is equal to Δx_0, the uncertainty in the particle's position at time $t = 0$. Thus, we have

$$\Delta x \simeq \frac{h}{m}\frac{t}{\Delta x_0}.$$

This result tells us that Δx, i.e., the spread of the wave-packet increases with time. The narrower the packet is initially, the quicker it spreads. This is the hidden influence of the uncertainty principle. If the confinement length Δx_0 is small, the uncertainty in its momentum and hence, its velocity is large $\left(\Delta v \simeq \frac{\hbar}{m\,\Delta x_0}\right)$. This means that the wave-packet will contain many waves of high velocity much greater than the average group velocity p_0/m. Due to the fluctuation in velocity, the distance covered by the particle will also be uncertain by an amount $\Delta x(t)$, i.e., its spread will be large.

UNIT 6 SCHRÖDINGER EQUATION

Structure

6.1 Introduction
Objectives
6.2 One-dimensional Schrödinger Equation
6.3 Statistical Interpretation of the Wave Function
Probability Current Density and the Continuity Equation
Normalisation of Wave Functions
6.4 Time independent Schrödinger Equation
Boundary Conditions and Acceptable Solutions
6.5 Summary
6.6 Terminal Questions
6.7 Solutions and Answers

6.1 INTRODUCTION

In Unit 4, you have seen that a microscopic particle is essentially represented by a matter wave with its wavelength given by the de Broglie relation. How do we describe the motion of such a particle or a system of such particles? Clearly, we cannot make use of Newton's laws of motion for this purpose. So a new theoretical description is needed for the motion of quantum mechanical particles. The new theory should be consistent with the wave nature of particles. It should also reduce to Newtonian mechanics for macroscopic particles. Recall that this condition is similar to the special theory of relativity which reduces to Newtonian mechanics at velocities much smaller than the velocity of light.

In this connection, we would like to recount a story. At the end of a seminar, in 1926, on de Broglie waves, the physicist Peter Debye said to another physicist that if matter is a wave, there should be a wave equation to describe a matter wave. Debye promptly forgot about it but the other physicist, Erwin Schrödinger proceeded to discover the wave equation for matter waves. This equation is named after him as the Schrödinger equation.

In this unit you will study the one-dimensional Schrödinger equation, and learn about its solutions. We shall also discuss the physical meaning of these solutions. These solutions are acceptable only under certain conditions about which you will study towards the end of the unit. An appendix has been given at the end of the Unit to explain the basic complex algebra used in the text. In the next unit we shall introduce another way of describing quantum mechanical systems, given by Heisenberg and Dirac which makes use of operators and observables.

Objectives

After studying this unit you should be able to

- write the one-dimensional time dependent Schrödinger equation and derive the time independent Schrödinger equation from it,
- give a statistical interpretation of the wave function,
- derive the continuity equation for the probability current density,
- normalise a given wave function,
- apply the boundary conditions to a given wave function.

6.2 ONE-DIMENSIONAL SCHRÖDINGER EQUATION

You have already learnt that wave nature is an inherent property of every particle. We now need a wave equation which suitably describes the time evolution of matter waves

representing the particle. In Unit 5, you have learnt that one way of localising a particle is by constructing a wave packet. However, by solving the terminal question 4 you have also seen that a wave packet spreads with time. This means that, a wave packet cannot represent a particle.

Hence, in quantum mechanics it is postulated that

Postulate 1: Description of the system

> **Every particle (or a system of particles) is represented by a "wave function", which is a function of space coordinates and time. The wave function determines all that can be known about the system it represents.**

For one-dimensional motion of a particle, the wave function may be represented by $\psi(x, t)$. Now you may ask: What is the form of $\psi(x, t)$ in terms of x and t? To answer this question, consider a classical (macroscopic) particle moving under the influence of a force. Its dynamical behaviour is described by Newton's second law, which is a differential equation. Similarly, Maxwell's equations of classical electromagnetism are also differential equations. Since all objects exhibit particle as well as wave nature, it is natural to expect that the quantum mechanical wave function will also be a solution of a certain differential equation, involving derivatives of x and t.

The credit for discovering such a differential equation goes to Erwin Schrödinger (Fig. 6.1). How did he arrive at his equation? The answer is, he created it intuitively breaking all traditions of such wave equations. The Schrödinger equation is one of the most successful equations of quantum mechanics because it predicts results which can be verified experimentally. We will now give you some idea of how he visualised the particular form of his equation.

Fig. 6.1 : Erwin Schrödinger, 1887-1961, Austrian theoretical physicist. Another of the founders of the new quantum mechanics, he received the Nobel Prize in 1933.

Let us first state certain preconditions for establishing this equation. Firstly, for a particle of mass m, energy E and momentum p, the equation should be consistent with

(i) the de Broglie relation $\lambda = \frac{h}{p}$, and (6.1)

(ii) Planck formula $\nu = \frac{E}{h}$ (6.2)

It should also satisfy the relation

(iii) $$E = \frac{p^2}{2m} + V(x, t) \tag{6.3}$$

for all x and t, where $V(x, t)$ is the potential energy of the particle.

(iv) Finally, the Schrödinger equation must be linear in x and t. That is, if $\psi_1(x, t)$ and $\psi_2(x, t)$ are two solutions of the Schrödinger equation for a given potential energy $V(x, t)$ then any linear combination of ψ_1 and ψ_2, say, $C_1 \psi_1 + C_2 \psi_2$, with C_1 and C_2 as arbitrary constants, must also be the solution of the same Schrödinger equation. This linearity is required so that two waves may be added to produce interference. If the linearity property is to be satisfied, the Schrödinger (differential) equation must be of degree 1, i.e., the wave functions and its derivatives appearing in it should be only of the first power.

We now put Eqs. (6.1) and (6.2) in (6.3) to obtain

$$\hbar\omega = \frac{\hbar^2 k^2}{2m} + V(x, t) \tag{6.4}$$

where $\omega = 2\pi\nu$ and $k = 2\pi/\lambda$. We now consider a simpler situation where the potential energy is constant, say equal to V_0. Under such a situation, if we take the particle to be a photon then it will have a fixed wavelength and frequency (as given by Eqs. (6.1) and (6.2)) and its wavefunction as given by the electromagnetic theory (see Unit 14, PHE-07 entitled Electric and Magnetic Phenomena), will be

$$\psi(x, t) = A\, e^{i(kx - \omega t)} \tag{6.5}$$

Differentiating the above equation once with respect to time you can easily obtain

$$\omega = i\frac{1}{\psi(x,t)}\frac{\partial\psi(x,t)}{\partial t} \tag{6.6}$$

and

$$k^2 = -\frac{1}{\psi(x,t)}\frac{\partial^2\psi(x,t)}{\partial x^2} \tag{6.7}$$

Putting Eqs. (6.6) and (6.7) in Eq. (6.4) we obtain a differential equation connecting the wave function $\psi(x, t)$ and its derivatives:

$$i\hbar\frac{\partial\psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2} + V_0\,\psi(x,t) \tag{6.8}$$

The above equation has been obtained for a special case of constant potential energy V_0. However, Schrödinger made a bold extrapolation and postulated that the form of the Eq. (6.8) does not change even for a particle of mass m moving in a potential which varies with x and t. Thus,

Postulate 2: Time Evolution of a System

The time-dependent Schrödinger equation for one-dimensional motion of a particle of mass m moving in a potential $V(x, t)$ is given by

$$i\hbar\frac{\partial\psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2} + V(x,t)\,\psi(x,t) \tag{6.9}$$

While reading this discussion, did you wonder what would happen if we had a wave equation with a second order time derivative? We would then have ended up with the relativistic energy-momentum relationship. Actually, to begin with, this is what Schrödinger himself attempted. But very soon he realised that such an equation did not work for electrons — it did not give the correct spectrum for the hydrogen atom. Incidentally, the correct relativistic equation for electrons, discovered by Dirac, does retain the first-order time derivative.

You should note that Eq. (6.9) is consistent with (i) to (iv). The appearance of $\hbar$ in the Schrödinger equation is, of course, crucial. This is how Schrödinger imposed the "quantum condition" on the wave equation of matter.

The Schrödinger equation is unlike any of the wave equations you have come across so far. Recall that wave equations usually connect a second-order time derivative of the function with its second order spatial derivative. But Schrödinger's equation contains only the first-derivative with respect to time but the second derivative with respect to space. Hence, time and space coordinates are not treated on an equal footing in this equation. Thus Eq. (6.9) cannot be correct in the relativistic domain. Hence, it is a **non-relativistic time dependent Schrödinger equation.**

And, there is a price to pay for having only a first order time derivative in the wave equation. The solutions of the Schrödinger equation are not real physical waves; they are complex functions with both a real and an imaginary part. This gives rise to the problem of interpretation of the wave function. What exactly does the wavefunction $\psi(x, t)$ mean physically? We shall discuss the interpretation of ψ given by Max Born in the next section.

But before that you might like to verify the linearity property of $\psi(x, t)$. Try the following SAQ.

Special
5 min

SAQ 1

If $\psi_1(x, t)$ and $\psi_2(x, t)$ are two solutions of the Schrödinger equation (6.9), show that $a\psi_1$ and $a\psi_1 + b\psi_2$ are also solutions of Eq. (6.9), where a and b are arbitrary constants.

6.3 STATISTICAL INTERPRETATION OF THE WAVE FUNCTION

The coefficient of the time derivative of ψ in Eq. (6.9) is imaginary. Therefore, it is evident that the wave function ψ, which is the solution of (6.9), will, in general, be complex. Thus, in order to extract any physical information from $\psi(x, t)$, we must establish a quantitative connection between $\psi(x, t)$ and the observables of the particle. In 1926, Max Born proposed the following connection:

Postulate 3: Probabilistic interpretation of the wave function

If, at any instant t, a measurement is made to locate the object represented by the wave function $\psi(x, t)$, then the probability $P(x, t)\,dx$ that the object will be found between the coordinates x and $x + dx$ is

$$P(x, t)\,dx = \psi^*(x, t)\,\psi(x, t)\,dx = |\psi(x, t)|^2 dx, \tag{6.10}$$

where "*" on a function represents its complex conjugate.

You can see that $|\psi(x, t)|^2$ is the modulus square of the wave function. Here, $P(x, t) = \psi^*(x, t)\,\psi(x, t) = |\psi(x, t)|^2$ is also termed the **probability density.** To put it in words:

The probability of finding a quantum mechanical object in a small interval dx is given by the product of the modulus square of the wave function representing the object and the interval itself.

The probability of finding the particle within some finite length $L = (x_2 - x_1)$ is given by

$$P_L(t) = \int_{x_1}^{x_2} P(x, t)\,dx \tag{6.11}$$

Thus, according to Max Born, the Schrödinger equation gives probability waves. The wave function just tells us probabilistically where the likelihood of finding the particle will be greater: there the wave will be strong, its amplitude will be larger. If the probability of finding a particle in a region is small, the wave will be weak and its amplitude will be small. It may seem from the above probabilistic interpretation that the phase of the wave function is not important, since it is the modulosquare $|\psi(x, t)|^2$ that we interpret as the probability. However, this is not so. We shall very briefly discuss this aspect.

To visualise this concept, imagine you are in a metropolis like Delhi and looking for traffic jams on its roads from a helicopter. If the vehicles were described by Schrödinger waves, we would say that the wave was strong at the location of traffic jams. Elsewhere, the wave would be weak.

The phase of the wave function

From the linearity property of Schrödinger equation and SAQ 1, you know that if ψ_1 and ψ_2 are solutions of the Schrödinger equation, the linear combination

$$\psi(x, t) = a_1\,\psi_1(x, t) + a_2\,\psi_2(x, t) \tag{6.12}$$

is also a solution of the Schrödinger equation where a_1 and a_2 are arbitrary complex numbers. This is the **superposition principle.** Now, calculate $|\psi(x, t)|^2$ using Eq. (6.5) with a phase difference, say ϕ, between ψ_1 and ψ_2. You will see that it depends on the relative phase of ψ_1 and ψ_2. Such a superposition is called **coherent superposition** — it gives rise to the interference of matter waves. Thus, the phase of the wave function is important and cannot be ignored.

You should also notice a crucial difference between the use of probability in classical physics and in quantum physics. In classical physics, probabilities add as you have studied in Unit 5 of the physics elective PHE-04 (Mathematical Methods in Physics-I). But in quantum physics, *the probability amplitudes add*, as in Eq. (6.12) and then we calculate the probabilities from Eq. (6.10) giving rise to interference.

Let us come back to the probabilistic interpretation. Since the object must always be somewhere in space, the **total probability of finding it in the whole space is unity.** We obtain this by integrating the probability over all space:

$$\int_{-\infty}^{\infty} \psi^*(x, t)\,\psi(x, t)\,dx = 1, \text{ at each instant of time } t \tag{6.13}$$

The interpretation of the wave function given by Eqs. (6.10) to (6.13) requires that ψ should be **finite** and **single-valued** everywhere, otherwise the probability of finding an object in a region of space will not be finite and unique. Further, Eq. (6.13) requires that we restrict the wave functions used in quantum mechanics to the class of **square integrable functions** for which

$$\int |\psi(x, t)|^2\,dx < \infty$$

i.e., $\psi(x, t)$ must approach 0 as $x \to \infty$ at least as fast as $x^{-1/2-\epsilon}$, with $\epsilon > 0$ and arbitrarily small. In addition, certain continuity conditions need to be imposed on ψ. We will discuss these in Sec. 6.4.1 in detail. Meanwhile, let us further explore the meaning of ψ.

6.3.1 Probability Current Density and the Continuity Equation

Since Eq. (6.13) is true for every t, *the total probability is conserved.* But this can be accomplished only when *the probability is conserved at each point, and at all times.* Let us examine this aspect in some detail.

Let us consider the concrete example of a fluid moving in the positive x-direction between two points $x = x_1$ and $x = x_2$ with a velocity v which changes with x. Let $\rho(x)$ denote the mass per unit length of the fluid around the point x. The quantity $S_x = v(x)\,\rho(x)$ is then the mass of the fluid crossing a given point x per unit time. What is the net mass accumulated per unit time in between the two points $x = x_1$ and $x = x_2$? Clearly, it is equal to $S_{x=x_1} - S_{x=x_2}$. And if the fluid is to be conserved in this region then this should be equal to the rate of change of mass in this region. Thus

$$\frac{\partial}{\partial t}\int_{x_1}^{x_2} \rho(x)\,dx = S_{x=x_1} - S_{x=x_2} \tag{6.14}$$

We can extend this analogy to the probabilistic interpretation. We say that *if the total probability is to be conserved, the conservation equation should look like* Eq. (6.14) *where* ρ *should be replaced by the probability density* $P(x, t)$ *and* S_x *by a function which we term the* **probability flux** *or the* **probability current density** $S(x, t)$. Thus, we must have

$$\frac{\partial}{\partial t}\int_{x_1}^{x_2} P(x, t)\,dx = S(x_1, t) - S(x_2, t) \tag{6.15}$$

Let us now obtain the definition of the probability flux $S(x, t)$ so that the probability conservation equation (6.15) is obeyed. For this we shall use the Schrödinger equation:

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + V\psi \tag{6.16a}$$

Here and in future we will omit the arguments of ψ and V as long as it does not create any confusion. The complex conjugate of Eq. (6.16a) is

$$-i\hbar\frac{\partial\psi^*}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi^*}{\partial x^2} + V\psi^* \tag{6.16b}$$

where we have assumed that V is real so that $V^* = V$. You can now multiply Eq. (6.16a) from the left by ψ^* and Eq. (6.16b) by ψ and then subtract (6.16b) from (6.16a) and obtain

$$i\hbar\left(\psi^*\frac{\partial\psi}{\partial t} + \psi\frac{\partial\psi^*}{\partial t}\right) = -\frac{\hbar^2}{2m}\left(\psi^*\frac{\partial^2\psi}{\partial x^2} - \psi\frac{\partial^2\psi^*}{\partial x^2}\right)$$

Now carry out a simple algebraic manipulation of the above equation to show that

$$\frac{\partial(\psi^*\psi)}{\partial t} = -\frac{\hbar}{2mi}\frac{\partial}{\partial x}\left(\psi^*\frac{\partial\psi}{\partial x} - \psi\frac{\partial\psi^*}{\partial x}\right) \tag{6.17}$$

We now integrate Eq. (6.17) with respect to x from x_1 to x_2 and get

$$\frac{\partial}{\partial t}\int_{x_1}^{x_2}\psi^*\,\psi\,dx = -\frac{\hbar}{2mi}\left(\psi^*\frac{\partial\psi}{\partial x} - \psi\frac{\partial\psi^*}{\partial x}\right)\Bigg|_{x_1}^{x_2} \tag{6.18}$$

A comparison of Eqs. (6.15) and (6.18) shows that the **probability density** $P(x, t)$ and the **probability flux** or **probability current density** $S(x, t)$ should be defined as

$$P(x, t) = \psi^*(x, t)\,\psi(x, t)$$

and

$$S(x, t) = \frac{\hbar}{2mi}\left(\psi^* \frac{\partial \psi}{\partial x} - \psi \frac{\partial \psi^*}{\partial x}\right) \tag{6.19}$$

We may rewrite Eq. (6.18) in terms of P and S as

$$\frac{\partial P(x, t)}{\partial t} + \frac{\partial S(x, t)}{\partial x} = 0 \tag{6.20}$$

Doesn't this equation look familiar to you? It has the form of a *continuity equation* analogous to the continuity equation between the charge density and current density in electrodynamics. This is one reason why $P(x, t)$ is referred to as the probability density and $S(x, t)$ is called the probability current density. In this equation, the first term $\frac{\partial P}{\partial t}$ denotes the rate of change of probability density in a certain fixed length. The second term denotes the net outward flux coming out of the same length. Eq. (6.20) then says that *the time rate of change of probability density (which is a negative quantity because as t increases ∂P/∂t decreases) is numerically equal to the net outward flux* (an outward flux is always assumed to be positive quantity). This means that within the above length, the *particles are neither created nor destroyed* (i.e., there are neither sources nor sinks)

Eq. (6.20) is, therefore, *the conservation law expressing the fact that a change in the particle density in a region of space is compensated for by a net change of flux from that region.* You can now also see why we need to impose continuity conditions on ψ: both ψ and its derivative $\frac{\partial \psi}{\partial x}$ must be finite and continuous for all values of x provided $V(x)$ is finite. That these restrictions on the solutions are necessary may be judged from the following considerations.

The probability density $P(x)$ and the probability flux $S(x, t)$ represent physical quantities and, therefore, have to be well defined. If $\psi(x)$ or its first derivative $\psi'(x)$ were not finite for some values of x then $P(x)$ and/or $S(x, t)$ would not be well defined for all values of x. Further, both $\psi(x)$ and $\psi'(x)$ must be continuous. Otherwise $S(x, t)$ would be singular at some points and these points would act as sources or sinks of probability current. In other words, creation or destruction of matter would take place. This, as you know, is impossible in non-relativistic physics.

Before proceeding further, we would like to point out that Eq. (6.20) is obtained under the condition that V is real. By writing $\psi = \psi_R + i\psi_I$ it is easy to see that both P and S are real and

$$S(x, t) = \frac{\hbar}{m}\ \mathrm{Im}\left(\psi^* \frac{\partial \psi}{\partial x}\right) = \mathrm{Re}\left(\psi^* \frac{\hbar}{im} \frac{\partial \psi}{\partial x}\right) \tag{6.21}$$

where Im(Z) denotes the magnitude of the imaginary part of Z and Re (Z) its real part. You may like to do an exercise to fix the ideas involved.

[illegible] energy [illegible] (6.20) we get

[illegible]

where V_I [illegible] potential energy

(b) The wave function of an object of mass m is given by

$$\psi(x, t) = e^{-(\alpha + i\beta)x}\, e^{-i\omega t}$$

Obtain the values of $P(x, t)$ and $S(x, t)$.

The probabilistic interpretation of ψ also leads us to the concept of normalising a wave function.

6.3.2 Normalisation of Wave Functions

Recall Eq. (6.13) which tells us that since the particle must be somewhere in space, the probability integrated over all x (i.e., all space in one-dimension) must equal 1. A wave function satisfying Eq. (6.13) is said to be **normalised**.

Now in quantum mechanics we deal with two types of wave functions. For one of them the value of the integral $\int_{-\infty}^{\infty} \psi^{*\prime}(x, t)\, \psi'(x, t)\, dx$ is finite, say equal to N, where ψ' is the solution of the Schrödinger equation. *Such functions are said to be normalisable* and N is known as the **norm of the wave function**. It also means that ψ' is a *square integrable function* which vanishes as $|x| \to \infty$.

You also know that Schrödinger equation is linear. And you have shown in SAQ 1 that if ψ' is a solution of Eq. (6.9) then ψ' multiplied by a constant (independent of t and x) is also a solution. Hence we can always choose the constant to be $N^{-1/2}$ and take $\psi = N^{-1/2}\, \psi'$. Then for ψ, Eq. (6.13) is satisfied and the wave function ψ is said to be a **normalised wave function**. Note that N is independent of time (otherwise ψ will not be a solution of Eq. (6.9)). Hence, *a wave function which is normalised at any instant of time stays normalised at all other times.*

However, there exists a category of wave functions for which the value of the integral in Eq. (6.13) is infinite. Such functions do not represent a physical system in the strictest sense. But, we shall see later that such functions are extensively used in quantum mechanics to describe free particles. In fact, we have already used $e^{i(kx - \omega t)}$ to represent a free particle. It is a wave function whose norm is infinite or, in other words, it is unnormalisable. Wave functions of this form do not go to zero as $x \to \pm\infty$. We shall discuss the normalisation of such functions in detail in Block 3.

Let us now further analyse the Schrödinger Equation.

6.4 TIME INDEPENDENT SCHRÖDINGER EQUATION

The Schrödinger equation (6.9) represents the time development of the wave functions $\psi(x, t)$. We have seen that the position probability density of the particle is related to $\psi(x, t)$. We can also transform $\psi(x, t)$ in such a way that we obtain momentum probability distributions of the particle. (Such transforms are called Fourier transforms). Thus, given the position and momentum distributions of the particle at one time, the same can be obtained at any later time with the help of the Schrödinger equation. Isn't this situation analogous to the one in classical mechanics? In classical mechanics, if the position and momentum of a particle are known at some initial time then we can use Newton's equation of motion to find the position and momentum of the particle at any later time. However, there is a difference between classical and quantum mechanics. What is it? Recall the uncertainty principle which tells us that in quantum mechanics both the position and the momentum of the particle cannot be known precisely at the same instant of time. We can know only their distributions.

Richard Feynman, whose Feynman Lectures on Physics are an essential reading for every physics student, once said, "Electron waves are probability waves in the ocean of uncertainty;" now you know what that means!

In many problems of classical mechanics such as Kepler's planetary orbits, Rutherford scattering, we assume that the particle has definite energy and/or angular momentum. However, in quantum mechanics if we assume the energy of the particle to be known precisely, i.e., if $\Delta E = 0$ then, according to the uncertainty relation $\Delta E\, \Delta t \geq \hbar$, and Δt is infinite. This means that an infinite amount of time should be available to make energy measurements. In other words, the probability density $\psi^*(x, t)\, \psi(x, t)$ should not change

with time. Hence, for a system of constant energy the wave function $\psi(x, t)$ should take the form

$$\psi(x, t) = \psi(x) \exp\{ig(t)\} \tag{6.22}$$

where $g(t)$ is any function of t. For a system represented by such a wave function, the energy of the system does not change with time, i.e., its energy is conserved. Hence, if the particle is initially in a specified energy state represented by Eq. (6.22) then it stays in it indefinitely unless it is disturbed by some external agency. Such energy states are called **stationary states.**

Classically, the energy (which is also termed the Hamiltonian) is a constant of motion if it does not contain time explicitly. Now, *the Hamiltonian is the sum of the kinetic energy and the potential energy.* Thus, for the total energy to be conserved, the potential in which the particle is moving should be independent of time. For such potentials, the Schrödinger equation (6.9) may be separated in x and t as follows. Substituting

$$\psi(x, t) = \psi(x) f(t) \tag{6.23}$$

in Eq. (6.9) and rearranging the terms, we get

$$-\frac{\hbar^2}{2m}\frac{1}{\psi(x)}\frac{\partial^2 \psi(x)}{\partial x^2} + V(x) = \frac{i\hbar}{f(t)}\frac{\partial f(t)}{\partial t} \tag{6.24}$$

The left side of this equation is constant for fixed x at all t. Similarly, the right side is constant for fixed t at all values of x. Hence, Eq. (6.24) will hold only when both the sides are equal to a constant C which is independent of x and t. Thus we obtain

$$i\hbar\frac{df}{dt} = C f(t) \tag{6.25}$$

and

$$-\frac{\hbar^2}{2m}\psi''(x) + V(x)\psi(x) = C\psi(x), \tag{6.26}$$

where $\psi''(x) \equiv d^2\psi(x)/dx^2$.

You can solve Eq. (6.25) to obtain

$$f(t) = A \exp\{-iCt/\hbar\} = A \exp\{-i\omega t\} \tag{6.27}$$

where A is the normalisation constant and $\omega = C/\hbar$. From Eqs. (6.26) and (6.27) it is clear that C should have dimensions of energy and be equal to the total energy E. Eq. (6.27) is, therefore, written as

$$\boxed{-\frac{\hbar^2}{2m}\psi''(x) + V(x)\psi(x) = E\psi(x)} \tag{6.28}$$

Time independent Schrödinger equation

The above equation is known as **time independent Schrödinger equation.**

We can now write the general solution or the stationary state solution of the time-dependent Schrödinger equation for a particle of definite energy E as

$$\psi(x, t) = \psi(x) \exp(-iEt/\hbar), \tag{6.29}$$

where $\psi(x)$ satisfies Eq. (6.28). The probability density and the probability flux in such cases are given by

$$P(x) = \psi^*(x)\psi(x) \tag{6.30}$$

and

$$S(x) = \frac{\hbar}{2mi}\left[\psi^*(x)\frac{d}{dx}\psi(x) - \psi(x)\frac{d}{dx}\psi^*(x)\right] \tag{6.31}$$

It is evident that both $P(x)$ and $S(x)$ are independent of time. You should note that Eq. (6.28) contains no imaginary quantities and hence $\psi(x)$ is not necessarily complex although $\psi(x, t)$ is. The normalisation condition (6.13) for the stationary state function reduces to a form

$$\int_{-\infty}^{\infty} \psi^*(x)\, \psi(x)\, dx = 1 \tag{6.32}$$

You may now like to apply the concepts discussed so far.

SAQ 3

The wave function for a steady state is given by

$$\psi(x) = N \exp\left(-\frac{x^2}{2}\right)$$

Calculate the value of the normalisation constant N.

In this course we shall confine our study to those problems which require solutions of time independent Schrödinger equation, i.e., we shall study stationary state problems. Let us now examine the conditions which the wave function $\psi(x)$ has to satisfy in order to be a physically acceptable solution.

6.4.1 Boundary Conditions and Acceptable Solutions

Recall that the probabilistic interpretation imposes the following conditions on the wave function $\psi(x)$:

1) $\psi(x)$ should be finite and single-valued everywhere,

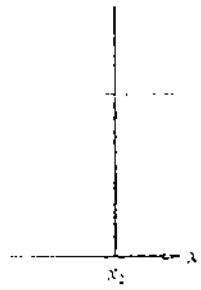
2) $\psi(x)$ should be square integrable and

3) both $\psi(x)$ and $\frac{\partial \psi}{\partial x}$ should be continuous everywhere.

We can rewrite the time independent Schrödinger equation (6.28) as

$$\frac{d^2\psi(x)}{dx^2} = \frac{2m}{\hbar^2}[V(x) - E]\,\psi(x) \tag{6.33}$$

You know that $\psi(x)$ represents a probabilistic wave satisfying Eq. (6.13). Further, whether $\psi(x)$ is an acceptable solution of Eq. (6.33) or not is also determined by $V(x)$ as well as by the boundary conditions which depend upon the nature of the problem. Let us consider an example of a particle bound in a potential well shown in Fig. 6.2. Here $V(x) > E$ for $x < x_1$ and $x > x_2$.



...cle in a potential

Classically, if the particle is initially between x_1 and x_2 then for all times to come it will be confined between the same space, i.e., the particle is bounded between x_1 and x_2. Then we say that the particle is in a bound state. Quantum mechanically we expect a large probability of finding the particle between the space $x_1 < x < x_2$. However, there also exists a decreasingly small probability of finding the particle outside this space, which is forbidden classically. This, in turn, demands for the (boundary) condition that the bound state wave functions must vanish at infinity. A very interesting result follows from the consideration of such a boundary condition which you shall study in the next block in detail. Here we just mention it: The (acceptable) solutions of the time independent Schrödinger equation exist only for certain discrete values of the total energy E. Thus *the energy quantization is an inherent property of the Schrödinger equation for the bound states.*

You have just learnt that the probabilistic interpretation puts another restriction on an acceptable (or well behaved) solution: *the wave function and its first derivative must be finite and continuous.* Eq. (6.28) shows that if $V(x)$, E and $\psi(x)$ are finite then $\psi''(x)$ is also finite. Thus, in turn, means that $\psi'(x)$ is continuous. However, if for certain values

of x, $V(x)$ becomes infinite then Eq. (6.28) yields an infinite value for $\psi''(x)$ at those points. Hence at those points $\psi'(x)$ may not be continuous.

We now end this section by summarising the properties and the boundary conditions that a wave function must satisfy to be acceptable:

Properties of the Wave Function

(1) $\psi(x)$ must be **single valued, finite** and **continuous** for all values of x.

(2) $\psi'(x)$ must be **finite** and **continuous** for all values of x, except at those points where $V \to \infty$. At these points $\psi'(x)$ has a finite discontinuity but ψ remains continuous.

(3) For bound states, the probability of finding the particle between x and $x + dx$, i.e., $|\psi|^2 dx$ must vanish as $|x| \to \infty$. Hence $|\psi(x)| \to 0$ as $|x| \to \infty$, i.e., $\psi(x)$ is a **square integrable** wave function.

Let us now summarise what you have studied in this unit.

6.5 SUMMARY

- In this unit we have concentrated on one-dimensional motion of particles. You have learnt **three postulates of quantum mechanics**:

  1. Every system can be represented by a wave **function**.

  2. The wave function satisfies a differential equation, called the **Schrödinger equation** given by

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} + V\psi$$

  3. The **probability** $P(x, t)$ of finding a particle at (x, t) in the elementary element dx is given by

$$P(x, t)\, dx = \psi^*(x, t)\, \psi(x, t)\, dx$$

  where $P(x, t)$ is the **probability density**.

- Schrödinger equation can be used to derive a **continuity equation** which connects the probability density with an associated probability current density $S(x, t)$ as follows:

$$\frac{\partial P(x, t)}{\partial t} + \frac{\partial S(x, t)}{\partial x} = 0$$

where

$$S(x, t) = \frac{\hbar}{2mi} \left(\psi^* \frac{\partial \psi}{\partial x} - \psi \frac{\partial \psi^*}{\partial x} \right)$$

The continuity equation tells us that a particle moving under a real potential is neither destroyed nor is another particle created; the change in particle density in a region equals the net change of flux into or away from that region.

- The total probability of finding a particle in the whole space always remains unity:

$$\int_{-\infty}^{\infty} \psi^*(x)\, \psi(x)\, dx = 1$$

Wave functions which satisfy this condition are said to be **normalised**.

- For a **conservative system** we can write

$$\psi(x, t) = \psi(x)\, e^{-iEt/\hbar}$$

where E is the total energy of the system and $\psi(x)$ is the solution of the **time independent Schrödinger equation** given by

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

- In order that $\psi(x)$ represent a physical system, the wave function $\psi(x)$ must be **single valued, finite** and **continuous** at all values of x. Its first derivative $d\psi(x)/dx$ must also be finite and continuous at all values of x, except at those points where $V(x) \to \infty$. At these points the first derivative has a finite discontinuity.

6.6 TERMINAL QUESTIONS

Spend 45 min

1. The wave function of an object of energy E and momentum p is given by

$$\psi(x, t) = A\, e^{i(px - Et)/\hbar}$$

   (i) Does ψ represent a bound state?

   (ii) Is the wave function normalisable?

   (iii) Using the above calculate the probability current density $S(x, t)$ in terms of the velocity v of an object and a constant A, which is complex.

2. The unnormalised wave function of a system is given by $x \exp(-x^2/2)$. Obtain the value of its normalisation constant.

3. A certain function is given by

$$\psi(x) = N(1 + ix) \exp(-x) \text{ for } x > 1$$
$$= 0 \qquad\qquad\qquad \text{for } x < 1$$

   Obtain the value of the normalisation constant N. Why can it not represent a physical system?

4. The potential energy of a simple harmonic oscillator of mass m and frequency ν is equal to $2m(\pi\nu x)^2$. Write down its time independent Schrödinger equation and show that it can be rewritten as

$$\frac{d^2\psi}{d\xi^2} + \left(\frac{\alpha}{\beta} - \xi^2\right)\psi = 0$$

where $\alpha = \dfrac{2m}{\hbar^2}E$, $\beta = 2\pi m\nu/\hbar$ and $\xi = \sqrt{\beta}\, x$, E being the total energy of the oscillator. For what values of α/β are the functions $\psi(\xi) = \exp(-\xi^2/2)$ and $\psi(\xi) = \xi \exp(-\xi^2/2)$ solutions of the above equation?

6.7 SOLUTIONS AND ANSWERS

Self-Assessment Questions

1. Substituting $\psi = a\psi_1 + b\psi_2$ in Eq. (6.9) we obtain,

$$i\hbar\left(a\frac{\partial\psi_1}{\partial t} + b\frac{\partial\psi_2}{\partial t}\right) = -\frac{\hbar^2}{2m}\left(a\frac{\partial^2\psi_1}{\partial x^2} + b\frac{\partial^2\psi_2}{\partial x^2}\right) + Va\psi_1 + Vb\psi_2$$

$$a\left(i\hbar\frac{\partial\psi_1}{\partial t}+\frac{\hbar^2}{2m}\frac{\partial^2\psi_1}{\partial x^2}-V\psi_1\right)+b\left(i\hbar\frac{\partial\psi_2}{\partial t}+\frac{\hbar^2}{2m}\frac{\partial^2\psi_2}{\partial x^2}-V\psi_2\right)=0$$

Since ψ_1 and ψ_2 satisfy Eq. (6.9), the above equation is identically zero. Hence ψ satisfies Eq. (6.9). Similarly, you can show that $a\psi_1$ is also a solution of Eq. (6.9).

2. (a) Let us express the complex potential energy V as

$$V = V_R + iV_I$$

where V_R is its real part and V_I, its imaginary part. The Schrödinger equation (6.16a) and its complex conjugate are then, respectively, written as:

$$i\hbar\frac{\partial\psi}{\partial t}=-\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2}+(V_R+iV_I)\psi$$

$$-i\hbar\frac{\partial\psi^*}{\partial t}=-\frac{\hbar^2}{2m}\frac{\partial^2\psi^*}{\partial x^2}+V_R\psi^*-iV_I\psi^*$$

Repeating the process of obtaining Eq. (6.20), we get

$$i\hbar\left(\psi^*\frac{\partial\psi}{\partial t}+\psi\frac{\partial\psi^*}{\partial t}\right)=-\frac{\hbar^2}{2m}\left(\psi^*\frac{\partial^2\psi}{\partial x^2}-\psi\frac{\partial^2\psi^*}{\partial x^2}\right)+2iV_I\psi^*\psi$$

or

$$i\hbar\frac{\partial}{\partial t}(\psi^*\psi)=-\frac{\hbar^2}{2m}\frac{\partial}{\partial x}\left(\psi^*\frac{\partial\psi}{\partial x}-\psi\frac{\partial\psi^*}{\partial x}\right)+2iV_I\psi^*\psi$$

$$\frac{\partial P(x,t)}{\partial t}+\frac{\partial S(x,t)}{\partial x}=\frac{2V_I}{\hbar}P(x,t)$$

(b) $P(x,t) = \psi^*\psi$

$$= e^{-(\alpha-i\beta)x}e^{i\omega t}e^{-(\alpha+i\beta)x}e^{-i\omega t}$$

$$= e^{-2\alpha x}$$

Using Eq. (6.21) we obtain

$$S(x,t) = \frac{\hbar}{m}\operatorname{Im}\left[\psi^*\frac{\partial\psi}{\partial x}\right]$$

$$= \frac{\hbar}{m}\operatorname{Im}\left[e^{-(\alpha-i\beta)x}e^{i\omega t}(-\alpha-i\beta)e^{-(\alpha+i\beta)x}e^{-i\omega t}\right]$$

$$= \frac{\hbar}{m}\operatorname{Im}\left[e^{-2\alpha x}(-\alpha-i\beta)\right] = -\frac{\hbar\beta}{m}e^{-2\alpha x}$$

3. The normalisation condition is

$$\int_{-\infty}^{\infty}\psi^*\psi\,dx = 1$$

or

$$|N|^2\int_{-\infty}^{\infty}e^{-x^2}\,dx = 1$$

or

$$2\,|N|^2 \int_0^\infty e^{-x^2}\,dx = 1 \text{ since } e^{-x^2} \text{ is an even function.}$$

By making the substitution $x^2 = t$, the integral takes the form

$$|N|^2 \int_0^\infty t^{-1/2}\, e^{-t}\,dt. \text{ Thus we have}$$

$$|N|^2 \int_0^\infty t^{-1/2}\, e^{-t}\,dt = 1$$

or $$|N|^2 \cdot \Gamma\left(\frac{1}{2}\right) = 1 \qquad \Gamma\left(\frac{1}{2}\right) = \int_0^\infty t^{-1/2}\, e^{-t}\,dt$$

$$= \sqrt{\pi}$$

or $$|N|^2 \sqrt{\pi} = 1$$

or $$|N|^2 = \frac{1}{\sqrt{\pi}}$$

and $$N = \left(\frac{1}{\pi}\right)^{1/4}$$

Terminal Questions

1. (i) No, it does not as $\psi(x)$ does not go to zero as $|x| \to \infty$

   (ii) No, its norm is infinite.

   (iii) $S(x, t) = \frac{\hbar}{2mi}\left[A^* A e^{-i(px - Et)/\hbar}\left(\frac{ip}{\hbar}\right) e^{i(px - Et)/\hbar}\right.$

   $\left. - AA^* e^{i(px - Et)/\hbar}\left(-\frac{ip}{\hbar}\right) e^{-i(px - Et)/\hbar}\right]$

   $= \frac{\hbar}{2mi} AA^* \left[\frac{2ip}{\hbar}\right]$

   $= v\,AA^*$ ($\because p = mv$)

2. Applying the normalisation condition we get

$$|N|^2 \int_{-\infty}^{\infty} x^2 e^{-x^2}\,dx = 1$$

or

$$2\,|N|^2 \int_0^\infty x^2 e^{-x^2}\,dx = 1 \text{ since the integrand is even}$$

Substituting $x^2 = t$, we get

$$|N|^2 \int_0^\infty t^{1/2} e^{-t}\,dt = 1$$

or

$$|N|^2 \frac{\sqrt{\pi}}{2} = 1, \text{ since } \int_0^\infty t^{1/2} e^{-t}\,dt = \frac{\sqrt{\pi}}{2}$$

Since $\Gamma(n + 1) = n\,\Gamma(n)$, $\Gamma\left(\frac{3}{2}\right) = \frac{1}{2}\,\Gamma\left(\frac{1}{2}\right) = \frac{\sqrt{\pi}}{2}$

$$\therefore |N|^2 = \frac{2}{\sqrt{\pi}} \text{ and } N = \left(\frac{4}{\pi}\right)^{1/4}.$$

3. The normalisation condition is

$$\int_{-\infty}^{\infty} \psi^* \psi \, dx = 1$$

or

$$|N|^2 \int_1^{\infty} (1 + ix)(1 - ix) \exp(-2x) \, dx = 1$$

since

$$\psi(x) = 0 \text{ for } x < 1,$$

or

$$|N|^2 \int_1^{\infty} (1 + x^2) e^{-2x} \, dx = 1$$

or

$$|N|^2 \int_1^{\infty} e^{-2x} \, dx + |N|^2 \int_1^{\infty} x^2 e^{-2x} \, dx = 1$$

or

$$|N|^2 \frac{e^{-2}}{2} + |N|^2 \int_1^{\infty} x^2 e^{-2x} \, dx = 1$$

Integrating the second term by parts, we get

$$|N|^2 \left(\frac{e^{-2}}{2} + \frac{5e^{-2}}{4}\right) = 1$$

or

$$|N| = 2e/\sqrt{7}$$

This wave function cannot represent a physical system because it is discontinuous at $x = 1$. Check this out by taking the limit $x \to 1$ from right ($+\infty$) and left ($-\infty$). Both the limits are unequal.

4. The time independent Schrödinger equation for a simple harmonic oscillator is

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + 2m(\pi\nu x)^2 \psi = E\psi \qquad (1)$$

where we have substituted $V(x) = 2m(\pi\nu x)^2$. Changing the variable to $\xi = \sqrt{\beta}\, x$, where $\beta = \frac{2\pi m\nu}{\hbar}$, we get

$$\frac{d\psi}{dx} = \frac{d\psi}{d\xi} \frac{d\xi}{dx} = \sqrt{\beta} \frac{d\psi}{d\xi}$$

$$\frac{d^2\psi}{dx^2} = \beta \frac{d^2\psi}{d\xi^2}$$

Thus, (1) becomes

$$-\frac{\hbar^2}{2m} \beta \frac{d^2\psi}{d\xi^2} + 2m\pi^2\nu^2 \frac{\xi^2}{\beta} \psi = E\psi$$

or

$$-\frac{\hbar^2}{2m} \frac{2\pi m\nu}{\hbar} \frac{d^2\psi}{d\xi^2} + \frac{2m\pi^2\nu^2\hbar}{2\pi m\nu} \xi^2 \psi = E\psi$$

or $$\frac{d^2\psi}{d\xi^2} - \frac{\pi\nu\hbar}{\pi\nu\hbar}\xi^2\psi = -\frac{2m}{\hbar^2\beta}E\psi$$

or $$\frac{d^2\psi}{d\xi^2} + \left(\frac{2mE}{\hbar^2\beta} - \xi^2\right)\psi = 0$$

Defining $\alpha = \frac{2mE}{\hbar^2}$, we can write the equation as

$$\frac{d^2\psi}{d\xi^2} + \left(\frac{\alpha}{\beta} - \xi^2\right)\psi = 0 \qquad \text{(A)}$$

Substituting $\psi(\xi) = \exp\left(-\frac{\xi^2}{2}\right)$ in this equation, we get:

$$-\psi + \xi^2\psi + \frac{\alpha}{\beta}\psi - \xi^2\psi = 0 \quad \left[\because \frac{d^2\psi}{d\xi^2} = -\psi + \xi^2\psi\right]$$

or $$\left(\frac{\alpha}{\beta} - 1\right)\psi = 0$$

or $$\frac{\alpha}{\beta} = 1 \text{ since } \psi \neq 0$$

Similarly, substituting $\psi(\xi) = \xi \exp(-\xi^2/2)$ in (A), we get:

$$-3\psi + \xi^2\psi + \left(\frac{\alpha}{\beta} - \xi^2\right)\psi = 0$$

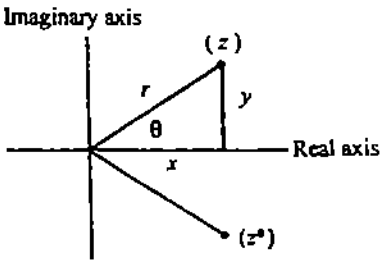
or $$\left(\frac{\alpha}{\beta} - 3\right)\psi = 0$$

or $$\frac{\alpha}{\beta} = 3 \text{ since } \psi \neq 0.$$

Appendix Complex Numbers

Imaginary number	$i = \sqrt{-1}$
Complex number	$z = x + iy$ (Cartesian form)
Real part	$x = \text{Re } z$
Imaginary part	$y = \text{Im } z$
Complex conjugate	$z^* = x - iy$
Modulus	$\lvert z \rvert$ where $\lvert z \rvert^2 = zz^* = x^2 + y^2 = r^2$
Phase	θ where $\tan \theta = \frac{y}{x}$

Complex plane



Power series

$$x = r \cos \theta = r\left(1 - \frac{\theta^2}{2!} + \ldots\right)$$

$$y = r \sin \theta = r\left(\theta - \frac{\theta^3}{3!} + \ldots\right)$$

$$z = r(\cos \theta + i \sin \theta)$$

$$= r\left[1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \ldots\right]$$

Complex number	$z = re^{i\theta}$ (polar form)
Phase factors	$e^{i\theta} = \cos \theta + i \sin \theta$ and $e^{-i\theta} = \cos \theta - i \sin \theta$
Real part	$\text{Re } e^{i\theta} = \cos \theta = (e^{i\theta} + e^{-i\theta})/2$
Imaginary part	$\text{Im } e^{i\theta} = \sin \theta = (e^{i\theta} - e^{-i\theta})/2i$

UNIT 7 OBSERVABLES AND OPERATORS

Structure

7.1 Introduction
Objectives
7.2 Quantum Mechanical Operators
Properties of Operators
Expectation Values
7.3 Eigenfunctions and Eigenvalues
7.4 Ehrenfest theorem
7.5 Summary
7.6 Terminal Questions
7.7 Solutions and Answers

7.1 INTRODUCTION

In the block introduction we had said that quantum mechanics developed along two tracks. In Units 4 to 6 we have presented one of these. You have learnt about the wave-particle duality in Unit 4. In the previous unit, you have studied how Erwin Schrödinger discovered the wave equation for matter waves. which is now called the Schrödinger equation. You have also learnt that in quantum mechanics any state of a system is represented by a wave function ψ which can be obtained by solving the Schrödinger equation. You have studied the probabilistic interpretation of ψ given by Max Born, that the de-Broglie Schrödinger waves are waves of probability which also satisfy the uncertainty relation. Thus Units 4 to 6 present the *wave mechanics version* of quantum mechanics. You should, of course, realise that the existence of ψ and the Schrödinger equation which form the basis of wave mechanics are postulates. Thus quantum mechanics is based on certain postulates which are not proved but are like axioms of geometry.

In this unit, we will introduce the other track of the development of quantum mechanics which is known as *matrix mechanics* and was developed by Werner Heisenberg, Max Born and P. Jordan in the years 1925 and 1926. In this approach, *only physically observable quantities appear*. Each physical quantity is associated with an **'operator'** which can be represented by a matrix. What is an operator? You will find an answer to this question in the next section. You will learn the method of converting a classical function into a quantum mechanical operator. This method is also one of the postulates of quantum mechanics. The essential difference between classical mechanics and matrix mechanics version of quantum mechanics is this: quantum mechanical operators obey non-commutative algebra. We will explain what this means in Sec. 7.2.2 and briefly discuss commutator algebra along with some of its applications.

The two tracks (Schrödinger's wave mechanics and Heisenberg's matrix mechanics) were integrated by Paul A.M. Dirac who invented an abstract formalism for quantum mechanics in 1930. In the remaining unit we shall present some basic concepts of this unified formulation of quantum mechanics given by Dirac (see Fig. 7.1). One of the basic postulates in this formalism connects the measured value of a dynamical variable with its theoretical value obtained with the help of the wave function ψ. We introduce it in Sec. 7.2.2. Thus we shall be able to relate quantum mechanical operators to physically observable quantities.

In certain circumstances, it is possible that when an operator operates on a wave function, the result may be a multiple of the same wave function. This gives rise to what we call the **eigenvalue-eigenfunction equation** which you will study in Sec. 7.3. Finally, you will learn about the **Ehrenfest theorem** which shows the similarity as well as one of the basic differences between classical and quantum mechanics.

The concepts presented in this unit may appear too mathematical and abstract to you in the first reading. However, the formalism presented here is a very powerful and elegant way of working with quantum mechanical systems.

Objectives

After studying this unit you should be able to

- express a classical dynamical variable as a quantum mechanical operator,
- define the hermitian operator and the parity operator and apply their properties to quantum mechanical systems,
- compute the expectation value of an operator,
- derive elementary results of commutator algebra,
- calculate the eigenvalues and eigenfunctions of a given operator,
- derive and interpret Ehrenfest theorem.

7.2 QUANTUM MECHANICAL OPERATORS

Fig. 7.1 : Paul A.M. Dirac, English physicist. He was one of the pioneers of quantum mechanics. He also formulated the relativistic wave equation of quantum mechanics which predicted the existence of positron. He was awarded the Nobel Prize in 1933.

What is a quantum mechanical '*operator*'? Let us begin with an analogy to explain this idea. When you exercise, your muscles build up by the action of the exercise: Exercise changes the muscles. The action of quantum mechanical operators on functions is a bit like that of exercise on muscles: they change the functions. You know from classical mechanics that the dynamical state of a system is determined at each instant of time by the knowledge of certain physical quantities, such as the position, velocity, linear momentum, angular momentum, energy etc. of the particles constituting the system. These physical quantities are also called **dynamical variables.**

The dynamical variables associated with a system can be measured and provide information about the system at a particular point in space-time. In quantum mechanics, *all dynamical variables are represented by operators* because they bring about changes in the wave functions upon which they act.

All measurable attributes of a quantum mechanical system are called **observables**, and yet another postulate of quantum mechanics states that

Every physical observable is associated with an operator which acts on the wave function.

Postulate 4 : Description of physical quantities

Most of these dynamical variables or observables are functions of position (x), linear momentum (p) and time (t) variables. Thus, if we can represent x and p by operators we shall be able to express most of the remaining dynamical variables as operators. A method to convert these variables into quantum mechanical operators is postulated in quantum mechanics as follows:

(i) $$\boxed{x_{op}\,\psi = x\,\psi} \qquad (7.1)$$

i.e., when operator x_{op} operates upon ψ, the result is simply the multiplication of ψ by the variable x. In other words, the operator corresponding to the dynamical variable x is x itself.

(ii) $$(p_x)_{op}\,\psi = -i\hbar\,\frac{\partial \psi}{\partial x} \qquad (7.2a)$$

Thus, the momentum operator acting on the wave function results in its differentiation with respect to the conjugate position coordinate x and the result is multiplied by $-i\hbar$. Thus, the operator of p_x is $-i\hbar\partial/\partial x$.

$$\boxed{(p_x)_{op} = -i\hbar\,\frac{\partial}{\partial x}} \qquad (7.2b)$$

Remember that Eqs. (7.1) and (7.2b) are postulated, i.e., they can't be proved. The

position and momentum operators are used in the construction of operators of other dynamical variables such as angular momentum, energy etc. How do we do this? For this, we take the classical expression for any operator D in terms of x and p_x and use Eqs. (7.1 and 7.2b) to obtain its operator form $D_{op}(x, -i\hbar\partial/\partial x, t)$. Notice that the time variable has been retained as itself in the operator formalism. In quantum mechanics, time is not treated as an operator. It is a dynamical variable.

To understand this method further, consider the example of the kinetic energy of a free particle of mass m executing one-dimensional motion given by $p_x^2/2m$. The quantum mechanical operator of the kinetic energy is obtained by replacing p_x by $(p_x)_{op}$. Then using Eq. (7.2b) we get

$$(\text{K.E.})_{op} = \frac{(p_x^2)_{op}}{2m} = \frac{1}{2m}\left(-i\hbar\right)^2 \frac{\partial^2}{\partial x^2} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} \tag{7.3a}$$

Furthermore, if its potential energy is given by the function $V(x)$, then its potential energy operator will also be $V(x)$ since x_{op} is x itself:

$$[V(x)]_{op} = V(x) \tag{7.3b}$$

Now do you notice that the sum $\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right)_{op}$ is nothing but the quantum mechanical operator of the Hamiltonian which appears in the Schrödinger equation?

Thus, we obtain the Hamiltonian operator:

$$\boxed{(H)_{op} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)} \tag{7.4}$$

In this manner, you can convert most of the dynamical variables into quantum mechanical operators. An important quantum mechanical operator that you will encounter in our subsequent discussions is the **parity operator**. We will introduce it here.

The Parity Operator

Parity is a simple but very useful concept in quantum mechanics. Consider a wave function $\psi(x)$: If on changing x to $-x$, the following relationship is obtained

$$\psi(-x) = \pm\,\psi(x) \tag{7.5}$$

then we say that the function $\psi(x)$ has a definite parity. If $\psi(-x) = +\,\psi(x)$, then $\psi(x)$ is of **even** parity. On the other hand, for $\psi(-x) = -\,\psi(x)$, the parity of $\psi(x)$ is said to be **odd**. All functions which do not obey (7.5) are said to have mixed parity. The parity operation is equivalent to transforming a right-handed system of coordinates into a left-handed one. Do you recall where you have first encountered this operation? It was introduced in Unit 1 of the elective PHE-04 (Mathematical Methods in Physics-I) where you have studied about this operation in relation to vectors. The parity operator is defined by

$$\boxed{\begin{aligned} &P\,\psi(x, t) = \psi(-x, t) && (7.6a)\\ \text{and}\quad &P\,A\left[x, -i\hbar\frac{\partial}{\partial x}, t\right] = A\left(-x, i\hbar\frac{\partial}{\partial x}, t\right) && (7.6b)\end{aligned}}$$

You can readily see that the parity operator is a space inversion operator, i.e., under its operation $x \to -x$. Thus if $\psi(x)$ describes the state of a system, $P\,\psi(x)$ describes its mirror image.

We have discussed a method to obtain quantum mechanical operators from the corresponding classical expression by changing x by x, t by t and p_x by $-i\hbar\frac{\partial}{\partial x}$. However, there is no classical expression, in terms of x and p_x which changes the sign of the argument of a function by its operation. Hence, we say that *the parity operator has no classical analog*.

You should now carry out a couple of quick exercises to fix all the ideas presented so far in your mind.

SAQ 1

Spend 10 min

(a) Express the variables p_y and p_z in operator form.

(b) Write the three components of the angular momentum **L** in terms of x, y, z and p_x, p_y and p_z and thus obtain quantum mechanical operators for L_x, L_y and L_z.

You have just studied that in quantum mechanics, the measurable classical dynamical variables like position, momentum etc. are represented by operators. These operators act on a wave function and change it in some way. We have summarised the results obtained so far for ready reference.

Dynamical Variables and Corresponding Operators

Dynamical Variable	*Operator*
Position coordinate x	x
x component of momentum p_x	$-i\hbar\frac{\partial}{\partial x}$
Kinetic energy $T=\frac{p_x^2}{2m}$	$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}$
Potential energy $V(x, t)$	$V(x, t)$
Total energy $\frac{p_x^2}{2m}+V(x, t)$	Hamiltonian $H=-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}+V(x)$
Angular Momentum	
L_x	yp_z-zp_y
L_y	zp_x-xp_z
L_z	xp_y-yp_x

Let us now discuss some important properties of these operators.

7.2.1 Properties of Operators

Firstly, operators in quantum mechanics are generally linear operators. What is a linear operator? By definition, a **linear operator** satisfies the following properties:

$$D_{op}(\phi+\psi)=D_{op}\phi+D_{op}\psi \tag{7.7a}$$

$$D_{op}\,c\phi=c\,D_{op}\,\phi \tag{7.7b}$$

where c is an arbitrary complex number. In general, we can combine Eqs. (7.7a and b) and write for a linear operator:

$$D_{op}(\lambda\phi+\mu\psi)=\lambda\,(D_{op}\phi)+\mu\,(D_{op}\psi) \tag{7.8}$$

where λ and μ are complex numbers. It is easy to see that both x and p_x satisfy Eqs. (7.7). You may like to check it out.

Spend 5 min

SAQ 2

Show that x and p_x satisfy the criterion of linearity.

Secondly, in general, quantum mechanical operators do not necessarily commute. What do we mean by this? To understand it, recall that in classical mechanics, we define angular momentum **L** as $\mathbf{r} \times \mathbf{p}$ and not as $\mathbf{p} \times \mathbf{r}$, which is equal to $-\mathbf{L}$. Putting this result in a mathematical language we can say that **r** and **p** do not commute under the operation of vector product. A similar situation exists for quantum mechanical operators. If two operators A and B operate one after the other on a function ψ, then their order of operation is important. In general, $BA\psi$ is not equal to $AB\psi$ for any arbitrary ψ, i.e.

$$\left[AB - BA \right] \psi \neq 0 \tag{7.9a}$$

The expression $AB - BA$ is denoted by the commutator bracket $[A, B]$. Thus, we define the commutator of two operators as the difference $AB - BA$ and denote it by the symbol $[A, B]$:

$$\boxed{[A, B] \equiv AB - BA} \tag{7.9b}$$

and, in general,

$$[A, B] \neq 0 \tag{7.9c}$$

In other words, in general, the operators A and B do not commute with one another and the value of the commutation bracket $[A, B]$ is non-zero. What does this result mean? It means that we have to be careful about the order of operators in considering operator products in quantum mechanics. However, if the commutator of the operators A and B vanishes, A and B commute, i.e., $AB = BA$. Then we can interchange their order.

To understand these concepts better, let us take a concrete example of operators. Let us examine whether the operators x and p_x commute with one another. For this purpose we evaluate

$$[x, p_x]\,\psi = x\left(-i\hbar\,\frac{\partial \psi}{\partial x}\right) + i\hbar\,\frac{\partial}{\partial x}\,(x\,\psi) = i\hbar\,\psi$$

Since ψ is arbitrary we obtain

$$\boxed{[x, p_x] = i\hbar} \tag{7.10}$$

Thus we have found that x and p_x operators do not commute with one another and the value of the commutation bracket is $i\hbar$. This result also tells us that we have to take care of the order of these operators when we apply them on a system. For instance, if the momentum operator acts first on a system followed by the position operator, it yields a certain result. The result is different if the position operator operates first and is followed by the momentum operator. This result has an interesting fallout. Sometimes, you may come across a situation where the product of x and p occurs in a classical dynamical variable. Now in quantum mechanics, the order of operators matters. So in which order do we put x and p? In such a case, we simply symmetrize the product, i.e., we replace the variable xp_x by the operator $\frac{1}{2}(xp_x + p_x x)$:

$$xp_x \rightarrow \frac{1}{2}(xp_x + p_x x)$$

You should note that x and p_x are what are termed in classical mechanics as *canonically conjugate variables*. In classical mechanics we do not have an equation like Eq. (7.10), since x and p_x are dynamical variables which have complex numerical values. So they occur interchangeably in classical expressions of physical quantities. You

should feel completely at home with these concepts before proceeding further. So work out this exercise.

SAQ 3

5 min

(a) Show that x_{op} commutes with $(p_y)_{op}$ and $(p_z)_{op}$.

(b) Determine $[y, p_y]$ and $[z, p_z]$.

Thus you have found that the quantum mechanical operators corresponding to classical canonically conjugate position and momentum variables do not commute with one another: x does not commute with p_x, y does not commute with p_y and z does not commute with p_z. The value of the commutation bracket is $i\hbar$ in each case. Because of this non-commutability we are required to write xp_x (yp_y or zp_z) in a symmetric form while converting a dynamical variable D containing such terms into its quantum mechanical operator. Let us now make use of the definition of $[A, B]$ given by Eq. (7.9b) to derive some interesting basic results of commutator algebra.

Basic Commutator Algebra

1. The following results satisfied by operators are useful and readily proved

$$[A, B] = -[B, A] \tag{7.11a}$$

$$[A, B + C] = [A, B] + [A, C] \tag{7.11b}$$

$$[AB, C] = A[B, C] + [A, C]B \tag{7.11c}$$

and

$$[A, BC] = B[A, C] + [A, B]C \tag{7.11d}$$

You should quickly verify Eqs. (7.11a) to (7.11d) before studying further.

2. Any operator always commutes with its own power, i.e.,

$$[A^n, A] = 0, \; n = 1, 2, 3 \ldots. \tag{7.12}$$

It follows from Eqs. (7.11) and (7.12) that if $f(x)$ is an operator which can be expanded in the powers of x then

$$[f(x), p_x] = i\hbar \frac{\partial}{\partial x}(f(x)) \tag{7.13}$$

Similarly, if $f(p_x)$ can be expanded in the powers of p_x we have

$$[x, f(p_x)] = i\hbar \frac{\partial f(p_x)}{\partial p_x} \tag{7.14}$$

You may like to prove Eqs. (7.13) and (7.14) before studying further.

SAQ 4

(a) Prove Eqs. (7.13) and (7.14).

(b) Show that the parity operator commutes with

$$A(x) = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + ax^2 + bx^4$$

Now, since *observables are measurable attributes of any physical system, they are real quantities*. Therefore, they should be represented by operators which, when operating on a physical system, yield real values of the observables. In quantum mechanics, all

observables are represented by such operators, which are called **hermitian operators.** Let us now study briefly about them.

Hermitian Operators

A **hermitian operator** is defined as follows:

$$\boxed{\int \phi^* (D_{op}\psi)\, d\tau = \int (D_{op}\phi)^* \psi\, d\tau} \qquad (7.15a)$$

For a one-dimensional system the volume element $d\tau$ is simply dx and the limit of integration is from $-\infty$ to $+\infty$. However, for a three dimensional systems $d\tau$ is the volume element $dx\,dy\,dz$ and all the three variables cover the whole space, i.e., the variables vary from $-\infty$ to $+\infty$.

Integrals like Eq. (7.15a) will occur quite often in this course. Hence we adopt a short hand notation and take

$$\int \phi^* (D_{op}\psi)\, d\tau = (\phi, D\psi) \qquad (7.15b)$$

and

$$\int (D_{op}\phi)^* \psi d\tau = (D\phi, \psi) \qquad (7.15c)$$

Henceforth, we shall use the same symbol D for the dynamical variable and also for its operator if there is no confusion. The integral $(\phi, D\psi)$ is also known as **inner product** or **scalar product** of ϕ with $D\psi$. For $\phi = \psi$ and $D = I$, the *identity operator*, the integral (ψ, ψ) is known as the **norm of the wave function** ψ. You should notice that *for a normalised* ψ *the norm is equal to unity.* The norm of a wave function representing a state of a system is always positive. It can be zero only when $\psi = 0$, i.e., that state of the system does not exist.

To understand these concepts concretely, let us now consider the linear momentum operator **p** and show that it is a hermitian operator. We have $\mathbf{p} = \hat{\mathbf{l}}\, p_x + \hat{\mathbf{m}}\, p_y + \hat{\mathbf{n}}\, p_z$ where $\hat{\mathbf{l}}$, $\hat{\mathbf{m}}$ and $\hat{\mathbf{n}}$ are unit vectors along x, y and z axes, respectively. You have already proved in SAQ 2 that

$$p_x (\lambda\phi + \mu\psi) = \lambda(p_x \phi) + \mu\, (p_x\psi)$$

This result holds for p_y and p_z as well. Hence, p_x, p_y and p_z are linear operators. Now let us consider the integral

$$I = \int_{-\infty}^{\infty} \phi^* (p_{op}\psi)\, dx = -i\hbar \int_{-\infty}^{\infty} \phi^* \frac{\partial \psi}{\partial x}\, dx$$

where p_{op} stands for p_x, p_y or p_z. Integrating by parts we get

$$I = -i\hbar\ [\phi^*\psi]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \left(-i\hbar \frac{\partial \phi}{\partial x}\right)^* \psi\, dx$$

If at least one of the functions is a normalizable wave function then the first term vanishes because the normalizable wave functions go to zero at $x = \pm\infty$. Thus

$$I = \int_{-\infty}^{\infty} (p_{op}\, \phi)^*\, \psi\, dx \qquad (7.16)$$

implying that p_{op} is a hermitian operator. Thus p_x, p_y, p_z and **p** are all hermitian operators.

The position operator x is obviously linear and hermitian. Hence the *angular momentum and Hamiltonian operators are also linear and hermitian.*

The operators which satisfy Eq. (7.15a) are also known as **self adjoint** operators. Here it is useful to introduce the *adjoint* or *Hermitian conjugate* of an operator D, by the relation

$$\int \phi^*\, D\psi\, d\tau = \int (D^\dagger \phi)^*\, \psi d\tau \qquad (7.17a)$$

If $D = D^\dagger$ then the operator D is said to be **self adjoint**. You can readily compare Eqs. (7.15a) and (7.17a) and see that for a hermitian operator

$$D^\dagger = D \tag{7.17b}$$

In quantum mechanics, operators are represented by matrices. In terms of matrix algebra, the matrix representing $D^\dagger$ is obtained by taking the complex conjugate and the transpose of the matrix representing D.

Now suppose $D = AB$ then according to Eq. (7.17a)

$$\int \phi^* (AB\psi)d\tau = \int \{(AB)^\dagger \phi\}^* \psi d\tau \tag{7.18a}$$

But we can also write

$$\int \phi^* (AB\psi)d\tau = \int \phi^* \{A(B\psi)\}d\tau$$

Thus, applying Eq. (7.17a) twice we have

$$\int \phi^* (AB\psi)d\tau = \int \{(A^\dagger \phi\}^* (B\psi)d\tau = \int (B^\dagger A^\dagger \phi)^* \psi d\tau \tag{7.18b}$$

Hence, comparing Eqs. (7.18a) and (7.18b), we obtain an important result for adjoint operators which applies to hermitian operators also:

$$\boxed{(AB)^\dagger = B^\dagger A^\dagger} \tag{7.18c}$$

So far, we have introduced you to the concepts of observables and operators. We have said that every observable is associated with an operator. Now you may ask: Exactly what is the connection between observables and operators? That is what we shall discover in the next section.

7.2.2 Expectation Values

Let us consider the measurement of a dynamical variable or the observable D of a system. Keeping the system always in a particular state ψ we measure D repeatedly. In general, each individual measurement will yield a different result. Hence we take the average of these measurements $\langle D \rangle$ as the value of the dynamical variable for that particular state. Since we have always started with the same state ψ, it is reasonable to assume that knowing ψ we should be able to calculate $\langle D \rangle$. Such a relationship between ψ and $\langle D \rangle$ is provided by another postulate of the quantum mechanics. According to this postulate,

Postulate 5: The measurement postulate

The average of the measured value of D is given by

$$\langle D \rangle = \frac{\int \psi^* D\psi \, d\tau}{\int \psi^* \psi \, d\tau} = \frac{(\psi, D\psi)}{(\psi, \psi)} \tag{7.19}$$

$\langle D \rangle$ is known as the **expectation value** of the operator D.

If $\langle D \rangle$ as obtained from Eq. (7.19) comes out to be real then the dynamical variable D is said to be an **observable**. Hence we can say that

> **An observable is a dynamical variable having a real expectation value.**

Now we can understand the significance of a hermitian operator. Hermitian operators have real expectation values. To prove this result, we have

$$\langle D \rangle = \frac{\int \psi^* D\psi \, d\tau}{\int \psi^* \psi \, d\tau} = C \int \psi^* (D\psi) \, d\tau \tag{7.20a}$$

where C is some constant representing the normalisation of ψ. The complex conjugate of this equation gives

$$\langle D\rangle^* = C\int \psi\,(D\psi)^*\,d\tau \tag{7.20b}$$

The difference is

$$\langle D\rangle - \langle D\rangle^* = C\int \psi^*\,(D\psi)\,d\tau - C\int \psi\,(D\psi)^*\,d\tau$$

Using Eq. (7.15a) this becomes

$$\langle D\rangle - \langle D\rangle^* = C\int \psi^*\,(D\psi)\,d\tau - C\int \psi^*\,(D\psi)\,d\tau$$

$$= 0$$

or $$\langle D\rangle = \langle D\rangle^* \tag{7.21}$$

which means that $\langle D\rangle$ is real. Thus, we have proved that *in quantum mechanics all observables are represented by hermitian operators.*

We will now introduce you to another interesting feature arising out of the discussion so far: Representing the Schrödinger equation as an eigenvalue-eigenfunction equation.

7.3 EIGENFUNCTIONS AND EIGENVALUES

So far you have studied that, in general, when an operator D operates upon ψ we get a new function ψ'. However, under special circumstances ψ' may just be a multiple of ψ itself, i.e.,

$$D\psi = d\psi \tag{7.22}$$

where d is a complex number. Under this situation, ψ is said to be an **eigenfunction** of the operator D having d as its **eigenvalue**. Eq. (7.22) is called the **eigenvalue-eigenfunction equation** for the operator D. Now recall the time independent Schrödinger equation given in Unit 6:

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V\right]\psi = E\psi$$

Do you recognise that the left hand side is nothing but the Hamiltonian H? We can also write this equation as

$$H\psi = E\psi \tag{7.23}$$

Thus, the time independent Schrödinger equation is an eigenvalue eigenfunction equation for the operator H (see Eq. 7.4). It tells us that H operating on a special class of wave functions gives back the same wavefunctions multiplied by the eigenvalue E of H. Since H is hermitian, the eigenvalue E is real. This result can be proved for any hermitian operator satisfying an eigenvalue eigenfunction equation as follows:

From Eq. (7.15a) we have

$$(\psi, D\psi) = (D\psi, \psi)$$

Now with the help of Eq. (7.22) we get

$$d\,(\psi, \psi) = d^*\,(\psi, \psi)$$

But for a given state ψ is not zero, hence (ψ, ψ) is finite. Thus we obtain

$$d = d^*$$

Hence, the eigenvalue of a hermitian operator is always real. In this case the expectation value $\langle D\rangle$ is equal to d itself, which is real.

Using the concepts presented so far, we would like to introduce an important class of eigenfunctions namely, eigenfunctions which are **normalised to unity** and satisfy the **orthogonality property**. Such eigenfunctions are called **orthonormal eigenfunctions**. In this connection, we will also introduce another useful concept of the **degeneracy of eigenfunctions**.

Orthonormal Eigenfunctions

Suppose for a system there are more than one eigenfunctions of an operator having the same eigenvalue. Then all such functions are called **degenerate eigenfunctions**. Eigenfunctions of an operator having different eigenvalues are called **non-degenerate eigenfunctions**. Let us now take two non-degenerate eigenfunctions ϕ and ψ of a hermitian operator D, having eigenvalues d_1 and d_2, respectively:

$$D\phi = d_1\phi \quad \text{and} \quad D\psi = d_2\psi$$

Let these eigenfunctions be normalised to unity. Then from Eqs. (7.15a) and (7.22) we obtain:

$$\int \phi^* (D\psi)\, d\tau = \int (D\phi)^* \psi\, d\tau$$

or $$d_2 (\phi, \psi) = d_1(\phi, \psi)$$

or $$(d_2 - d_1)(\phi, \psi) = 0 \tag{7.24}$$

Since $d_1 \neq d_2$ we find that the inner product (ϕ, ψ) of ϕ and ψ in Eq. (7.24) is zero. Eigenfunctions having inner product equal to zero are said to be **orthogonal** to each other:

$$(\psi_i, \psi_j) = 0, \quad \text{for} \quad i \neq j \tag{7.25}$$

We can generalise this statement. If $\psi_1, \psi_2, \ldots \psi_n$ are *non-degenerate eigenfunctions of a hermitian operator, normalised to unity* then they satisfy the following *orthonormality condition*

$$(\psi_i, \psi_j) = 0 \quad \text{for } i \neq j \tag{7.26a}$$

$$(\psi_i, \psi_j) = 1 \quad \text{for } i = j \tag{7.26b}$$

Thus, the eigenfunctions are normalised to unity, and all eigenfunctions ψ_i, ψ_j satisfy the orthogonality property (7.26a) for $i \neq j$. We can make use of the Kronecker delta symbol δ_{ij} and write Eqs. (7.26a and b) in a compact form:

$$(\psi_i, \psi_j) = \delta_{ij} \tag{7.26}$$

where δ_{ij} is defined as:

$$\delta_{ij} = 0 \quad \text{for } i \neq j$$

$$\delta_{ij} = 1 \quad \text{for } i = j$$

Such eigenfunctions which satisfy Eq. (7.26) are called **orthonormal functions** and form an **orthonormal set**.

Using these ideas we can show that if ψ is a non-degenerate eigenfunction of an operator D and D commutes with another operator B then ψ is also an eigenfunction of B. To prove it let us operate B on Eq. (7.22) from the left to obtain

$$BD\psi = d(B\psi) \tag{7.27a}$$

since d is a number. Furthermore, B commutes with D, hence, we also have

$$D(B\psi) = B(D\psi) = d(B\psi) \tag{7.27b}$$

The above equation clearly shows that $(B\psi)$ is an eigenfunction of D with the same eigenvalue d. Since ψ is not degenerate with $B\psi$, the eigenfunction $B\psi$ must be a multiple of ψ, i.e.,

$$B\psi = b\psi \tag{7.28}$$

From Eq. (7.28) we conclude that ψ is also an eigenfunction of the operator B with the eigenvalue b. In general, if there are n commuting operators and ψ is a non-degenerate eigenfunction of any one of them then it is an eigenfunction of the remaining $(n-1)$ operators also. These n operators form a set of commuting operators.

We now proceed to demonstrate that *if an operator A commutes with the parity operator P, then the non-degenerate eigenfunctions of A have definite parity.*

Let

$$A(x, p_x)\,\psi(x) = \lambda\,\psi(x) \tag{7.29a}$$

Applying P to Eq. (7.29a) from left and using the condition $[P, A] = 0$ we get

$$A(x, p_x)\left\{P\,\psi(x)\right\} = \lambda\left\{P\,\psi(x)\right\} \tag{7.29b}$$

Thus we notice that both $\psi(x)$ and $P\psi(x)$ are eigenfunctions of A with the same eigenvalue. Since $\psi(x)$ is non-degenerate, the two functions $\psi(x)$ and $P\psi(x)$ can differ at the most by a constant. Hence

$$P\psi(x) = p\,\psi(x) \tag{7.30a}$$

Thus $\psi(x)$ is an eigenfunction of the parity operator with p as the eigenvalue. Applying once again the parity operator we get

$$P^2\psi(x) = p\,P\psi(x) = p^2\psi(x) \tag{7.30b}$$

But $\psi(x)$ and $P^2\psi(x)$ are identically the same. Hence $p^2 = 1$, i.e., $p = \pm 1$. Thus, $\psi(x)$ are of definite parity. For degenerate eigenfunctions it is possible to take linear combinations of $\psi(x)$ and $\psi(-x)$ to obtain eigenfunctions of definite parity. Equation (7.30b) also gives us the eigenvalues of the parity operator: these are ± 1.

We shall use the parity operators in the next block to obtain eigenfunctions and eigenvalues of some simple systems.

We end this discussion with an exercise for you.

SAQ 5

In the last section of this unit, it would not be out of place to establish a correspondence between the quantum mechanical and classical concepts. Remember that in quantum mechanics we have operators and in classical mechanics there are only dynamical variables which may be complex numbers. Thus we have to consider the expectation values of operators. Now, according to the correspondence principle we expect that the motion of a quantum object, represented by ψ, should agree with that of a classical particle whenever the distances and momenta become so large that we can ignore the uncertainty principle. When we try to explore this point, we arrive at the **Ehrenfest theorem**.

7.4 EHRENFEST THEOREM

Let us consider the rate of change of an observable D, which does not depend explicitly on time. From Eq. (7.19) we obtain for a normalized wave function ψ

$$\frac{d\langle D\rangle}{dt} = \left(\frac{\partial\psi}{\partial t}, D\psi\right) + \left(\psi, D\frac{\partial\psi}{\partial t}\right) \tag{7.31}$$

Now we use time dependent Schrödinger equation in Eq. (7.31) to replace $\frac{\partial\psi}{\partial t}$ by $\frac{1}{i\hbar}H\psi$. Thus

$$\frac{d\langle D\rangle}{dt} = -\frac{1}{i\hbar}(H\psi, D\psi) + \frac{1}{i\hbar}(\psi, DH\psi)$$

or

$$\frac{d\langle D\rangle}{dt} = \frac{1}{i\hbar}\left[\int\psi^* DH\psi\, d\tau - \int\psi^* HD\psi\, d\tau\right], \text{ since } H \text{ is hermitian } H^\dagger = H.$$

or

$$\frac{d\langle D\rangle}{dt} = \left(\psi, \frac{1}{i\hbar}[D, H]\psi\right) \tag{7.32a}$$

or

$$\frac{d\langle D\rangle}{dt} = \frac{1}{i\hbar}\langle[D, H]\rangle \tag{7.32b}$$

Let us now take D to be position operator then

$$[D, H] = [x, H] = \frac{1}{2m}[x, p_x^2] = i\hbar\frac{p_x}{m} \tag{7.33}$$

Putting Eq. (7.33) into Eq. (7.32b) we obtain

$$\frac{d\langle x\rangle}{dt} = \frac{1}{m}\langle p_x\rangle \tag{7.34}$$

Furthermore, let us take D to be the linear momentum operator. In this case

$$\frac{d\langle p_x\rangle}{dt} = \left\langle -\frac{\partial V}{\partial x}\right\rangle \tag{7.35}$$

You should note that Eqs. (7.34) and (7.35) are very similar to the equations which define linear momentum and force in classical mechanics. However, the basic difference between the two mechanics is that x, p_x and $\frac{\partial V}{\partial x}$ of classical mechanics are *replaced by their average values in quantum mechanics*. For a macroscopic system there is hardly any difference between x, p_x and $\partial V/\partial x$ and their average values. However, for microscopic systems they are quite different. As a matter of fact you have seen that for a microscopic system the precise values of x and p_x do not exist simultaneously but their average values $\langle x\rangle$ and $\langle p_x\rangle$ can be obtained.

Eqs. (7.34) and (7.35) constitute the **Ehrenfest theorem** which shows the correspondence as well as a basic difference between classical and quantum mechanics. You may like to apply these ideas and make use of Eq. (7.32b) to arrive at an interesting result.

SAQ 6

Show that when an operator commutes with the Hamiltonian, the expectation value of the observable associated with it is a constant of motion. Hence prove that the linear momentum of a system is conserved when no net force fields acts on the system.

Let us now summarise what you have studied in this unit.

7.5 SUMMARY

- In this unit you have learned about two more postulates of the quantum mechanics, in addition to the postulates given in the previous unit. According to the first of these postulates, **every observable is associated with an operator.** The operators corresponding to the dynamical variables x and p_x are x and $-i\hbar \frac{\partial}{\partial x}$.
- To construct an operator of any other dynamical variable we write that function in terms of x, p_x (in a symmetric form) and then replace p_x by $-i\hbar\partial/\partial x$.
- Most of the operators relevant to quantum mechanics are **linear** and **hermitian**, i.e.,

$$D(a\psi + b\phi) = aD\psi + bD\phi$$

$$(\psi, D\phi) = (D\psi, \phi)$$

- According to another postulate of quantum mechanics the **expectation value** of a dynamical variable D, is equal to the average value of D, obtained by the repeated measurement of D for that system in the same state. A *dynamical variable having real expectation value is said to be an* **observable**.
- For quantum mechanical operators in general $AB\psi \neq BA\psi$ and the value of the commutation bracket $[A, B] = AB - BA$ is non-zero.
- If the operation of D on ψ produces a multiple of ψ say $d\psi$ then ψ is said to be an **eigenfunction** of D having **eigenvalue** d. *The eigenvalues of a hermitian operator are real.*
- The rate of change of average $\langle x \rangle$ and $\langle p_x \rangle$ for a system of mass m and potential energy $V(x)$ are equal to $\langle p_x \rangle/m$ and $\langle -\partial V/\partial x \rangle$, respectively. These relations are called **Ehrenfest theorems** and are very similar to those obtained in classical mechanics with the difference that in classical mechanics we consider x, p_x and $\partial V/\partial x$ themselves instead of their averages.

7.6 TERMINAL QUESTIONS

Spend 45 min

1. A state of a particle of mass m is given by $e^{-\alpha r^2}$. Normalise the wave function and calculate the expectation value of the kinetic energy of the particle.
2. If for two operators A and B

$$[A, B] = 1$$

then show that $[A^2, B^2] = 2(AB + BA)$.

3. If two non-commuting operators A and B commute with their commutator $[A, B]$, show that

$$[A, B^n] = nB^{n-1}[A, B]$$

where n is an integer. Hence obtain the value of $[e^r, p_x]$.

4. If for a quantum mechanical system

$$(p_x^2/2m + V(x))\ \psi(x) = E\,\psi(x)$$

show that $\langle \text{K.E.} \rangle = \frac{1}{2} \langle x\, \partial V/\partial x \rangle$.

The expression is known as **Virial theorem.**

Hint: Start with $\langle [xp_x, H] \rangle = 0$.

5. a) Determine whether the parity operator P is hermitian or not.

   b) Show that all operators which are invariant under space inversion commute with the parity operator.

6. The Hamiltonian of a system is given by

$$H = \frac{p_x^2}{2m} + V(x)$$

and $\psi_1(x)$ and $\psi_2(x)$ are two degenerate energy eigenfunctions of the system. Show that

$$(\psi_1, (xp_x + p_x x)\ \psi_2) = 0$$

Hint: Start with $(\psi_1, [H, x^2]\ \psi_2) = 0$.

7. Show that $[L_z, L_x] = i\hbar\, L_y$. Hence prove that if o is an eigenfunction of L_z,

$$\langle L_y \rangle = \langle L_x \rangle = 0.$$

8. Suppose $\psi(x) = \sum_i c_i\, \phi_i(x)$

where $\phi_i(x)$ are eigenfunctions of a hermitian operator D with eigenvalues d_i and $(\phi_i, \phi_j) = \delta_{ij}$ then show that

$$(\psi, D\psi) = \sum_i d_i |c_i|^2$$

7.7 SOLUTIONS AND ANSWERS

Self-Assessment Questions

1. a) $p_y = -i\hbar \dfrac{\partial}{\partial y}$

$$p_z = -i\hbar \frac{\partial}{\partial z}$$

b) $L_x = yp_z - zp_y$

$$= -i\hbar y \frac{\partial}{\partial z} + i\hbar z \frac{\partial}{\partial y}$$

$$L_y = zp_x - xp_z$$

$$= -i\hbar z \frac{\partial}{\partial x} + i\hbar x \frac{\partial}{\partial z}$$

$$L_z = xp_y - yp_x$$

$$= -i\hbar x \frac{\partial}{\partial y} + i\hbar y \frac{\partial}{\partial x}$$

2. $x\,[a\psi + b\phi] = ax\psi - ix\phi$

$$p_x\,[a\psi + b\phi] = -i\hbar \frac{\partial}{\partial x}\,[a\psi + b\phi]$$

$$= -i\hbar\, a \frac{\partial \psi}{\partial x} - i\hbar\, b \frac{\partial \psi}{\partial x}$$

$$= a\, p_x \psi + b\, p_x \phi$$

Therefore, x and p_x are linear.

3. (a) $[x, p_y] = [xp_y - p_y x]$

$$= -x\, i\hbar \frac{\partial \psi}{\partial y} + i\hbar \frac{\partial \psi}{\partial y}(x\psi)$$

$$= -i\hbar\, x \frac{\partial \psi}{\partial y} + i\hbar\, x \frac{\partial \psi}{\partial y} \quad (\because x \text{ and } y \text{ are independent})$$

$$= 0$$

Since ψ is arbitrary,

$$xp_y - p_y x = 0$$

Thus operators x and p_y commute.

Similarly, we can show that x commutes with p_z.

(b) $[y, p_y] = i\hbar$

$[z, p_z] = i\hbar$

Proof is similar to that of $[x, p_x] = i\hbar$

4. (a) Since $f(x)$ can be expanded in powers of x, we may write

$$[f(x), p_x] = [\sum_{n=0}^{\infty} x^n, p_x]$$

$$= [x + x^2 + \ldots + x^n + \ldots, p_x]$$

$$= [x, p_x] + [x^2, p_x] + \ldots + [x^n, p_x] + \ldots$$

Now using Eq. (7.11c) we can write

$$[x^2, p_x] = x\,[x, p_x] + [x, p_x]x$$

$$= 2\, i\hbar\, x$$

and

$$[x^3, p_x] = x\,[x^2, p_x] + [x, p_x]x^2$$

$$= x\,[2\, i\hbar\, x] + i\hbar\, x^2$$

$$= 3\, i\hbar\, x^2$$

Similarly,

$$[x^n, p_x] = n x^{n-1} i\hbar$$

Thus, we have

$$[f(x), p_x]$$

$$= i\hbar\,[1 + 2x + 3x^2 + \ldots + n x^{n-1} + \ldots]$$

$$= i\hbar \frac{\partial}{\partial x} f(x)$$

You can prove Eq. (7.14) in the same way.

b) Now $A(x) = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + ax^2 + bx^4$

Since $A(-x) = A(x)$ hence $PA(x)\,\psi(x) = A(-x)\,\psi(-x) = A(x)\,P\,\psi(x)$. Since $\psi(x)$ is arbitrary, $A(x)$ commutes with P.

5. We have to show that

$$\left(-\frac{d^2}{dx^2} + x^2\right) f_i = \lambda_i f_i$$

and calculate λ where

(i) $f_1 = \exp\left(-\frac{x^2}{2}\right)$

and

(ii) $f_2 = x \exp\left(-\frac{x^2}{2}\right)$

(i) $\left(-\frac{d^2}{dx^2} + x^2\right) \exp\left(-\frac{x^2}{2}\right)$

$= -x^2 \exp\left(\frac{-x^2}{2}\right) + x^2 \exp\left(\frac{-x^2}{2}\right) + e^{-x^2/2}$

$= e^{-x^2/2}$ implying $\lambda_1 = 1$.

(i) $\left(-\frac{d^2}{dx^2} + x^2\right)\left(x\, e^{-x^2/2}\right)$

$= [3x\, e^{-x^2/2} - x^3 e^{-x^2/2} + x^3 e^{-x^2/2}]$

$= 3x\, e^{-x^2/2}$ implying $\lambda_2 = 3$.

Terminal Questions

1. The normalisation condition is $N^2 \int_{-\infty}^{\infty} e^{-2\alpha x^2}\, dx = 1$

or

$N^2 (\pi/2\alpha)^{1/2} = 1$ or $N = (2\alpha/\pi)^{1/4}$ $\left[\because \int_{-\infty}^{\infty} e^{-2\alpha x^2} dx = \left(\frac{\pi}{2\alpha}\right)^{1/2}\right.$

$(K.E.) = (p^2/2m) = (2\alpha/\pi)^{1/2} \left(e^{-\alpha x^2}, -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} e^{-\alpha x^2}\right)$

$= -\frac{\hbar^2}{2m} (2\alpha/\pi)^{1/2} \left(e^{-\alpha x^2}, -2\alpha (1-2\alpha x^2) e^{-\alpha x^2}\right)$

$= -\frac{\hbar^2}{2m} \left(\frac{2\alpha}{\pi}\right)^{1/2} \left[-2\alpha \left(\frac{\pi}{2\alpha}\right)^{1/2} + \frac{4\alpha^2}{4\alpha}\left(\frac{\pi}{2\alpha}\right)^{1/2}\right]$

$\because \int_{-\infty}^{\infty} e^{-2\alpha x^2} dx = \left(\frac{\pi}{2\alpha}\right)^{1/2}$ and $\int_{-\infty}^{\infty} x^2 e^{-2\alpha x^2} dx = \frac{1}{2}\left(\frac{\pi}{8\alpha^3}\right)^{1/2}$

or $(K.E.) = \frac{1}{2}\alpha\hbar^2/m.$

2. $[A^2, B] = A[A, B] + [A, B]A$ (Using Eq. 7.11c)

$= A\{B[A, B] + [A, B]B\} + \{B[A, B] + [A, B]B\}A$

$= AB + AB + BA + BA \quad \because [A, B] = 1$

$= 2(AB + BA)$

3. Let $[A, B^n] = n B^{n-1}[A, B]$ (1)

Hence from Eq. 7.11d)

$[A, B^{n+1}] = B[A, B^n] + [A, B]B^n$

$$= n B^n [A, B] + [A, B] B^n \quad \text{(Using Eq. 7.14)}$$

$$= (n+1) B^n [A, B]. \qquad (2)$$

Hence if (1) is true for n it is also true for $n+1$. Since (1) is certainly true for $n=1$ hence it is also true for $n=2$. Thus Eq. (1) is true for any n.

Now $e^x = \sum_{n=0}^{\infty} x^n/n!$

$$\therefore \quad [e^x, p_x] = \sum_{n=0}^{\infty} \frac{1}{n!} [x^n, p_x]$$

$$= \sum_{n=1}^{\infty} \frac{1}{(n-1)!} x^{n-1} [x, p_x]$$

$$= i\hbar \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} = i\hbar\, e^x.$$

4. We have $(\psi, [xp_x, H]\, \psi) = E\,(\psi, xp_x\, \psi) - E\,(\psi, xp_x\, \psi) = 0.$ (1)

Now $[xp_x, H] = [xp_x, p_x^2/2m + V(x)]$

$$= \frac{1}{2m} [xp_x, p_x^2] + [xp_x, V(x)]$$

From Eqs. (7.11b, 7.11c and 7.11d)

$$[xp_x, H] = \frac{1}{2m} [x, p_x^2]\, p_x + x\,[p_x, V(x)] \quad (\because [p_x, p_x^n] = 0, [x, V(x)] = 0$$

$$= \frac{1}{2m} [x, p_x]\, 2p_x^2 + x\,[p_x, x]\,(\partial V/\partial x)$$

$$= i\hbar\, p_x^2/m - i\hbar\, x \frac{\partial V}{\partial x}$$

$\therefore$ Using (1) we get (K.E) = $1/2\ \langle x\, \partial V/\partial x \rangle$.

5. a) We have

$$\int_{-\infty}^{\infty} \psi^*(x)\, P\, \psi(x)\, dx = \int_{-\infty}^{\infty} \psi^*(x)\, \psi(-x) dx = -\int_{\infty}^{-\infty} \psi^*(-x')\, \psi(x') dx'$$

where $x' = -x$

$$= \int_{-\infty}^{\infty} [P\psi(x')]^*\ \psi(x')\, dx'$$

Hence P is a Hermitian operator.

b) $PA(x, p_x)\, \psi(x) = A(-x, -p_x)\, \psi(-x) = A(-x, -p_x)\, P\, \psi(x) = A(x, p_x) P\, \psi(x)$

Hence $(PA(x, p_x) - A(x, p_x)P)\, \psi(x) = 0$

Since $\psi(x)$ is arbitrary we have

$$[P, A] = 0.$$

6. Since ψ_1 and ψ_2 are degenerate, we have

$$(\psi_1, [H, x^2]\, \psi_2) = 0$$

Now $[H, x^2] = [p_x^2/2m + V(x), x^2]$

$$= (1/2m)\, [p_x^2, x^2]$$

$$= (1/2m)\, 2(p_x x + x p_x]$$

$$= (1/m)\, (p_x x + x p_x] \text{ (Using the result of terminal question 2)}$$

$$\therefore \quad (\psi_1, (p_x x + x p_x)\, \psi_2) = 0.$$

7. $[L_z, L_x] = L_zL_x - L_xL_z$

Now from the definition of L_x and L_z (see SAQ 1(b)) we have

$$L_x = yp_z - zp_y \text{ and } L_z = xp_y - yp_x$$

Hence $[L_z, L_x] = [xp_y - yp_x, yp_z - zp_y]$

Using Eq. (7.11b) we get

$$[L_z, L_x] = [xp_y, yp_z] - [xp_y, zp_y] - [yp_x, yp_z] + [yp_x, zp_y]$$

Using Eqs. (7.11c and d)

$$[L_z, L_x] = y\,[xp_y, p_z] + [xp_y, y]p_z - 0 - 0 + y\,[p_x, zp_y] + [y, zp_y]p_x$$

$$= 0 - i\hbar\, xp_z + i\hbar\, zp_x$$

$$= i\hbar\,(zp_x - xp_z) = i\hbar\, L_y$$

Let $L_z\phi = m\,\phi$

But $L_zL_x - L_xL_z = i\hbar\, L_y$

$$\therefore \quad (\phi, L_z L_x\phi) - (\phi, L_xL_z\phi) = i\hbar\,\langle L_y\rangle$$

$$m\,\langle L_x\rangle - m\,\langle L_x\rangle = i\hbar\,\langle L_y\rangle$$

$$\therefore \quad \langle L_y\rangle = 0.$$

Similarly, $\langle L_x\rangle = 0.$

8. Since $\psi(x) = \sum_i c_i\,\phi_i(x)$

we have $\langle D\rangle = \sum_i \sum_j c_i^*\, c_j\, d_j\, (\phi_i, \phi_j)$

$$= \sum_i \sum_j c_i^* c_j\, d_j\, \delta_{ij}$$

$$= \sum_i |\,c_i\,|^2\, d_i.$$

since only those terms of the j series will survive for which $j = i$.

FURTHER READING

1. Concepts of Modern Physics, A. Beiser, McGraw-Hill International Book Company, 1990.
2. Introduction to Quantum Mechanics, B.H. Bransden, C.J. Joachain, ELBS, 1990.
3. Quantum Mechanics, J.L. Powell, B. Crasemann, Addison Wesley Inc, 1961.
4. A Textbook of Quantum Mechanics, P.M. Mathews, K. Venkatesan, Tata McGraw Hill Publishing Company Limited, 1987.

A PERSPECTIVE ON QUANTUM MECHANICS

In this block we have introduced you to those basic ideas and concepts which form the bulwark of the new quantum mechanics. In the process, it may have seemed to you that the entire edifice of classical physics has been turned upside down: the classical ideas of causal determinism, continuity, unambiguous and precise language descriptions lie squarely challenged. What has replaced it is an entirely new way of thinking and understanding our world. And because the behaviour of objects in the quantum world is so unlike ordinary experience, you may have found it very difficult to get used to it, in the first instance. Do not worry. It appears peculiar and mysterious to everyone who encounters it for the first time — whether a novice or an experienced physicist.

All of us know how large objects act — all of our direct experience and intuition applies to such objects. But as you have studied in this block, things on a small scale just do not act that way. Quantum objects are wave-particles represented by wave functions. Though the time-evolution of wave functions is governed by an equation of motion, its solutions give us only a probability of finding the wave-particles in a certain region at a given time. Measurement of the physical observables like position, momenta, energy, etc. (which can be determined precisely for classical objects) is governed by the uncertainty principle in the quantum mechanical world. Then there is the idea of quantum jumps (or discontinuities) in quantum mechanics. To put it in a nutshell, as per the **Copenhagen interpretation** of quantum mechanics developed through the ideas of Born, Heisenberg and Bohr, we calculate quantum objects probabilistically, we determine their attributes somewhat uncertainly and we understand them complementarily. Quantum mechanics, thus, presents a new and exciting world-view that challenges old concepts such as deterministic trajectories of motion and causal continuity. It springs unexpected surprises on us, and keeps our minds in a constant flurry of animated activity.

For those of you philosophically inclined, we present here an excerpt from Feynman's Lectures on Physics which gives us a perspective on quantum mechanics. It is a masterly reflection upon one of the most fundamental concepts of quantum mechanics — the uncertainty principle which has unendingly troubled the best of minds. Through Feynman's eyes, we, the students of physics, get to look deeply and philosophically into the nature of quantum mechanics and the nature of science. This, in our opinion, constitutes a befitting finale to an introductory foray into the world of quantum mechanics.

"Philosophical Implications

Let us consider briefly some philosophical implications of quantum mechanics. As always, there are two aspects of the problem; one is the philosophical implication for physics, and the other is the extrapolation of philosophical matters to other fields. When philosophical ideas associated with science are dragged into another field, they are usually completely distorted. Therefore we shall confine our remarks as much as possible to physics itself.

First of all, the most interesting aspect is the idea of the uncertainty principle; making an observation affects a phenomenon. It has always been known that making observations affects a phenomenon, but the point is that the effect cannot be disregarded or minimized or decreased arbitrarily by rearranging the apparatus. When we look for a certain phenomenon we cannot help but disturb it in a certain minimum way, and *the disturbance is necessary for the consistency of the viewpoint*. The observer was sometimes important in prequantum physics, but only in a rather trivial sense. The problem has been raised: if a tree falls in a forest and there is nobody there to hear it, does it make a noise? A *real* tree falling in a *real* forest makes a sound, of course, even if nobody is there. Even if no one is present to hear it, there are other traces left. The sound will shake some leaves, and if we were careful enough we might find somewhere that some thorn had rubbed against a leaf and made a tiny scratch that could not be explained unless we assumed the leaf were vibrating. So in a certain sense we would have to admit that there is sound made. We might ask; was there a *sensation* of sound? No, sensations have to do, presumably, with consciousness. And whether ants are

conscious and whether there were ants in the forest, or whether the tree was conscious, we do not know. Let us leave the problem in that form.

Another thing that people have emphasized since quantum mechanics was developed is the idea that we should not speak about those things which we cannot measure. (Actually relativity theory also said this.) Unless a thing can be defined by measurement, it has no place in a theory. And since an accurate value of the momentum of a localized particle cannot be defined by measurement it therefore has no place in the theory. The idea that this is what was the matter with classical theory *is a false position*. It is a careless analysis of the situation. Just because we cannot measure position and momentum precisely does not *a priori* mean that we *cannot* talk about them. It only means that we *need* not talk about them. The situation in the sciences is this: A concept or an idea which cannot be measured or cannot be referred directly to experiment may or may not be useful. It need not exist in a theory. In other words, suppose we compare the classical theory of the world with the quantum theory of the world, and suppose that it is true experimentally that we can measure position and momentum only imprecisely. The question is whether the *ideas* of the exact position of a particle and the exact momentum of a particle are valid or not. The classical theory admits the ideas: the quantum theory does not. This does not in itself mean that classical physics is wrong. When the new quantum mechanics was discovered, the classical people—which included everybody except Heisenberg, Schrödinger, and Born—said: "Look, your theory is not any good because you cannot answer certain questions like: what is the exact position of a particle?, which hole does it go through?, and some others." Heinsenberg's answer was: "I do not need to ask such questions because you cannot ask such a question experimentally." It is that we do not *have* to. Consider two theories (a) and (b): (a) contains an idea that cannot be checked directly but which is used in the analysis, and the other, (b) does not contain the idea. If they disagree in their predictions, one could not claim that (b) is false because it cannot explain this idea that is in (a) because that idea is one of the things that cannot be checked directly. It is always good to know which ideas cannot be checked directly, but it is not necessary to remove them all. It is not true that we can pursue science completely by using only those concepts which are directly subject to experiment.

In science we go by experiments — even conceptual experiments. And if the limitation is not of the actual measuring devices used but is set by the fundamental processes of measurement then we have to accept it.

In quantum mechanics itself there is a wave function amplitude, there is a potential, and there are many constructs that we cannot measure directly. The basis of a science is its ability to *predict*. To predict means to tell what will happen in an experiment that has never been done. How can we do that? By assuming that we know what is there, independent of the experiment. We must extrapolate the experiments to a region where they have not been done. We must take out concepts and extend them to places where they have not yet been checked. If we do not do that, we have no prediction. So it was perfectly sensible for the classical physicists to go happily along and suppose that the position—which obviously means something for a baseball—meant something also for an electron. It was not stupidity. It was a sensible procedure. Today we say that the law of relativity is supposed to be true at all energies, but somebody may come along and say how stupid we were. We do not know where we are "stupid" until we "stick our neck out," and so the whole idea is to put our neck out. And the only way to find out that we are wrong is to find out *what* our predictions are. It is absolutely necessary to make constructs.

Let us give you an idea about the debate on determinism (causality) versus free will: According to Newtonian dynamics, if the position, velocity and the forces acting on a body at any instant of time are known its 'state' at all later times can be predicted accurately. That is, if know the 'cause' we can predict the effect. This applies to any object, however large or small! Extending this to all objects in the universe it was thought that every event and its time evolution can be determined for all times to come. This determinism would apply even to human body and mind. This means that even the human mind has no free will, no freedom of choice: the future of every living creature, being a part of the mechanistic universe is completely determined

We have already made a few remarks about the indeterminacy of quantum mechanics. That is, that we are unable now to predict what will happen in physics in a given physical circumstance which is arranged as carefully as possible. If we have an atom that is in an excited state and so is going to emit a photon, we cannot say *when* it will emit the photon. It has a certain amplitude to emit the photon at any time, and we can predict only a probability for emission; we cannot predict the future exactly. This has given rise to all kind of nonsense and questions on the meaning of freedom of will, and of the ideas that the world is uncertain.

Now things are entirely different in quantum mechanics. We can only predict the probability of an event taking place and the attributes of a system are governed by the uncertainty principle. Thus, there seems to be a complete breakdown of determinism (or the cause-effect relationship in the microscopic world). This is interpreted by some as restoration of 'freedom of will' to choose an alternative in the probabilistic world which is also uncertain.

Of course we must emphasise that classical physics is also indeterminate, in a sense. It is usually thought that this indeterminacy, that we cannot predict the future, is an important quantum-mechanical thing, and this is said to explain the behaviour of the mind, feelings of free will, etc. But if the world were classical—if the laws of mechanics were classical—it is not quite obvious that the mind would not feel more or less the same. It is true classically that if we knew the position and the velocity of

every particle in the world, or in a box of gas, we could predict exactly what would happen. And therefore the classical world is deterministic. Suppose, however, that we have a finite accuracy and do not know *exactly* where just one atom is, say to one part in a billion. Then as it goes along it hits another atom, and because we did not know the position better than to one part in a billion, we find an even larger error in the position after the collision. And that is amplified, of course, in the next collision, so that if we start with only a tiny error it rapidly magnifies to a very great uncertainty. To give an example: if water falls over a dam, it splashes. If we stand nearby, every now and then a drop will land on our nose. This appears to be completely random, yet such a behavior would be predicted by purely classical laws. The exact position of all the drops depends upon the precise wigglings of the water before it goes over the dam. How? The tiniest irregularities are magnified in falling, so that we get complete randomness. Obviously, we cannot really predict the position of the drops unless we know the motion of the water *absolutely exactly*.

Speaking more precisely, given an arbitrary accuracy, no matter how precise, one can find a time long enough that we cannot make predictions valid for that long a time. Now the point is that this length of time is not very large. It is not that the time is millions of years if the accuracy is one part in a billion. The time goes, in fact, only logarithmically with the error, and it turns out that in only a very, very tiny time we lose all our information. If the accuracy is taken to be one part in billions and billions and billions—no matter how many billions we wish, provided we do stop somewhere—then we can find a time less than the time it took to state the accuracy—after which we can no longer predict what is going to happen! It is therefore not fair to say that from the apparent freedom and indeterminancy of the human mind, we should have realized that classical "deterministic" physics could not even hope to understand it, and to welcome quantum mechanics as a release from a "completely mechanistic" universe. For already in classical mechanics there was indeterminability from a practical point of view."

Table of fundamental constants

Quantity	*Symbol*	*Value*
Planck's constant	h	6.62618×10^{-34} J s
	$\hbar = \frac{h}{2\pi}$	1.05459×10^{-34} J s
Velocity of light in vacuum	c	2.99792×10^{8} m s^{-1}
Elementary charge (absolute value of electron charge)	e	1.60219×10^{-19} C
Permeability of free space	μ_0	$4\pi \times 10^{-7}$ H m^{-1} $= 1.256\,64 \times 10^{-6}$ H m^{-1}
Permittivity of free space	$\varepsilon_0 = \frac{1}{\mu_0 c^2}$	$8.854\,19 \times 10^{-12}$ F m^{-1}
Gravitational constant	G	6.672×10^{-11} N m^2 kg^{-2}
Fine structure constant	$\alpha = \frac{e^2}{4\pi\varepsilon_0 \hbar c}$	$\frac{1}{137.036} = 7.297\,35 \times 10^{-3}$
Avogadro's number	N_A	$6.022\,05 \times 10^{23}$ mol^{-1}
Faraday's constant	$F = N_A e$	$9.648\,46 \times 10^{4}$ C mol^{-1}
Boltzmann's constant	k	$1.380\,66 \times 10^{-23}$ J K^{-1}
Gas constant	$R = N_A k$	8.314 41 J mol^{-1} K^{-1}
Atomic mass unit	a.m.u. $= \frac{1}{12} M_{^{12}\mathrm{C}}$	$1.660\,57 \times 10^{-27}$ kg
Electron mass	m or m_e	$9.109\,53 \times 10^{-31}$ kg $= 5.485\,80 \times 10^{-4}$ a.m.u.
Proton mass	M_p	$1.672\,65 \times 10^{-27}$ kg $= 1.007\,276$ a.m.u.
Neutron mass	M_n	$1.674\,92 \times 10^{-27}$ kg $= 1.008\,665$ a.m.u.
Ratio of proton to electron mass	M_p/m_e	1836.15
Electron charge to mass ratio	$\lvert e \rvert / m_e$	$1.758\,80 \times 10^{11}$ C kg^{-1}
Classical radius of electron	$r_0 = \frac{e^2}{4\pi\varepsilon_0 m c^2}$	$2.817\,84 \times 10^{-15}$ m
Bohr radius for atomic hydrogen (with infinite nuclear mass)	$a_0 = \frac{4\pi\varepsilon_0 \hbar^2}{m e^2}$	5.29177×10^{-11} m
Rydberg's constant for infinite nuclear mass	$R_\infty = \frac{m e^2}{8\varepsilon_0^2 h^3 c} = \frac{\alpha}{4\pi a_0}$	$1.097\,37 \times 10^{7}$ m^{-1}
Rydberg's constant for atomic hydrogen	R_H	$1.096\,78 \times 10^{7}$ m^{-1}
Bohr magneton	$\mu_B = \frac{e\hbar}{2m}$	9.27408×10^{-24} J T^{-1}
Nuclear magneton	$\mu_N = \frac{e\hbar}{2M_p}$	5.05082×10^{-27} J T^{-1}

NOTES

NOTES

NOTES

UTTAR PRADESH
RAJARSHI TANDON OPEN UNIVERSITY

UGPHS-08

Modern Physics

Block

3

APPLICATION OF QUANTUM MECHANICS TO SOME SYSTEMS

UNIT 8

Some Simple Systems 5

UNIT 9

Spherically Symmetric Systems: Hydrogen Atom 28

UNIT 10

Atomic Spectra 45

UNIT 11

X-Ray Spectra 61

APPLICATION OF QUANTUM MECHANICS TO SOME SYSTEMS

In the two previous blocks of this course we have introduced you to certain ideas and concepts (related to the special theory of relativity and quantum mechanics) which have shaken many of our age-old beliefs about the world around us. You must have noticed that most of these concepts and theories were developed in the first quarter of the twentieth century. So we can well say that this period was a period of revolution in the foundations of physics. Speaking of revolutions, the development of quantum mechanics led to the greatest conceptual revolution of our century and probably to the greatest mankind has ever experienced. It most likely exceeded the great revolutions in our thinking brought about by the Copernican revolution, Darwinian revolution (recall Units 9 and 13 of FST), and the special as well as general theory of relativity.

Quantum mechanics has forced us to reconsider our deepest convictions about the reality of nature. For example, the concept of all entities in nature being **wave-particles** (Unit 4) does away with the distinction between matter and radiation (which is at the root of classical physics). We have successfully resolved the classical physicist's dilemma of "how can an electron (or any other particle) be both a particle and a wave?" We now say, "The electron is neither; it is neither a classical particle nor a classical wave. It is a quantum wave-particle which possesses both momentum and wavelength." It reveals its particle properties in certain experiments and wave properties in others (**complementarity**).

Then you have seen how classical determinism and causality are called into question by quantum mechanics through the **uncertainty principle** (Unit 5) and the **probabilistic description**. Uncertainty is not a matter of our inability to do better. Quantum mechanics claims it to be intrinsic to the nature of the quantum world. That a particle may not have a well defined position and momentum at the same time is a fundamentally new notion which many physicists including Einstein found hard to accept. The uncertainty principle maintains that this is an *intrinsic limitation*. It is indeed in the nature of things — either we define the particle's position and lose all information about its momentum, or we pinpoint its momentum and do not know where in space it is. This is surely a far cry from the classical world.

These fundamental concepts together with the equation of time evolution of matter waves (Unit 6) lead us to the probabilistic description of the quantum world. Quantum mechanics tells us that *the probabilistic description is also the fundamental description; there is no deeper level.* It gives us only the probability amplitude from which a probability can be computed. For example, we can compute the *probability* that an electron would have a certain velocity. But quantum mechanics does not give the velocity of a specific particle. Instead it claims that no such detailed information even *exists*. The classical question itself, 'what is the velocity of a specific particle in an ensemble of particles?' is in most cases considered meaningless.

And did you realise while studying Block 2 that this conceptual foundation of quantum mechanics was laid down by a very small group of people? Of the five most prominent ones, three were about forty years of age and the other two, in their early twenties. Two of the older ones, the Danish physicist Neils Bohr and the German Max Born, contributed more wisdom and interpretation than formalism or mathematical structure. That was done by the two younger ones, the Englishman Paul Adrien Maurice Dirac and the German Werner Heisenberg. The fifth one, the Austrian Erwin Schrödinger, played a curious role in that he contributed greatly and fundamentally to the theory but turned his back on the interpretation which others gave to it.

Schrödinger developed his 'wave mechanics', and simultaneously and independently Heisenberg constructed 'matrix mechanics'. These were apparently different theories even in the kind of mathematics they used: wave mechanics used calculus and matrix mechanics used algebra. However, Dirac alongwith Schrödinger and the mathematician John von Neumann established that these two theories are completely equivalent. From that time on, the general term quantum mechanics for both of them has become standard usage in scientific literature.

Having provided you once again with a bird's eye view of the fundamental concepts of quantum mechanics discussed in Block 2, we draw your attention to another feather in its cap — its ability to predict. You will discover its power in this block when you study its applications to various systems, such as the **potential well** and **potential barrier**, and the **harmonic oscillator** (Unit 8). The phenomenon of barrier penetration (also known as **tunneling**) by quantum particles, finds wide use today in tunnel diodes and scanning tunneling electron microscopes which enable us to 'see' the quantum world with much more clarity than hitherto possible (of course, without violating the uncertainty principle). When the formalism of quantum mechanics was developed, it was applied extensively to study the structure of **hydrogen atom** — its predicted energy levels were in beautiful agreement with the experimental results. And the optical spectrum resulting from the transition of an electron from a higher energy state to a lower one was explained very well. Therefore, in Unit 9 we discuss the application of quantum mechanics to spherically symmetric systems, and especially to hydrogen atom. The evolution of quantum mechanics has been guided from the beginning by discoveries in atomic spectroscopy. Explaining the internal structure of atoms and the resulting **spectra** (in the **optical** as well as **X-ray** regions) has been a major success of quantum mechanics and we take these up in Units 10 and 11, respectively.

As far as studying the material is concerned, follow the advice given in Block 2 — learn to think and calculate quantum mechanically! You may find the going tough initially as there is a lot of mathematical abstraction in this material. Our advice to you is to solve each and every step on your own. Do not try to rush through the material if you want to really understand it. In our estimate, you would need to spend 8 h on Unit 8, 7 h on Unit 9, 5 h on Unit 10 and just 3 h on Unit 11, for studying the material and assimilating the ideas. And at the end of it all, we hope that you experience a genuine sense of achievement and exhilaration at having understood one of the greatest and most beautiful intellectual creations of twentieth century science. Our best wishes are with you!

Acknowledgement

We are thankful to Shri Gopal Krishan Arora for his invaluable contribution in word-processing of the entire course, and for secretarial assistance.

UNIT 8 SOME SIMPLE SYSTEMS

Structure

8.1 Introduction
Objectives
8.2 A Free Particle
8.3 Particle in a Box
8.4 One-dimensional Rectangular Potential Barrier
8.5 One-dimensional Potential Well
8.6 One-dimensional Harmonic Oscillator
8.7 Summary
8.8 Terminal Questions
8.9 Solutions and Answers

8.1 INTRODUCTION

In Unit 6 of Block 2 you have studied the time independent Schrödinger equation. In this unit we shall apply it to some simple one-dimensional systems so that you get some experience of solving the time-independent Schrödinger equation. You will learn how to solve it to obtain its eigenfunctions satisfying the conditions given in Sec. 6.4 of Unit 6 and determine the corresponding eigenvalues for these systems. Although the real world is three-dimensional, these one-dimensional systems are of great interest. This is not only because several physical situations are effectively one-dimensional but also because we can use them effectively to model the real world. A number of more complicated physical problems can be reduced to the solution of equations similar to the one-dimensional Schrödinger equation.

So in this unit we shall first study a **free particle** and then confine it to move within a limited space like a box. You will notice that the eigenvalues and eigenfunctions of the two systems are different. Then we shall solve the Schrödinger equation for a particle moving in a one-dimensional **rectangular potential barrier** and a **one-dimensional potential well**. These examples of simple potentials model natural processes like alpha decay of radioactive nuclei, and the energy spectrum of atoms, molecules and nuclei. Thus our models are expected to give us some insight into the physics of these systems. They will also bring out the difference between classical and quantum description of motion of objects. We shall end this unit by studying the quantum mechanical behaviour of a **simple harmonic oscillator**. As its application, we shall consider the vibrations of a diatomic molecule in terms of the oscillations of a simple harmonic oscillator. In the next unit, you will learn how to solve the Schrödinger equation for the hydrogen atom.

Objectives

After studying this unit you should be able to

- solve the time independent one-dimensional Schrödinger equation and obtain the eigenfunctions and eigenvalues for simple one-dimensional systems like

  free particle and particle in a box,

  particle in a one-dimensional potential barrier and one-dimensional potential well,

  simple harmonic oscillator.

- apply the one-dimensional potential models to simple applications in quantum physics.

8.2 A FREE PARTICLE

Let us consider the simplest case of a potential which is constant:

$$V(x) = V_0$$

The force acting on the particle, $F(x) = -\frac{dV}{dx}$, then vanishes, so that the particle is free. Without loss of generality, we can take the constant V_0 to be zero. So now we have a particle of mass m moving freely in space. Since the particle is not subjected to any force, its total energy E, which is equal to its kinetic energy, and also its linear momentum $\mathbf{p}$ do not change with time. E and p are related by the equation

$$\frac{p^2}{2m} = E \tag{8.1}$$

Recall Eq. (6.28) from Unit 6. Putting $V(x) = 0$ in it, we can write the time independent Schrödinger equation for a free particle as follows:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} = E\psi(x) \tag{8.2}$$

where the direction of the vector $\mathbf{p}$ is taken to be along x axis. Since $\mathbf{p} = \hbar\mathbf{k}$, where k is the wave number, we can combine Eqs. (8.1) and (8.2) to obtain

$$\frac{d^2\psi(x)}{dx^2} = -k^2\psi(x) \tag{8.3a}$$

where, $$E = \frac{\hbar^2 k^2}{2m} \tag{8.3b}$$

The two linearly independent solutions of Eq. (8.3a) are given by

$$\psi_{\pm k}(x) = e^{\pm ikx} \tag{8.4a}$$

Thus for one value of E we have two eigenfunctions e^{+ikx} and e^{-ikx}. We denote these eigenfunctions by $\psi_k(x)$ and $\psi_{-k}(x)$. Recall that such **eigenfunctions which have the same eigenvalue for a given eigenvalue-eigenfunction equation are called degenerate**. (Otherwise they are **non-degenerate**). Thus $\psi_k(x)$ and $\psi_{-k}(x)$ are degenerate eigenfunctions. The general solution of Eq. (8.3a) is the linear combination

$$\psi(x) = Ae^{ikx} + Be^{-ikx} \tag{8.4b}$$

where A and B are arbitrary constants. It is clear that for the solution $\psi(x)$ to be physically acceptable, k cannot have an imaginary part. If it did, then $\psi(x)$ would increase exponentially at one of the limits $x = -\infty$ or $x = +\infty$, or possibly at both limits. You may know that the wavefronts of the waves represented by Eqs. (8.4) are planes, perpendicular to the direction of the propagation of the wave. Hence we call $e^{\pm ikx}$ as **plane waves**. They do not go to zero as $x \to \pm\infty$. Hence they are unnormalizable wave functions (recall Sec. 6.3.2 of Unit 6).

In order to interpret Eq. (8.4b) physically, let us consider some special cases. If we set $B = 0$, the resulting wavefunction is the plane wave

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

where we have included the time dependence (recall Eq. 6.27). This is associated with a free particle of mass m moving along the positive x-axis with a definite momentum of magnitude k and an energy $\hbar^2 k^2/2m$. The probability density corresponding to it is $\psi^*\psi = |A|^2$. It is independent of time as well as position. Thus, the position of the particle is completely unknown. This is in accordance with the uncertainty principle. You can obtain a similar analysis by setting $A = 0$ in Eq. (8.4b). In this case the plane wave will be travelling in the negative x-direction.

You can also verify easily that $\psi_{\pm k}(x)$ are also eigenfunctions of the operator p with the eigenvalues $\pm\hbar k$. Thus as far as the linear momentum operator p is concerned, the

two eigenfunctions $\psi_{\pm k}(x)$ are non-degenerate, i.e., they have different eigenvalues. Why don't you do it as an exercise?

SAQ 1 *Spend 2 min*

Show that $p_{op}\, e^{\pm ikx} = \pm \hbar k\, e^{\pm ikx}$.

For the above system both the constants of motion E and p are given in terms of k. Thus k characterises the eigenfunction $\psi_k(x)$ and hence we call k a **quantum number**. Since $E = \hbar^2k^2/2m$, we must have $E > 0$, i.e., the energy cannot remain lower than the potential (here $V = 0$) over the entire interval $(-\infty, +\infty)$. Since any non-negative value of E is allowed, the energy spectrum of a free particle is continuous, extending from $E = 0$ to $+\infty$. This, of course, is not surprising since E is the kinetic energy of a free particle. It also corresponds to the classical result.

Let us now see what happens when we confine the free particle to a box.

8.3 PARTICLE IN A BOX

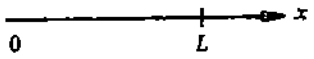


Fig. 8.1 : Particle confined in a line segment.

Let us consider a one-dimensional system and confine a free particle in a length segment lying between $x = 0$ and $x = L$ (Fig. 8.1). Then the probability of finding the particle at $x = L - \varepsilon$ and $x = L + \varepsilon$, where ε is an infinitesimal quantity, should be zero. Further, since the wave functions are single-valued and continuous, they should also satisfy the condition that they are zero at the boundaries $x = 0$ and $x = L$, i.e.,

$$\lim_{\varepsilon \to 0} \psi_k(-\varepsilon) = \lim_{\varepsilon \to 0} \psi_k(L + \varepsilon) = 0 \qquad (8.5a)$$

It follows from Eq. (8.5a) that

$$\psi_k(0) = \psi_k(L) = 0 \qquad (8.5b)$$

Now we may also write the wave function as

$$\psi_k(x) = A \cos kx + B \sin kx$$

where $k = (2mE/\hbar^2)^{1/2}$. Since $\psi_k(0) = 0$, we have

$$A = 0$$

and since $\psi_k(L) = 0$, we get

$$B \sin kL = 0 \qquad (\because B \neq 0)$$

whence $kL = n\pi, \qquad n = 1, 2, 3, \ldots$

Hence k is quantized; let us call it k_n:

$$k_n = \frac{n\pi}{L}, \qquad n = 1, 2, 3, \ldots \qquad (8.6a)$$

Hence, energy E is also quantized; let us call it E_n:

$$E_n = \frac{\hbar^2 k_n^2}{2m} = \frac{n^2\pi^2\hbar^2}{2mL^2}, \qquad n = 1, 2, 3, \ldots \qquad (8.6b)$$

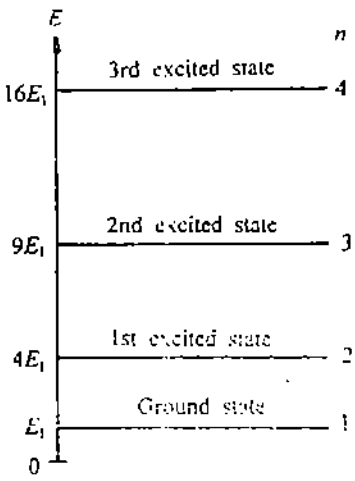


Fig. 8.2 : The four lowest allowed energy levels for a particle in a one-dimensional box, where $E_1 = \hbar^2/8mL^2$.

Notice that the quantization of energy follows from the condition of confinement or the boundary conditions imposed on the wave function. Thus, the moment we confine a free particle in a limited space, its energy can take only discrete values given by Eq. (8.6b); it is no longer continuous, it is quantized (Fig. 8.2). The eigenfunction of a state of quantized energy E_n of a particle in a box is, therefore, given by

$$\psi_k(x) = \psi_n(x) = N \sin\left(\frac{n\pi x}{L}\right) \qquad (8.6c)$$

Notice that we still have to compute the normalisation constant N. You may like to do this exercise yourself. Here's an SAQ for you.

SAQ 2

Spend 5 min

(a) Show that $N = \left(\frac{2}{L}\right)^{1/2}$

(b) Show that $\psi_n(x)$ and $\psi_m(x)$ are orthogonal when $m \neq n$.

Hence the eigenfunction for a free particle confined in a segment of length L is

$$\psi_n(x) = \left(\frac{2}{L}\right)^{1/2} \sin\left(\frac{n\pi x}{L}\right) \tag{8.7a}$$

The complete solution for a stationary state of a particle in a box is then given as

$$\begin{aligned}\psi_n(x, t) &= \psi_n(x) e^{-iE_n t/\hbar} \\ &= \left(\frac{2}{L}\right)^{1/2} \sin\left(\frac{n\pi x}{L}\right) e^{-iE_n t/\hbar}\end{aligned} \tag{8.7b}$$

and the probability for finding the particle at position x is independent of time:

$$P_n(x) = |\psi_n(x, t)^2| = \left(\frac{2}{L}\right) \sin^2\left(\frac{n\pi x}{L}\right) \tag{8.7c}$$

However, it now depends on the position of the particle. Fig. 8.3 shows the probabilities per unit distance for a particle in a box for the three lowest energy states. The classical result is also shown for comparison.

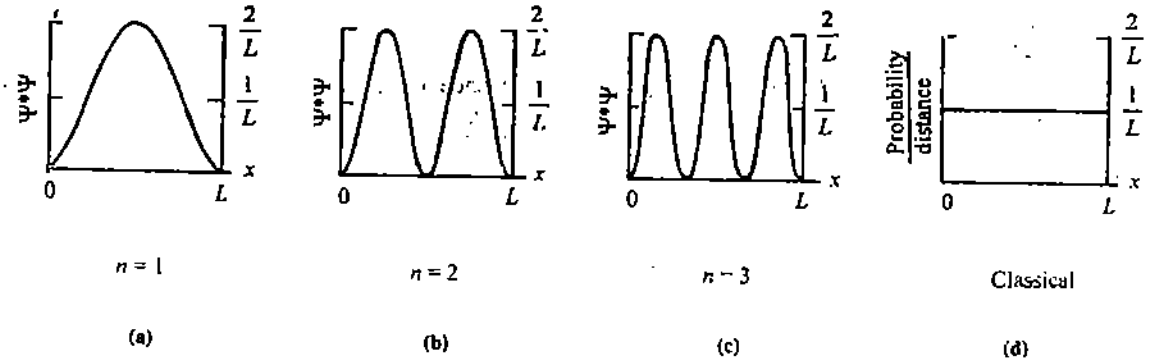


Fig. 8.3: The probability per unit distance for a particle in a box for the three lowest values of n, compared to the classical result.

The smallest value of n is $n = 0$, but $n = 0$ gives $\psi = 0$. This result means that the probability of an $n = 0$ state is zero. Now refer to Fig. 8.3. For $n = 1$, $\psi^*\psi = (2/L) \sin^2 (\pi/L)x$; for $n = 2$, $\psi^*\psi = (2/L) \sin^2 (2\pi/L)x$; and for $n = 3$, $\psi^*\psi = (2/L) \sin^2 (3\pi/L)x$. Thus $\psi^*\psi$ oscillates between 0 and $2/L$ with an average value in the box of $1/L$.

Classically, the particle bounces back and forth between the walls. Since its kinetic energy is constant, it moves at a constant velocity between collisions with the walls. Therefore, it spends the same amount of time in all equal-distance intervals. This gives it the constant probability per unit distance of $1/L$ of being found anywhere in the box (Fig. 8.3d).

As you can see from Fig. 8.3a to c, $\psi^*\psi$ has n humps. This behaviour is easy to understand if you remember that k_n, the wave number for state n, equals $2\pi/\lambda_n$. Equating this wave number to $n\pi/L$ gives $L = n\lambda_n/2$. That is, n is the number of the particle's half wavelengths being fitted between the walls. In the limit of large quantum numbers, the number of humps with any finite Δx becomes so large that $\psi^*\psi\ \Delta x$ approaches the classical value, $(1/L)\Delta x$.

The linear momentum of the particle is given by

$$p_n = \pm k_n \hbar = \pm \frac{\hbar n \pi}{L}, \qquad n = 1, 2, 3, \ldots \tag{8.8}$$

You should note that the integer n in Eqs. (8.6a, b) and (8.8) now plays the same role as was played by k for a free particle. Hence n is called a **quantum number**. Since n takes only integer values, the momentum and the energy of the particle no longer vary in a continuous manner but take only discrete values.

Let us see what are the other implications of these results about a particle in a one-dimensional box. According to quantum mechanics, there is a minimum energy, the ground-state energy $E_1 = \pi^2\hbar^2/2mL^2$ that the particle in a box must have. This result is in contrast to the classical result where all energy values including $E = 0$ are permitted. You can easily see that this is a consequence of the uncertainty principle. We are confining the particle within the length L. Thus there is an uncertainty in its position given by L. According to the uncertainty relation, Eq. (5.6a), there has to be a corresponding momentum uncertainty of $\hbar/L$. This means that the particle has to have a minimum energy; its energy can never be zero because that would contradict the uncertainty relation. This minimum energy is called the **zero point energy.**

Notice from Eq. (8.6b) that the energy separation between successive quantized levels increases with the decrease of L, the confining length. Conversely, as L increases, the energy separations decrease. When L is much larger than atomic dimensions, the energy separation is so small that we approach the classical correspondence limit. Note that for large L, the zero point energy also tends towards zero.

Thus far we have considered the qualitative situation. But what about the quantitative situation? Let us calculate the energy levels of an electron ($m = 9.1 \times 10^{-31}$ kg) confined to a box of atomic size ($L \approx 10^{-10}$ m). Substituting in the energy level formula, Eq. (8.6b), we get $E_n = 40n^2$ eV. The energy difference between the ground and the first excited state is then 120 eV. And a photon emitted from transition between these energy levels would have a frequency 3×10^{16} Hz, which is of the same order as that in many atomic transitions.

Now from Eq. (8.8) we have

$$p_{n+1} - p_n = \hbar\pi/L \tag{8.9}$$

Thus, the difference between p_{n+1} and p_n decreases with the increase of L. Now if we take a very large value of L, $(p_{n+1} - p_n)$ will approach zero, i.e., p will become continuous. Hence we may represent a free particle by a normalised wave function with a very large value of L. Such a procedure for the normalisation of a free particle is known as **box normalisation** and you can easily see that a one-dimensional box-normalised wave function is given by

$$\psi_k(x) = \left(\frac{1}{L}\right)^{1/2} e^{ikx} \tag{8.10}$$

We can generalise this analysis to a particle in a three-dimensional box. In fact, you should work this out yourself.

SAQ 3

Spend 10 min

Obtain eigenfunctions and eigenvalues of a particle confined in a three dimensional box having lengths L_x, L_y and L_z.

Let us now consider the problem of a particle in a one-dimensional rectangular potential barrier.

8.4 ONE-DIMENSIONAL RECTANGULAR POTENTIAL BARRIER

In Sec. 8.2, we considered the motion of a free particle, i.e., the potential in the entire

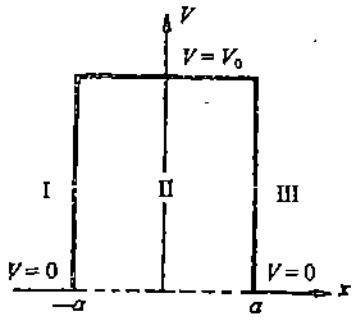


Fig. 8.4 : One-dimensional potential barrier.

space was zero. Let us now modify that situation and consider a one-dimensional space in which the potential energy of a particle is V_0 between $x = -a$ and $x = a$ but zero otherwise:

$$\begin{aligned} V(x) &= 0 \qquad x < -a \\ &= V_0 \qquad -a < x < a \\ V(x) &= 0 \qquad x > a \end{aligned} \tag{8.11}$$

Thus we divide the whole one-dimensional space in three regions. Region I extends from $-\infty$ to $-a$, region II from $-a$ to $+a$ and region III from $+a$ to $+\infty$ (see Fig. (8.4)). The central region is known as the **potential barrier**. Notice that as defined by Eq. (8.11), V is rectangular.

Let us consider the motion of a particle of mass m and total constant energy E in the above mentioned one-dimensional space. From purely physical considerations you can see that the particle is free in region I and there will be an incident plane wave as well as a plane wave reflected by the potential barrier. Hence, we may write the wave function for region I as

$$\psi_I(x) = A\, e^{ikx} + B\, e^{-ikx} \quad \text{for } x < -a \tag{8.12}$$

where

$$k = \sqrt{2mE}/\hbar \tag{8.13}$$

In region III the particle is again free ($V = 0$) but there we have only a transmitted wave. Hence

$$\psi_{III}(x) = F\, e^{ikx} \qquad \text{for } x > a \tag{8.14}$$

However, in region II the Schrödinger equation is

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_{II}(x)}{dx^2} + V_0\,\psi_{II}(x) = E\psi_{II}(x) \tag{8.15}$$

Let us rewrite this equation in the form

$$\frac{d^2\psi_{II}(x)}{dx^2} + \frac{2m}{\hbar^2}(E - V_0)\,\psi_{II}(x) = 0 \tag{8.16}$$

Introducing a real parameter γ, we may write Eq. (8.16) as

$$\frac{d^2\psi_{II}(x)}{dx^2} - \gamma^2\psi_{II}(x) = 0 \tag{8.17a}$$

where

$$\gamma^2 = \frac{2m}{\hbar^2}(V_0 - E) \tag{8.17b}$$

The solution of Eq. (8.17a) is given as

$$\psi_{II}(x) = Ce^{-\gamma x} + De^{\gamma x}, \qquad -a < x < a \tag{8.18}$$

Now, there are two possibilities: (i) E is less than V_0 and (ii) E is greater than V_0. Let us consider the case when $E < V_0$. Then γ is real and the solution of Eq. (8.17a) becomes

$$\psi_{II}(x) = Ce^{-\gamma x} + De^{\gamma x} \qquad \text{for } -a < x < a \tag{8.19}$$

With the help of Eqs. (8.19) and incident part (Ae^{ikx}) of Eq. (8.12) we find that the incident flux from the left of the potential barrier is

$$S_i = \frac{\hbar k}{m}\,|A|^2 \tag{8.20}$$

Similarly, using the reflected part of the wave function ψ_I and the transmitted wave function ψ_{III} we find that the reflected flux S_r and the transmitted flux S_t are given by

$$S_r = S_i\,|B/A|^2 \tag{8.21}$$

$$S_t = S_i\,|F/A|^2 \tag{8.22}$$

You should quickly verify Eqs. (8.20) to (8.22) before studying further. Thus we obtain

$$\text{Probability of reflection } P_r = S_r/S_i = \left|B/A\right|^2 \tag{8.23}$$

and

$$\text{Probability of transmission } P_t = S_t/S_i = \left|F/A\right|^2 \tag{8.24}$$

Now to determine P_r and P_t (i.e., the coefficients A, B and F) we match the appropriate wave functions ψ and their derivatives ψ' at the boundaries $x = -a$ and $x = a$. The boundary conditions are:

$$\psi_I(x = -a) = \psi_{II}(x = -a); \quad \psi'_I(x = -a) = \psi'_{II}(x = -a) \tag{8.25a}$$

and

$$\psi_{II}(x = a) = \psi_{III}(x = a); \quad \psi'_{II}(x = a) = \psi'_{III}(x = a) \tag{8.25b}$$

Substituting ψ_I, ψ_{II} and ψ_{III} from Eqs. (8.12), (8.19) and (8.14) and carrying out the necessary algebra we can obtain P_r and P_t as given below:

$$P_t = \left[1 + \frac{V_0 \sinh^2 (2\gamma a)}{4E\,(V_0 - E)}\right]^{-1} \tag{8.26}$$

$$P_r = \left[1 + \frac{4E\,(V_0 - E)}{V_0 \sinh^2 (2\gamma a)}\right]^{-1} \tag{8.27}$$

You may like to prove Eqs. (8.26) and (8.27) before studying further. So here is an exercise for you.

SAQ 4

Spend 20 min

Prove Eqs. (8.26) and (8.27).

You can easily verify that $P_t + P_r = 1$. One of the remarkable features of this analysis is that there is a non-zero probability of transmission even though $E < V_0$, i.e., the energy of the object is less than the height of potential barrier. This is in contrast to the classical situation. It is a purely quantum effect signifying the wave property of quantum objects. Hence, quantum objects are said to be able to **tunnel** through a classically impenetrable barrier. Thus, we conclude that the penetration through a potential barrier, usually referred to as **tunneling**, is a quantum mechanical phenomenon and should be taken into consideration for those objects for which wave-particle duality is significant. You may like to know that the phenomenon of tunneling has many applications. It explains the emission of α-particles in radioactive decay, and the passage of electrons through a forbidden region as in field emission. In electronics, it has been utilised in the construction of tunnel diodes. Let us take one of these applications, namely, alpha decay.

You should not take the word 'tunneling' literally. There is, of course, a finite probability for the particle to be inside the classically forbidden barrier region where its kinetic energy is negative. But the point is that nobody can "see" a particle actually go through a classically forbidden region. Particle detectors can detect only objects of kinetic energy greater than zero. If you insert a detector inside the barrier to 'see' the particle, you are not only making a hole in the potential but also in your objective, because the object will no longer belong to a classically forbidden region where you wanted to find it! Another way to say this is that our effort to observe the object with any measuring instrument will impart to it an uncontrollable amount of energy. This is how the uncertainty principle works in such measurement situations!

Similarly, we cannot say what kind of time the particle takes to tunnel through a barrier. It may do it now or later; quantum mechanics gives us only average behaviour — the average time it takes to travel to the other side. It seems best to think of quantum tunneling as "as if" tunneling — tunneling in potentia. Alternatively, we can say that quantum objects go across a barrier for which they do not have enough energy by doing a quantum "jump"— Bohr style; they never go through the intervening space, and they do not do it causally taking a finite amount of time. Now they are here, on this side of the barrier; and then they are there, on the other side.

Example 1: Application to Alpha Decay

We are giving this application for the sake of interest only. You will not be tested on the mathematics given here.

Let us consider the case of alpha radioactivity of nuclei. This was the original application of the barrier transmission phenomenon worked out by Gamow, Gurney and Condon.

In this case of a physical barrier, the shape is not so square (Fig. 8.5). Now when $\gamma a \gg 0$, the probability of transmission becomes

$$P_t = \left|T\right|^2 = \frac{16\,k^2\,\gamma^2}{(k^2 + \gamma^2)^2} \exp(-4\gamma a) \simeq \exp(-4\gamma a)$$

We can write this as

$$P_t = e^{-2\,[2m\,(V_0 - E)/\hbar^2]^{1/2}\,.\,2a}$$

For a broad range of barriers, we can now make a generalised estimate of the order of magnitude. Generalising the expression of P_t for $V(x)$, we can write

$$P_t = \exp\left[-2\int_a^b dx\left(\frac{2m}{\hbar^2}(V(x)-E)\right)^{1/2}\right]$$

Now we can approximate the situation for alpha emission from a nucleus by the potential barrier shown in Fig. 8.5b. Inside the nucleus, the alpha particle is a free particle, $E > 0$. (For, if the alpha particle were bound, how would the nucleus decay?). It has to tunnel through the coulomb barrier

$$V(r) = \frac{ZZ'e^2}{r}$$

where Z and Z' are the atomic numbers of the daughter nucleus and the alpha particle, respectively, into which the parent nucleus is splitting. Here r is the radial distance. Now, in the expression for P_t we can change the variable x to r, since x is just a dummy variable. Putting P_t equal to exp $(-G)$, we obtain

$$G = 2\,(2m/\hbar^2)^{1/2}\int_R^b\left[\frac{ZZ'e^2}{r} - E\right]^{1/2} dr$$

where R is the radius of the daughter nucleus (Fig. 8.5b) and the upper limit of the integral, b, is taken to be the classical "turning point" where the integrand vanishes, since

$$E = \frac{ZZ'e^2}{b}$$

Thus

$$G = 2\,(2m/\hbar^2)^{1/2}\,(ZZ'e^2)^{1/2}\int_R^b\left(\frac{1}{r} - \frac{1}{b}\right)^{1/2} dr$$

The value of the integral is

$$\int_R^b dr\left(\frac{1}{r} - \frac{1}{b}\right)^{1/2} = \sqrt{b}\left[\cos^{-1}\left(\frac{R}{b}\right)^{1/2} - \left(\frac{R}{b} - \frac{R^2}{b^2}\right)^{1/2}\right]$$

For low energies and high barriers, $b >> R$, and we get

$$G \approx 2\left(\frac{2m\,ZZ'e^2b}{\hbar^2}\right)^{1/2}\frac{\pi}{2}$$

But $b = ZZ'e^2/E = 2ZZ'e^2/mv^2$, where v is the velocity of the alpha particle inside the nucleus. Hence

$$G = \frac{2\pi\,ZZ'e^2}{\hbar v}$$

To calculate the alpha particle escape probability per second, we have to multiply the transmission coefficient exp $(-G)$ by the rate of the alpha particles hitting the barrier, which is $\sim v/R$. For a 1 MeV alpha particle, using $R = 1.2 \times 10^{-13}\,A^{1/3}$ cm with $A = 216$, we estimate $v/R = 10^{21}\,s^{-1}$. Consequently, escape probability per second $= \tau^{-1} = 10^{21}e^{-G}$ where τ denotes the decay time. Noting that $Z' = 2$ and that the mass of alpha particle in energy units is $m \approx 4 \times 10^3$ MeV, we get

$$G = \frac{2\pi\,ZZ'e^2}{\hbar\,(2E/m)^{1/2}} \approx \frac{4Z}{(E(\text{in MeV}))^{1/2}}$$

Therefore, we obtain

$$\log_{10}\frac{1}{\tau} = C_1 - C_2\,\frac{Z}{(E\ \text{MeV})^{1/2}}$$

where C_1 and C_2 are two constants, never mind our estimates for them. This formula, first derived by Gamow, Gurney and Condon, fits the data on alpha decay quite

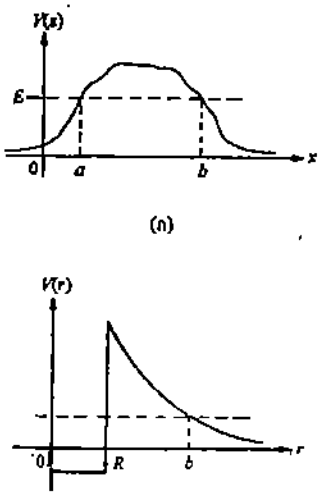


Fig. 8.5: (a) A realistic potential barrier has a more jagged look than a square barrier; (b) model potential barrier for alpha decay of nuclei.

remarkably. It is also remarkable that we can derive the formula from a one-dimensional calculation.

Let us now consider the situation for $E > V_0$. Classically the particle with $E > V_0$ will be transmitted with cent per cent probability. The quantum mechanical prediction may be easily obtained by appropriately modifying the above formulae for this case. The only modification is in γ which is now imaginary. Hence taking $\gamma = iq$ with $q^2 = (2m/\hbar^2)(E-V_0)$, we get

$$P_t = |T|^2 = \left[1 + \frac{V_0^2 \sin^2(2qa)}{4E(-V_0+E)}\right]^{-1} \tag{8.28}$$

and

$$P_r = |R|^2 = \left[1 + \frac{4E(-V_0+E)}{V_0^2 \sin^2(2qa)}\right]^{-1} \tag{8.29}$$

Thus, we see that quantum mechanically $P_r \neq 0$ and hence, even if $E > V_0$, a part of the incident flux is reflected. In Fig. 8.6 we have shown the variation of the transmission probability with the barrier height. We see in the figure that P_t is low for low E and for high E it approaches unity.

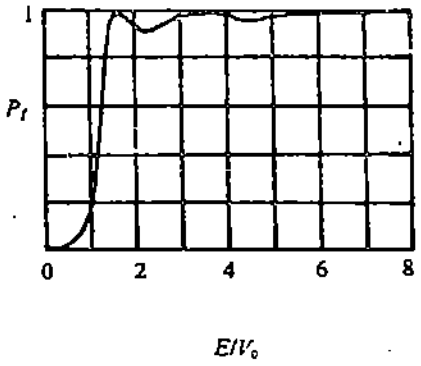


Fig. 8.6 : Transmission probability of a particle of energy E through a rectangular potential barrier of height V_0.

An interesting feature of the above curve is that for a certain value of (E/V_0), the transmission probability P_t is unity and hence $P_r = 0$ (see Fig. 8.6). This happens whenever $2qa = n\pi$ with $n = 0, 1, 2, \ldots$. The two boundaries at $x = \pm a$ where the reflection is taking place are producing reflected waves of equal amplitudes and opposite phases. Hence, there is no reflection at all.

Let us now study the case of a particle in a one-dimensional potential well.

8.5 ONE-DIMENSIONAL POTENTIAL WELL

Let us obtain the eigenvalues and eigenfunctions of a potential well, shown in Fig. 8.7. The well is such that a particle has V_0 potential energy for $x < -a$ as well as for $x > a$ but for $-a < x < a$ the potential energy is zero:

$$V = V_0 \qquad x < -a$$

$$V = 0 \qquad -a < x < a$$

$$V = V_0 \qquad x > a$$

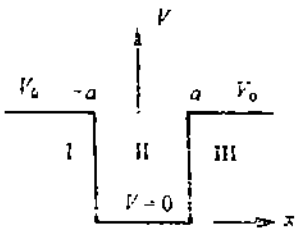


Fig. 8.7: Rectangular potential well.

The total energy E of a particle of mass m is again a constant of motion. Hence the Schrödinger equations of the particle in regions I and III are given by

$$H\psi = E\psi \tag{8.30a}$$

where

$$H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x) \tag{8.30b}$$

Substituting for $V(x)$, we obtain for region I

$$\psi''(x) + \frac{2m}{\hbar^2}(E - V_0)\,\psi(x) = 0 \qquad \text{for } |x| > a \tag{8.31a}$$

and in region II it is

$$\psi''(x) + \frac{2mE}{\hbar^2}\,\psi(x) = 0 \qquad \text{for } |x| < a \tag{8.31b}$$

We define

$$\gamma^2 = \frac{2m}{\hbar^2}(V_0 - E) \quad \text{and} \quad q^2 = \frac{2mE}{\hbar^2} \tag{8.32}$$

and consider the case when $V_0 > E$, i.e., γ is real. In region II the wave function is given by

$$\psi_{II}(x) = B\cos(qx) + C\sin(qx) \tag{8.33}$$

Similarly, general solutions in regions I and III are given by

$$\psi_I(x) = A\exp(\gamma x) \tag{8.34a}$$

and $$\psi_{III}(x) = D\exp(-\gamma x) \tag{8.34b}$$

Let us further examine the wave function $\psi_{II}(x)$. Notice that the Hamiltonian (Eq. 8.30b) of this system does not change when x is replaced by $-x$. This leads to an interesting result that H commutes with the parity operator P. We can prove this as follows:

$$P[H\psi(x)] = H(-x)\,\psi(-x) = H\psi(-x) = H(x)\,[P\psi(x)]$$

Hence

$$PH - HP = 0, \text{ since } \psi(x) \neq 0$$

or

$$[P, H] = 0$$

Now recall this result from Unit 7, Sec. 7.3: If an operator A commutes with the parity operator, then the non-degenerate eigenfunctions of A have definite parity, i.e., they are either of odd parity or of even parity. Let us again demonstrate this for H. Let

$$H(x, p_x)\,\psi(x) = \lambda\psi(x)$$

Then operating with P from the left we get

$$PH\,(x, p_x)\,\psi(x) = \lambda P\,\psi(x)$$

or

$$H(P\psi) = \lambda(P\psi) \qquad (\because PH = HP)$$

Thus both ψ and $P\psi$ are eigenfunctions of H for the same eigenvalue λ. Since $\psi(x)$ is a non-degenerate wave function, the two functions can differ at the most by a constant. Hence

$$P\psi(x) = p\psi(x)$$

where p is a constant.

Thus, $\psi(x)$ is an eigenfunction of P with p as the eigenvalue. Operating P once again on the left we get

$$P^2\psi(x) = P\,p\psi(x) = p^2\psi(x)$$

and

$$P^2\psi(x) = P\psi(-x) = \psi(x)$$

implying $p^2 = 1$ and $p = \pm 1$. The wave functions for which $p = +1$ are called **wave functions of even parity** and those for which $p = -1$ are called **wave functions of odd parity**. Since p can take only one of these values at one time, $\psi(x)$ has to be of either even or odd parity.

$\psi_{II}(x)$, as given by Eq. (8.33) is of mixed parity. Hence the above discussion shows that it is unacceptable. Therefore, either C or B should be zero. If we take $C = 0$ in Eq. (8.33) we get even parity solutions while $B = 0$ yields odd parity solutions. You can verify that equating the wave functions at boundaries of the well yields $D = \pm A$ for the even and odd parity solutions, respectively. Thus the **even parity solutions of the well** are given by

$$\psi_I(x) = A \exp(\gamma x) \quad \text{for } x < -a \tag{8.35a}$$

$$\psi_{II}(x) = B \cos(qx) \quad \text{for } -a < x < a \tag{8.35b}$$

and

$$\psi_{III}(x) = A \exp(-\gamma x) \quad \text{for } x > a \tag{8.35c}$$

On the other hand, the **odd parity solutions** are given by

$$\psi_I(x) = D \exp(\gamma x) \quad \text{for } x < -a \tag{8.36a}$$

$$\psi_{III}(x) = C \sin(qx) \quad \text{for } -a < x < a \tag{8.36b}$$

and

$$\psi_{III}(x) = -D \exp(-\gamma x) \quad \text{for } x > a \tag{8.36c}$$

Since ψ and $\partial\psi/\partial x$ have to be continuous, the logarithmic derivative $\partial/\partial x$ (ln ψ), i.e., $1/\psi(\partial\psi/\partial x)$ too has to be continuous at the boundaries. Imposing the continuity condition of the logarithmic derivatives at $x = \pm a$ on the even parity solution we obtain the condition that

$$\eta = \xi \tan \xi \tag{8.37}$$

where

$$\eta = \gamma a \text{ and } \xi = qa \tag{8.38}$$

You may like to prove Eq. (8.37) before studying further.

SAQ 5

Spend 5 min

Prove Eq. (8.37).

Eqs. (8.38) and (8.32) yield

$$\eta^2 + \xi^2 = 2m V_0 a^2/\hbar^2 = R^2 \tag{8.39}$$

The energy eigenvalues can now be obtained by solving Eqs. (8.37) and (8.39) for η and ξ, from which γ and q may be determined. A graphical method of solving these equations is illustrated in Fig. 8.8a.

The possible values of η and ξ and hence of γ and q are obtained from the intersections of the circle $\eta^2 + \xi^2 = R^2$ (where $R^2 = 2mV_0a^2/\hbar^2$) with the curve $\eta = \xi \tan \xi$. Since γ and q are real, the circle and the curve are to be considered in the first quadrant only. It is clear from the figure that the allowed values of ξ and hence E are discrete and the number of allowed values of E increases as R increases. It is easy to see that for $R = 1$, there is only one solution. The same is true for $R = 2$ but at $R = 3.5$ the number of even parity solutions are two.

A similar analysis for the odd parity solutions yields

$$\eta = -\xi \cot \xi \tag{8.40}$$

For the above case $R = 1$ yields no solution; however $R = 2$ and 3 support one bound state (Fig. 8.8b).

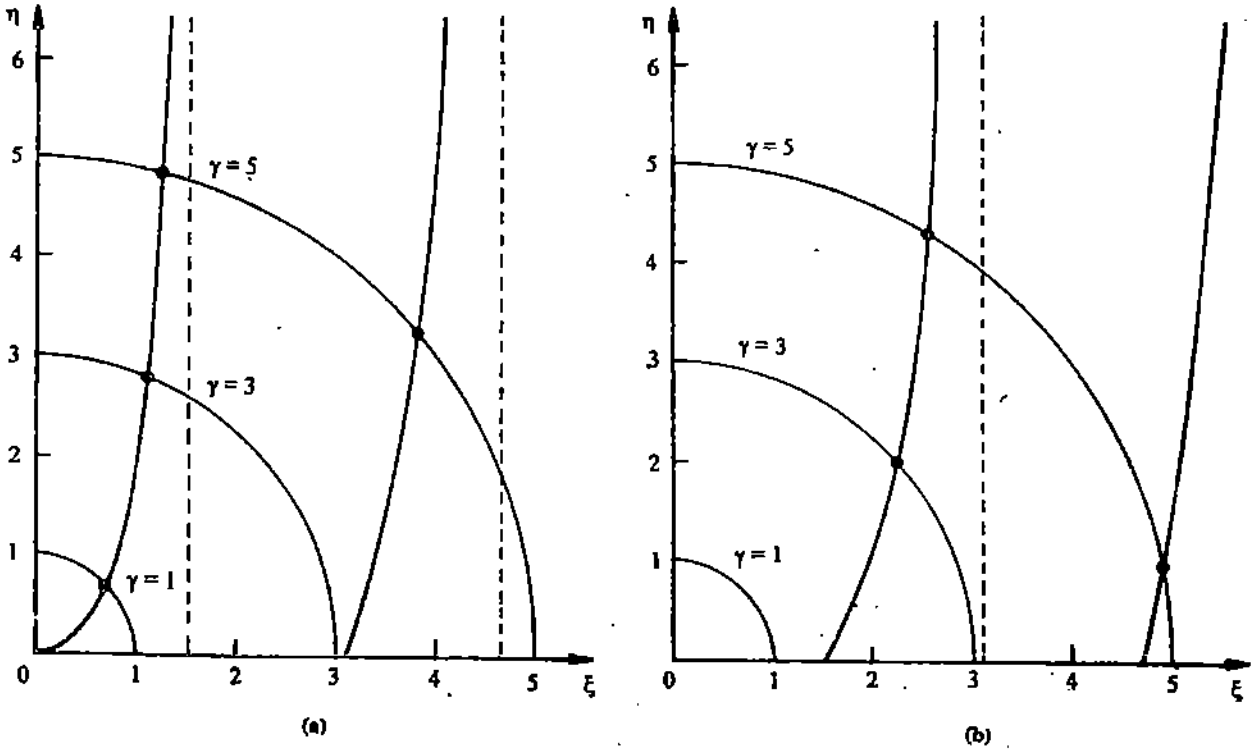


Fig. 8.8: (a) Variation of η with ξ in the two equations $\eta^2 + \xi^2 = R^2$ and $\eta = \xi \tan \xi$. The intersections give the solutions; (b) Variation of η with ξ in the two equations $\eta^2 + \xi^2 = R^2$ and $\eta = -\xi \cot \xi$. The intersections give the solutions.

We summarize that for $E < V_0$, the energy levels of a particle in a potential well depend upon the well parameters V_0 and a. For R lying between 0 and $\pi/2$, i.e., for V_0 (a^2) lying between 0 and $(\hbar^2/2m)\cdot(\pi/2)^2$ there is just one energy level of even class; for R between $(\pi/2)$ and $2(\pi/2)$ there is one even and one odd class solution. As R increases the total number of energy levels increases. Thus for $E < V_0$ the particle is bound to the well and the energy spectrum is discrete. On the other hand for $E > V_0$ we can show that the particle is in the continuum state and eigenenergy E varies continuously from V_0 to ∞.

Finally, let us consider a potential well having infinite depth. In this case the probability of finding the particle outside the well will be zero and the wave function will be non-zero only for $|x| < a$. The even and odd parity solutions will again be given by Eqs. (8.35b) and (8.36b), respectively. Equating $\psi_n(x)$ at $x = \pm a$ to zero yields

$$qa = (n + 1)\,\pi/2 \tag{8.41}$$

where n is a positive integer including zero. Putting Eq. (8.41) into (8.32) we obtain

$$E_n = \frac{\hbar^2\,\pi^2\,(n+1)^2}{8m\,a^2} \qquad \text{with } n = 0, 1, 2, \ldots \tag{8.42}$$

It is interesting to note that even for $n = 0$ the energy is finite. This energy E_0 $(= \hbar^2/8ma^2)$ is known as **zero point energy** and is a consequence of the Heisenberg uncertainty principle. Since the width of the well is '$2a$', the maximum uncertainty in the position of the particle is $2a$. Hence the uncertainty in the momentum of the particle is $\hbar/2a$. This may be taken as the minimum momentum of the particle. Thus the minimum energy of the particle will be $\hbar^2/8ma^2$. This energy is of the same order as E_0. Hence the zero point energy is consistent with the uncertainty principle.

You should note that in the present problem we have only one constant of motion, namely, total energy. Hence, the energy levels are specified by only one quantum number n. Again although $E < V_0$, the probability of finding the particle outside the well is non-zero. This result is non-classical. However, you should remember that this probability decreases exponentially with the increase of $|x|$.

Let us now consider the case of the one dimensional harmonic oscillator.

8.6 ONE DIMENSIONAL HARMONIC OSCILLATOR

The problem of a one-dimensional harmonic oscillator is of direct physical interest. Actually, a large number of systems are governed exactly or approximately by the harmonic oscillator equation. Recall the classical definition of a harmonic oscillator: It obeys Hooke's law according to which the force **F** on the particle is directly proportional to the displacement and is always directed towards the mean position, i.e., $\mathbf{F} = -k\mathbf{x}$: k, the constant of proportionality, is known as force constant (Fig. 8.9). It is related to the classical frequency ν of the oscillator and is given by $k = 4\pi^2 \nu^2 m$ where m is the mass of the particle. The potential energy of the particle at x is given by $\frac{1}{2} kx^2$ or $\frac{1}{2} m \omega^2 x^2$, where $\omega = 2\pi\nu$ is the angular frequency. Hence the time independent Schrödinger equation is

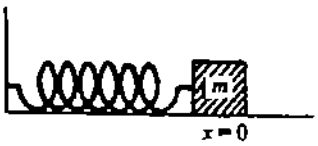


Fig. 8.9: Mass on a spring as a one-dimensional harmonic oscillator.

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2} m\omega^2 x^2\right] \psi(x) = E\psi(x) \tag{8.43}$$

where E is the total energy of the oscillator and is independent of time. It is evident from the above equation that the Hamiltonian of the system is invariant under space inversion, i.e., it commutes with the parity operator. Hence the eigenfunctions are of definite parity.

Now we define

$$\xi = ax, \text{ with } a^2 = m\omega/\hbar \tag{8.44}$$

and

$$E = \frac{\lambda\hbar\omega}{2}$$

With the help of the above definitions, Eq. (8.43) reduces to

$$\frac{d^2\psi(\xi)}{d\xi^2} + (\lambda - \xi^2)\, \psi(\xi) = 0 \tag{8.45}$$

You should quickly check out Eq. (8.45).

SAQ 6

Spend 5 min

Verify Eq. (8.45).

To obtain acceptable solutions of Eq. (8.45) we are required to go through a fairly lengthy algebra which is unnecessary at this stage. Here we simply state the results. For the solution to be acceptable λ must satisfy the relation

$$\lambda = 2n + 1, \qquad n = 0, 1, 2, \ldots \tag{8.46}$$

Hence from Eq. (8.46) it follows that

$$E_n = (n + \frac{1}{2})\, \hbar\omega, \qquad n = 0, 1, 2, \ldots \tag{8.47}$$

The integer n is known as **energy quantum number**. Since our simple harmonic oscillator is a one-dimensional system we have only one quantum number. Notice from Fig. 8.10 that

- The energy levels of a quantum mechanical oscillator are equally spaced. This is a characteristic of some parts of the molecular and nuclear experimental spectra. The harmonic oscillator provides a good model description of these spectra, so much so that the spectra are referred to as vibrational spectra. There are also excitations in solids called phonons that fall in the same category.

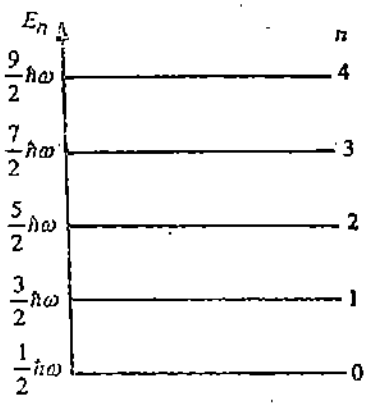


Fig. 8.10 : Energy level diagram of a simple harmonic oscillator.

- For each eigenvalue, there will be only one eigenfunction. Thus there is no degeneracy. This property seems to be a common characteristic of bound states for one-dimensional potentials that remain finite for all finite values of x.

We have zero point energy given by

$$E_0 = \frac{\hbar\omega}{2} \qquad \text{for } n = 0 \tag{8.48}$$

This zero point energy is again a consequence of the Heisenberg uncertainty principle. This can be seen as follows. Since

$$E = \frac{p^2}{2m} + \frac{1}{2}kx^2$$

E can be zero only when p and x both are equal to zero simultaneously. Under such a circumstance p and x will become definite (equal to zero) simultaneously. This will violate uncertainty principle. Therefore, the lowest eigenenergy has to be non-zero.

The normalised eigenfunctions of a simple harmonic oscillator are given by

$$\psi_n(x) = \left(\frac{a}{\sqrt{\pi}\, 2^n\, n!}\right)^{1/2} H_n(ax) \exp(-a^2x^2/2), \qquad n = 0, 1, 2, \ldots \tag{8.49}$$

where $H_n(ax)$ are Hermite polynomials. A few of the lower order $H_n(\xi)$ are given by

$$\begin{aligned} H_0(\xi) &= 1; \\ H_1(\xi) &= 2\xi; \\ H_2(\xi) &= 4\xi^2 - 2; \\ H_3(\xi) &= 8\xi^3 - 12\xi; \\ H_4(\xi) &= 16\xi^4 - 48\xi^2 + 12; \\ H_5(\xi) &= 32\xi^5 - 160\xi^3 + 120\xi. \end{aligned} \tag{8.50}$$

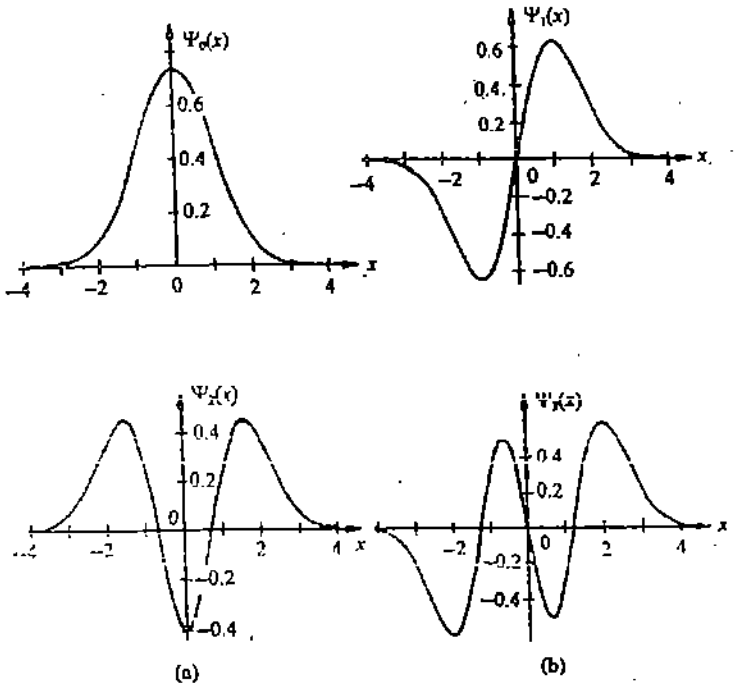


Fig. 8.11 : (a) Even parity solutions and (b) odd parity solutions for the one-dimensional harmonic oscillator.

We can show that the eigenfunctions of the Hamiltonian are also eigenfunctions of the parity operator. It is evident from Eqs. (8.49) and (8.50) that the eigenfunctions corresponding to zero or even value of n are of even parity. On the other hand odd parity eigenfunctions have odd values of n. The variation of even parity functions for $n = 0,2$ and 4 with ξ are shown in Fig. (8.11a). The odd parity functions for $n = 1$, 3 and 5 are shown in Fig. (8.11b).

Let us now compare the quantum oscillator with the classical oscillator. Let us first take up the question of time dependence. Classically, the simple harmonic oscillator oscillates in such a manner that the position of the particle represented by the oscillator changes from one moment to another. Quantum mechanics, on the other hand, tells us that for any state of energy E, although there is a distribution of probabilities for various positions, this distribution is constant as far as time is concerned; (these probabilities are 'frozen' in time). This is the usual meaning of energy eigenstates being stationary. Is it possible to reconcile these two very different pictures?

The answer lies in considering not one single eigenfunction but a superposition of eigenfunctions as in a wave packet. Consider, for example, the superposition $\psi(x, t)$ of the first two oscillator eigenfunctions:

$$\psi(x, t) = \frac{1}{\sqrt{2}} [\exp(-iE_0 t/\hbar)\, \psi_0(x) + \exp(-iE_1 t/\hbar)\, \psi_1(x)]$$

If we plot $|\psi(x, t)|^2$, we get Fig. 8.12, where the plot is made for four different values of time. Clearly, the probability oscillates with time with just the frequency of the harmonic oscillator as expected classically. It is therefore reasonable to expect that when we take a superposition of a large number of oscillator eigenfucntions we will get a classical behaviour.

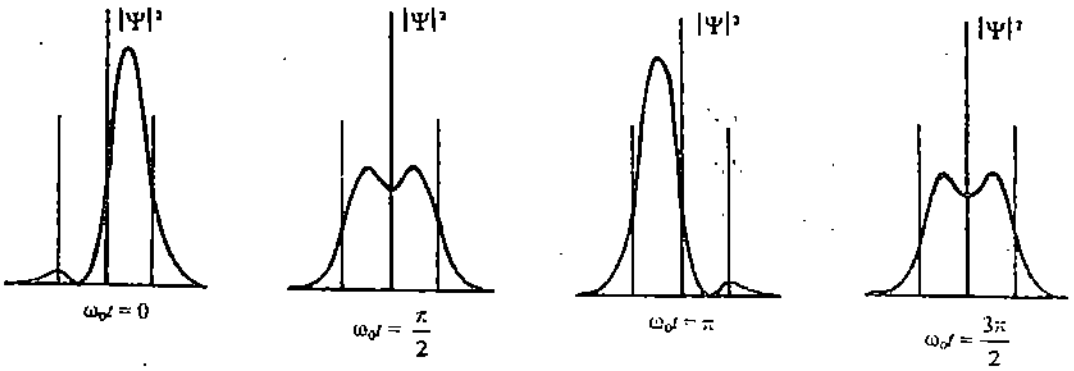


Fig. 8.12 : The probability corresponding to the superposition of the first two oscillator eigenfunctions of equal amplitude (with their time dependence included) plotted at four different times. Classical oscillatory behaviour is clearly seen. The vertical lines indicate the classical limits of motion, assuming an energy $E = \langle H \rangle = \hbar\omega$.

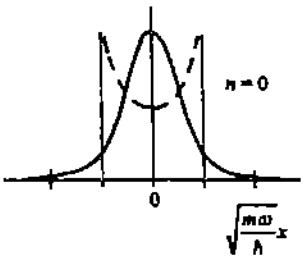
You should be very clear however, that the quantum solution of the harmonic oscillator is radically different from that for the classical oscillator. In classical mechanics, the oscillator is forbidden to go beyond the potential, beyond the turning points where its kinetic energy turns negative. But clearly, the quantum wave functions extend beyond the potential, and thus there is a finite probability for the oscillator to be found in a classically forbidden region.

To be specific, let us compare the quantum and classical probabilities for the states corresponding to $n = 0$ and $n = 1$. The quantum probabilities are easily calculated by taking the square of the appropriate wave functions, ψ_0 and ψ_1.

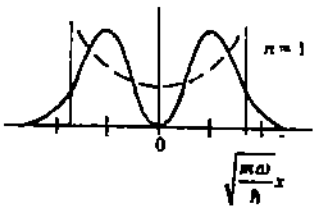
The expression for the probability of finding a classical harmonic oscillator of quantum number n and energy E governed by the equation $x = A \sin \omega t$ in a region Δx is given by

$$P(x)\,\Delta x = \frac{1}{2\pi A} \frac{1}{(1 - x^2/A^2)^{1/2}}\, \Delta x$$

where $A = (2E/m\omega^2)^{1/2}$. As expected, the classical probability is non-zero only for $-A < x < A$; the oscillator is confined within the turning points. For $x > A$, it is clear that the potential energy $\frac{1}{2} m\omega^2 x^2 > E$, and classically this is impossible.



(a)



(b)

Fig. 8.13 : Comparison of quantum (solid curve) and classical probability densities (dashed curve) for the harmonic oscillator for the two oscillator states having the same total energy, corresponding to $n = 0$ and $n = 1$.

The quantum and classical probabilities are compared in Figs. 8.13 (a) and (b) for $n = 0$ and $n = 1$, respectively. In both cases, the quantum probability does not vanish in the classically forbidden region. For $n = 1$, the classical probability is maximum at the turning points. But the quantum probability reaches the maximum much closer to the point of equilibrium. For large n, the average of the quantum mechanical probability distribution is found to be given by the classical probability curve.

Let us now consider an application of the simple harmonic oscillator to physical systems: the vibrations of the two atoms of a diatomic molecule.

Example 2: Diatomic Molecule

Let us assume that both the atoms in the molecule execute simple harmonic motion about their equilibrium positions. Hence, they satisfy the following equations

$$M_1 \frac{d^2 R_1}{dt^2} = -k\,(R - R_e) \tag{8.51}$$

and

$$M_2 \frac{d^2 R_2}{dt^2} = -k\,(R - R_e)$$

where R_1 and R_2 are the distance of the two atoms A and B from their centre of mass and $R = R_1 + R_2$ (Fig. 8.14). The masses of the two atoms are M_1 and M_2, k is the force constant and at equilibrium, the distance between the two atoms is R_e. Considering the moment about A (see Fig. 8.14) we obtain

$$M_2 R = (M_1 + M_2)\, R_1 \tag{8.52}$$

Putting Eq. (8.52) in Eq. (8.51) we obtain

$$\frac{M_1 M_2}{M_1 + M_2} \frac{d^2 R}{dt^2} = -k(R - R_e) \tag{8.53}$$

or

$$\mu \frac{d^2 x}{dt^2} = -k\,x \tag{8.54}$$

with $x = R - R_e$.

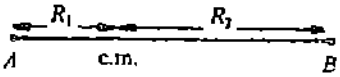


Fig. 8.14 : Vibration of two atoms A and B in a diatomic molecule.

Thus we have reduced a two body problem of masses M_1 and M_2 to a one-body problem of mass μ and the whole molecule behaves as a simple harmonic oscillator of mass μ (*reduced mass* of the molecule) having force constant k. Hence the eigenfunctions and eigenvalues of the molecules are given by Eq. (8.49) and Eq. (8.47), respectively. These equations have been very useful in understanding the vibrational spectrum (obtained in the near infrared region of the electromagnetic waves) of diatomic molecules. The analysis of the experimental spectrum has yielded force constants of a large number of heteronuclear molecules.

Would you like to apply the ideas discussed so far to some concrete situations? Try the following SAQ.

Spend 20 min

SAQ 7

(a) Consider a proton as a bound oscillator with a natural frequency of 3×10^{21} Hz. What is the energy of its ground and first excited states?

(b) Calculate $\langle x \rangle$ and $\langle p_x \rangle$ for the ground state harmonic oscillator eigenfunctions.

Let us now summarise what you have studied in this unit.

8.7 SUMMARY

In this unit you have solved time independent Schrödinger equation for a number of simple one-dimensional conservative systems. Some of the important results are summarised below:

- The eigenenergy E of a **free particle** moving in a one-dimensional space is given by $E = \frac{\hbar^2 k^2}{2m}$. It can take any value continuously from 0 to ∞. Hence its energy **spectrum is continuous**. The eigenfunctions are incoming and outgoing plane waves given by $\psi_{\pm k}(x) = Ae^{\pm ikx}$ and are unnormalizable. However, if the particle is confined to a finite line segment then the eigenenergies vary in a discrete manner and the bound state eigenfunctions can be normalised. The plane waves can also be **box-normalised**.

- For a **one-dimensional rectangular potential barrier**, the eigenenergies vary continuously from 0 to ∞. However, unlike in classical mechanics the quantum mechanical probabilities of reflection and transmission of the particle by the barrier are, in general, finite. Thus even for $E < V_0$, the particle can **tunnel** through the barrier. Another interesting result is that for

$$E = \frac{1}{2m}\left(\frac{n\hbar\pi}{2a}\right)^2 + V_0$$

there is a hundred percent transmission with no reflection. Here symbols have the meanings discussed in the text.

- For a **one-dimensional potential well** the eigenenergy spectrum breaks into two parts. For $E > V_0$, the eigenenergy varies in a continuous manner. On the other hand for $E < V_0$ the eigenenergy varies in a discrete manner and bound states are obtained. The lowest eigenenergy is non-zero and is in accordance with the Heisenberg uncertainty relation. The number of permissible bound states increase with V_0 : the eigenfunctions being alternately of even and odd parities, the lowest being of even parity.

- for a **particle executing simple harmonic motion** along a line all the eigenstates are bound states and eigenenergies vary in a discrete manner. Consecutive eigenenergy states are separated from each other by the same amount $\hbar\omega$. The eigenfunctions are given in terms of Hermite polynomials and are of definite parity, alternately of even and odd parities. The finite value of the lowest eigenenergy is again a consequence of the uncertainty principle. The results obtained for a simple harmonic oscillator can be applied to study the vibrational spectrum of diatomic molecules.

8.8 TERMINAL QUESTIONS

Spend 45 min

1. The wave function of a particle of mass m inside an infinite square well of width $2a$ ($-a$ to $+a$) is given by

$$\psi(x) = A\cos\frac{3\pi x}{2a} + B\sin\frac{3\pi x}{2a}$$

Obtain the values of A and B and the eigenenergy corresponding to the above eigenfunciton.

2. The potential energy of a particle of mass m is given by

$$V(x) = \frac{1}{2}m\omega^2 x^2 \qquad \text{for } x > 0$$
$$= \infty \qquad \text{for } x < 0.$$

Show that its eigenenergies are given by

$$E_{2m+1} = 2\left(m + \frac{3}{4}\right)\hbar\omega \text{ for } m = 0, 1, 2, \ldots.$$

3. Show that the average value of x for a simple harmonic oscillator in the n^{th} quantum state is zero.

4. Calculate the mean kinetic and potential energies of a simple harmonic oscillator which is in its ground state.

8.9 SOLUTIONS AND ANSWERS

Self-Assessment Questions

1. $$p_{op}e^{\pm ikx} = -i\hbar\frac{\partial}{\partial x}e^{\pm ikx}\left[\because p_{op} = -i\hbar\frac{\partial}{\partial x}\right]$$
$$= -i\hbar(\pm ik)\,e^{\pm ikx}$$
$$= \pm k\hbar\, e^{\pm ikx}$$

2. (a) The normalisation condition is
$$\int_{-\infty}^{\infty}\psi_n^*(x)\,\psi_n(x)\,dx = 1$$
or
$$N^2\int_0^L \sin^2\left(\frac{n\pi x}{L}\right)dx = 1 \qquad [\because \psi_n(x) \text{ is finite only for } 0 < x < L]$$
or
$$N^2\left(\frac{L}{2}\right) = 1 \qquad \left[\because \int_0^L \sin^2\left(\frac{n\pi x}{L}\right)dx = \frac{L}{2}\right]$$
or
$$N = \left(\frac{2}{L}\right)^{1/2}$$

(b) To prove that $\psi_m(x)$ and $\psi_n(x)$ are orthogonal for $m \neq n$, we have to show that
$$I = \int_{-\infty}^{\infty}\psi_m^*(x)\psi_n(x)\,dx = 0 \text{ for } m \neq n$$
or
$$I = N^2\int_{-\infty}^{\infty}\sin\frac{n\pi x}{L}\sin\frac{m\pi x}{L}\,dx = \int_0^L \sin\frac{n\pi x}{L}\sin\frac{m\pi x}{L}\,dx$$
$$= 0, \text{ for } m \neq n \quad \left[\because \int_0^L \sin\frac{n\pi x}{L}\sin\frac{m\pi x}{L}\,dx = 0 \text{ for } m \neq n\right]$$

3. A three-dimensional extension of Eq. (8.3a) is
$$\left[\frac{d^2}{dx^2}+\frac{d^2}{dy^2}+\frac{d^2}{dz^2}\right]\psi(x, y, z) = -(k_x^2 + k_y^2 + k_z^2)\,\psi(x, y, z)$$

Since x, y, z are independent variables, we may write
$$\psi(x, y, z) = \phi_{n_x}(x)\,\phi_{n_y}(y)\,\phi_{n_z}(z) \tag{1}$$
where
$$\phi_{n_x}(x) = \left(\frac{2}{L}\right)^{1/2}\sin\left(\frac{n_x\pi x}{L_x}\right) \text{ etc.} \tag{2}$$
and
$$E_{n_x, n_y, n_z} = \frac{\hbar^2\pi^2}{2m}\left(\frac{n_x^2}{L_x^2}+\frac{n_y^2}{L_y^2}+\frac{n_z^2}{L_z^2}\right) \tag{3}$$

Here n_x, n_y and n_z are positive integers ranging from zero to infinity, but all the three can't be zero simultaneously. It is evident from (3) that E_{n_x, n_y, n_z} takes only discrete values. Hence the energy spectrum of a particle enclosed in a box is discrete and the eigenfunctions given by (1) and (2) form an orthonormal set.

4. Applying the boundary conditions given by Eqs. (8.25a and b) to the wave functions ψ_I, ψ_{II} and ψ_{III} we obtain the following set of equations:
$$Ae^{-ika} + Be^{ika} = Ce^{\gamma a} + De^{-\gamma a} \tag{1}$$
$$Ae^{-ika} - Be^{ika} = \frac{i\gamma}{k}(Ce^{\gamma a} - De^{-\gamma a}) \tag{2}$$
$$Ce^{-\gamma a} + De^{\gamma a} = Fe^{ika} \tag{3}$$
$$Ce^{-\gamma a} - De^{\gamma a} = -\frac{i\gamma}{k}Fe^{ika} \tag{4}$$

To obtain P_r we have to eliminate C and D from Eqs. (1 and 2). Adding Eqs. 1 and 2 and subtracting Eq. 2 from Eq. 1, we get

$$2Ae^{-ika} = Ce^{\gamma a}\left(1 + \frac{i\gamma}{k}\right) + De^{-\gamma a}\left(1 - \frac{i\gamma}{k}\right) \tag{5}$$

$$2Be^{ika} = Ce^{\gamma a}\left(1 - \frac{i\gamma}{k}\right) + De^{-\gamma a}\left(1 + \frac{i\gamma}{k}\right) \tag{6}$$

Similarly, adding (3) and (4) and subtracting (4) from (3), we get

$$2Ce^{-\gamma a} = Fe^{ika}\left(1 - \frac{ik}{\gamma}\right)$$

$$2De^{\gamma a} = Fe^{ika}\left(1 + \frac{ik}{\gamma}\right)$$

whence

$$Ce^{\gamma a} = \frac{Fe^{ika+2\gamma a}}{2}\left(1 - \frac{ik}{\gamma}\right)$$

and

$$De^{-\gamma a} = \frac{Fe^{ika-2\gamma a}}{2}\left(1 + \frac{ik}{\gamma}\right)$$

Substituting these expressions in Eqs. (5) and (6) we get

$$2Ae^{-ika} = \frac{F}{2}e^{ika}\left[e^{2\gamma a}\left(1 - \frac{ik}{\gamma}\right)\left(1 + \frac{i\gamma}{k}\right) + e^{-2\gamma a}\left(1 + \frac{ik}{\gamma}\right)\left(1 - \frac{i\gamma}{k}\right)\right] \tag{7}$$

and

$$2Be^{ika} = \frac{F}{2}e^{ika}\left[e^{2\gamma a}\left(1 - \frac{ik}{\gamma}\right)\left(1 - \frac{i\gamma}{k}\right) + e^{-2\gamma a}\left(1 + \frac{ik}{\gamma}\right)\left(1 + \frac{i\gamma}{k}\right)\right] \tag{8}$$

Dividing Eq. (8) by (7) we have

$$\frac{B}{A}e^{2ika} = \frac{e^{2\gamma a}(\gamma - ik)(k - i\gamma) + e^{-2\gamma a}(\gamma + ik)(k + i\gamma)}{e^{2\gamma a}(\gamma - ik)(k + i\gamma) + e^{-2\gamma a}(\gamma + ik)(k - i\gamma)}$$

$$= \frac{-ie^{2\gamma a}(\gamma^2 + k^2) + ie^{-2\gamma a}(\gamma^2 + k^2)}{-ie^{2\gamma a}(k + i\gamma)^2 + ie^{-2\gamma a}(k - i\gamma)^2}$$

$$= \frac{(\gamma^2 + k^2)(e^{2\gamma a} - e^{-2\gamma a})}{e^{2\gamma a}(k + i\gamma)^2 - e^{-2\gamma a}(k - i\gamma)^2}$$

$$= \frac{2(\gamma^2 + k^2)\sinh 2\gamma a}{(k^2 - \gamma^2)(e^{2\gamma a} - e^{-2\gamma a}) + 2ik\gamma(e^{2\gamma a} + e^{-2\gamma a})}$$

$$= \frac{2(\gamma^2 + k^2)\sinh 2\gamma a}{2(k^2 - \gamma^2)\sinh 2\gamma a + 4ik\gamma\cosh 2\gamma a}$$

Now

$$P_r = \left|\frac{B}{A}\right|^2 = \frac{(\gamma^2 + k^2)^2 \sinh^2 2\gamma a}{(k^2 - \gamma^2)^2 \sinh^2 2\gamma a + 4k^2\gamma^2 \cosh^2 2\gamma a}$$

$$= \frac{(\gamma^2 + k^2)^2 \sinh^2 2\gamma a}{(k^4 + \gamma^4 - 2k^2\gamma^2)\sinh 2\gamma a + 4k^2\gamma^2 \cosh^2 2\gamma a}$$

$$= \frac{(\gamma^2 + k^2)^2 \sinh^2 2\gamma a}{(k^4 + \gamma^4 + 2k^2\gamma^2)\sinh^2 2\gamma a + 4k^2\gamma^2}$$

where we have added and subtracted $4k^2\gamma^2 \sinh^2 2\gamma a$ in the denominator and used the relation $\cosh^2\theta - \sinh^2\theta = 1$.

or $$P_r = \frac{(\gamma^2 + k^2)^2 \sinh^2 2\gamma a}{(\gamma^2 + k^2)^2 \sinh^2 2\gamma a + 4k^2\gamma^2}$$

Now $$\gamma^2 + k^2 = \frac{2m}{\hbar^2}(V_0 - E + E) = \frac{2m}{\hbar^2} V_0$$

and $$\gamma^2 k^2 = \left(\frac{2m}{\hbar^2}\right)^2 E(V_0 - E)$$

Therefore

$$P_r = \frac{V_0^2 \sinh^2 2\gamma a}{V_0^2 \sinh^2 2\gamma a + 4E(V_0 - E)}$$

$$= \left[\frac{V_0^2 \sinh^2 2\gamma a + 4E(V_0 - E)}{V_0^2 \sinh^2 2\gamma a}\right]^{-1}$$

$$= \left[1 + \frac{4E(V_0 - E)}{V_0^2 \sinh^2 2\gamma a}\right]^{-1}$$

To obtain an expression for P_t, we can use Eq. (7) whence

$$\frac{F}{A} = \frac{4e^{-2ika}}{e^{2\gamma a}\left(1 - \frac{ik}{\gamma}\right)\left(1 + \frac{i\gamma}{k}\right) + e^{-2\gamma a}\left(1 + \frac{ik}{\gamma}\right)\left(1 - \frac{i\gamma}{k}\right)}$$

$$\frac{F}{A} = \frac{4e^{-2ika}}{e^{2\gamma a}\left(2 - \frac{ik}{\gamma} + \frac{i\gamma}{k}\right) + e^{-2\gamma a}\left(2 + \frac{ik}{\gamma} - \frac{i\gamma}{k}\right)}$$

$$= \frac{4e^{-2ika}}{4\cosh 2\gamma a + 2i\left(\frac{\gamma}{k} - \frac{k}{\gamma}\right)\sinh 2\gamma a}$$

Thus

$$P_t = \left|\frac{F}{A}\right|^2 = \frac{4}{4\cosh^2 2\gamma a + \left(\frac{\gamma}{k} - \frac{k}{\gamma}\right)^2 \sinh^2 2\gamma a}$$

Now $$\left(\frac{\gamma}{k} - \frac{k}{\gamma}\right)^2 = \frac{(\gamma^2 - k^2)^2}{k^2\gamma^2} = \frac{\left(\frac{2m}{\hbar^2}\right)^2 (V_0 - E - E)^2}{\left(\frac{2m}{\hbar^2}\right)^2 (V_0 - E)E}$$

$$= \frac{(V_0 - 2E)^2}{E(V_0 - E)}$$

Therefore,

$$P_t = \frac{4E(V_0 - E)}{4\cosh^2 2\gamma a\,(EV_0 - E^2) + (V_0^2 + 4E^2 - 4V_0E)\sinh^2 2\gamma a}$$

$$= \frac{4E(V_0 - E)}{4EV_0 - 4E^2 + V_0^2 \sinh^2 2\gamma a}$$

$$= \frac{4E(V_0 - E)}{4E(V_0 - E) + V_0^2 \sinh^2 2\gamma a}$$

$$= \left[\frac{4E(V_0 - E) + V_0^2 \sinh^2 2\gamma a}{4E(V_0 - E)} \right]^{-1}$$

$$= \left[1 + \frac{V_0^2 \sinh^2 2\gamma a}{4E(V_0 - E)} \right]^{-1}$$

5. The condition

$$\frac{1}{\psi_I} \left. \frac{\partial \psi_I}{\partial x} \right|_{x=-a} = \frac{1}{\psi_{II}} \left. \frac{\partial \psi_{II}}{\partial x} \right|_{x=-a}$$

yields

$$\frac{-\gamma A e^{-\gamma a}}{A e^{-\gamma a}} = \frac{-q B \sin qa}{B \cos qa}$$

or

$$-\gamma a = -qa \tan qa$$

or

$$\eta = \xi \tan \xi$$

where

$$\eta = \gamma a \text{ and } \xi = qa$$

6. Since $\xi = ax$

$$\frac{d\psi}{dx} = \frac{d\psi}{d\xi}\frac{d\xi}{dx} = a\frac{d\psi}{d\xi}$$

and

$$\frac{d^2\psi}{dx} = a^2 \frac{d^2\psi}{d\xi^2}$$

Thus Eq. (8.43) may be written as

$$-\frac{\hbar^2 a^2}{2m}\frac{d^2\psi}{d\xi^2} + \frac{1}{2}\frac{m\omega^2}{a^2}\xi^2\psi = \frac{\lambda\hbar\omega}{2}\psi(\xi)$$

or

$$-\frac{\hbar^2 m\omega}{2m\hbar}\frac{d^2\psi}{d\xi^2} + \frac{1}{2}\frac{m\omega^2\hbar}{m\omega}\xi^2\psi = \frac{\lambda\hbar\omega}{2}\psi$$

or

$$-\frac{\hbar\omega}{2}\frac{d^2\psi}{d\xi^2} + \frac{\hbar\omega}{2}\xi^2\psi = \frac{\lambda\hbar\omega}{2}\psi$$

or

$$\frac{d^2\psi}{d\xi^2} + (\lambda - \xi^2)\psi(\xi) = 0$$

7. (a) $E_n = \left(n + \frac{1}{2}\right)\hbar\omega$

For the ground state $n = 0$, $E_0 = \frac{1}{2}\hbar\omega = \frac{1}{2}h\nu$

$$= \frac{1}{2} \times 6.626 \times 10^{-34} \text{ Js} \times 3 \times 10^{21} \text{ Hz}$$

$$= 9.939 \times 10^{-13} \text{ J}$$

For the first excited state, $n = 1$

$$E_1 = \frac{3}{2}\hbar\omega = \frac{3}{2}h\nu$$

$$= 3E_0 = 2.982 \times 10^{-12} \text{ J}.$$

(b) For the ground state harmonic oscillator wavefunction

$$\langle x \rangle = \int_{-\infty}^{\infty} dx\, \psi_0^*(x)\, x\, \psi_0(x)\, dx$$

where $\psi_0(x) = \left(\frac{a}{\sqrt{\pi}}\right)^{1/2} H_0(ax) \exp\left(-\frac{a^2x^2}{2}\right)$

$$= \left(\frac{a}{\sqrt{\pi}}\right)^{1/2} \exp\left(-\frac{a^2x^2}{2}\right) \quad \text{since } H_0(\xi) = 1$$

Thus

$$\langle x \rangle = \left(\frac{a}{\sqrt{\pi}}\right) \int_{-\infty}^{\infty} x \exp(-a^2x^2)\, dx$$

The integrand is an odd function. Hence the integral over the interval will be zero.

$$\therefore \langle x \rangle = 0$$

Similarly

$$\langle p_x \rangle = -i\hbar \int_{-\infty}^{\infty} \psi_0^*(x) \frac{\partial}{\partial x} \psi_0(x)\, dx$$

$$= -i\hbar \left(\frac{a}{\sqrt{\pi}}\right) \int_{-\infty}^{\infty} \exp\left(-\frac{a^2x^2}{2}\right) (-a^2x) \exp\left(-\frac{a^2x^2}{2}\right) dx$$

$$= i\frac{a^3}{\sqrt{\pi}}\hbar \int_{-\infty}^{\infty} x \exp(-a^2x^2)\, dx$$

Again the integrand is an odd function of x. Hence the integral over this interval will be zero.

$$\therefore \langle p_x \rangle = 0$$

Terminal Questions

1. Since the well is of infinite depth $\psi(+a) = 0$. Hence $B = 0$. A is obtained as follows:

$$A^2 \int_{-a}^{+a} \cos^2 \frac{3\pi x}{2a}\, dx = 1$$

or

$$A = \left(\frac{1}{a}\right)^{1/2}$$

Now inside the well $V(x) = 0$ and the Schrödinger equation yields

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\left(A \cos\frac{3\pi x}{2a}\right) = +A\frac{\hbar^2}{2m}\left(\frac{3\pi}{2a}\right)^2 \cos\left(\frac{3\pi x}{2a}\right).$$

$$\therefore E = \frac{9\pi^2\hbar^2}{8m\,a^2}$$

2. The problem is similar to that of a simple harmonic oscillator but

$$\psi(0) = 0.$$

Hence n must be odd (See Eq. 8.50).

Therefore,

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega, \text{ with } n = 1, 3, 5, \ldots$$

Putting $n = 2m + 1$ we get

$$E_{2m+1} = 2\left(m + \frac{3}{4}\right)\hbar\omega \text{ with } m = 0, 1, 2, 3, \ldots$$

3. $\langle x \rangle = (\psi_n, x\psi_n) = \int_{-\infty}^{\infty} \psi_n^*(x)\, x\, \psi_n(x)\, dx$

Since $\psi_n(x)$ is of definite parity, x is either an odd or an even function of x. In either case, the product $\psi_n^*(x)\,\psi_n(x)$ will be even. Since x is odd, the integrand will be an odd function of x and hence the integral will be zero.

4. The average value of the potential energy V is given by

$$\langle V \rangle = \int_{-\infty}^{\infty} \psi_0^*(x) \frac{1}{2} kx^2 \psi_0(x)\, dx.$$

Evaluation of the integral with $\psi_0 = \left(\frac{a}{\sqrt{\pi}}\right)^{1/2} \exp\left(-\frac{a^2 x^2}{2}\right)$ yields

$$\langle V \rangle = \frac{1}{4} \hbar\omega$$

Since $E_0 = \frac{1}{2} \hbar\omega$

$$\therefore \langle \text{K.E.} \rangle = E_0 - \frac{1}{4} \hbar\omega = \frac{1}{4} \hbar\omega.$$

UNIT 9 SPHERICALLY SYMMETRIC SYSTEMS: HYDROGEN ATOM

Structure

9.1 Introduction
Objectives
9.2 Three Dimensional Schrödinger Equation for a Central Potential
Eigenfunctions and Eigenvalues of the Angular Momentum Operator
Space Quantization
Radial Eigenfunctions
9.3 The Hydrogen Atom
Spectra of the Hydrogen Atom
Quantum Numbers and Constants of Motion
9.4 Summary
9.5 Terminal Questions
9.6 Solutions and Answers

9.1 INTRODUCTION

In the previous unit you have obtained eigenfunctions and eigenvalues of a number of one-dimensional systems. In this unit, we shall extend our study to a three-dimensional system. Thus now there will be three-independent variables, x, y and z, in Cartesian coordinate system or r, θ and ϕ in spherical polar coordinate system. Hence, the degree of freedom of the particle will increase from one to three and the time independent Schrödinger equation will be a three-dimensional differential equation.

In general, the potential in which a particle moves in a three-dimensional space is a function of all the three coordinates. However, in the present unit we shall confine ourselves to those potentials which depend only upon the radial coordinate r and are independent of the polar coordinates θ and ϕ. Such potentials are known as **spherically symmetric potentials** and the corresponding systems are called **spherically symmetric systems**.

When quantum mechanics was developed in the 1920s, one of its first (and also one of the most important) applications was to the understanding of hydrogen and hydrogen-like atoms (atoms with one valence electrons). In this unit, our main focus will be on the hydrogen atom. As you know, a hydrogen atom consists of a proton and an electron moving in the Coulomb potential of the proton. The motion of an electron in the Coulomb potential of the nucleus is also referred to as the Kepler problem of quantum mechanics — it is exactly solvable. You know that the Coulomb potential of the proton at a distance r is $-\frac{-e^2}{4\pi\varepsilon_0 r}$. Hence the potential is spherically symmetric.

In this unit we shall solve the three-dimensional Schrödinger equation for the hydrogen atom to obtain eigenfunctions and eigenenergies for the stationary states of hydrogen atom. In the course of solving this problem, when we find that the mathematics is becoming too difficult, we shall restrict ourselves to a qualitative discussion of the problem. We shall consider bound as well as continuum states of the hydrogen atom. To further simplify our study we shall begin by considering the motion of a particle in a spherically symmetric potential and then extend these ideas to the hydrogen atom. In the next unit we shall apply the results of this unit to hydrogen-like and other multielectron atoms and understand their optical spectra.

Objectives

After studying this unit you should be able to

- separate the time independent Schrödinger equation for a spherically symmetric system into its radial and angular parts,

- show that the angular momentum is a constant of motion for such systems,
- explain the concept of space quantization,
- reduce the two-body hydrogen atom system to two one-body systems,
- obtain the eigenfunctions and energy eigenvalues for the stationary state of a hydrogen atom,
- explain the spectra of the hydrogen atom,
- specify the constants of motion and the corresponding quantum numbers for the hydrogen atom problem.

9.2 THREE DIMENSIONAL SCHRÖDINGER EQUATION FOR A CENTRAL POTENTIAL

Let us consider the three-dimensional motion of a particle of mass μ in a spherically symmetric potential. For its stationary states, the time independent Schrödinger equation is given by

$$\left[-\frac{\hbar^2}{2\mu}\nabla^2 + V(r)\right]\psi(\mathbf{r}) = E\psi(\mathbf{r}) \tag{9.1}$$

where E is the total energy of the particle and $V(r)$ is its potential energy. Note that $V(r)$ is independent of the polar angles θ and ϕ. The force $\mathbf{F}$ acting on such a particle will be directed along $\mathbf{r}$. So classically, the torque $\boldsymbol{\tau}$ on the particle is equal to $\mathbf{r} \times \mathbf{F}$. Since $\mathbf{F}$ and $\mathbf{r}$ are in the same direction, $|\boldsymbol{\tau}|$ is equal to zero. Furthermore, torque is equal to the rate of change of angular momentum $\mathbf{L}$. Hence, the angular momentum of a particle moving under a spherically symmetric potential (also known as **central potential**) will not change with time. Thus, it will be a constant of motion for that object. Recall that you have solved this problem in Unit 6 of the elective PHE-01 in the classical domain to obtain Keplerian orbits for planetary motion around the sun.

However, a constant $\mathbf{L}$ means that all its three components L_x, L_y, and L_z are constant simultaneously. This is not possible in quantum mechanics because the three components of $\mathbf{L}$ do not commute among themselves (see Unit 7). Thus, there is a difference between the classical and quantum introductions to angular momentum. Instead of relating angular momentum to torque (as in classical mechanics), in quantum mechanics we find that the Hamiltonian can be written in such a way that it depends only on the angular momentum. This is how angular momentum makes its entry into the scheme of quantum mechanics.

Let us see how it is done. You know from Unit 3 of PHE-04, that in spherical polar coordinates, ∇^2 is given by

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}$$

Putting this expression for ∇^2 in Eq. (9.1) we obtain

$$\left[\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{2\mu}{\hbar^2}(E - V(r))r^2\right]\psi = -\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]\psi \tag{9.2}$$

Eq. (9.2) suggests that (r, θ, ϕ) can be separated in the variables r, θ and ϕ as follows:

$$\psi(r, \theta, \phi) = R(r)\,Y(\theta, \phi) \tag{9.3}$$

Putting Eq. (9.3) in Eq. (9.2) and using the method of separation of variables (refer to Unit 6 of PHE-05 entitled Mathematical Methods in Physics-II), we get the following two equations

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial R(r)}{\partial r}\right)+\left\{\frac{2\mu}{\hbar^2}(E-V(r))-\frac{K}{r^2}\right\}R(r)=0 \tag{9.4}$$

and

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y(\theta,\phi)}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2 Y(\theta,\phi)}{\partial\phi^2}=-KY(\theta,\phi) \tag{9.5}$$

where K is a constant. We can further show that the operators L^2 and L_z in spherical polar coordinates are given by

You should quickly establish Eqs. (9.4) and (9.5) before studying further.

$$L^2=-\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right] \tag{9.6a}$$

$$L_z=\frac{h}{2\pi i}\frac{\partial}{\partial\phi} \tag{9.6b}$$

In fact, you may like to do this exercise yourself. Try the following SAQ.

Spend 15 min

SAQ 1

(a) Prove Eqs. (9.6a) and (9.6b).

(b) Show that L^2 and L_z commute with the Hamiltonian.

Note: We will make use of this result in Sec. 9.2.1.

Using Eq. (9.6a), we can write Eq. (9.5) as:

$$L^2Y(\theta,\phi)=K\hbar^2Y(\theta,\phi) \tag{9.7}$$

Thus $Y(\theta,\phi)$ is an eigenfunction of the operator L^2 with $K\hbar^2$ as the eigenvalue. Let us study the functions $Y(\theta,\phi)$ in somewhat greater detail.

9.2.1 Eigenfunctions and Eigenvalues of the Angular Momentum Operator

Let us obtain the eigenvalues of Eq. (9.7) and determine the form of $Y(\theta,\phi)$. From Eq. (9.7) we can readily see that $Y(\theta,\phi)$ is an eigenfunction of the operator L^2 with the eigenvalue $K\hbar^2$. Further you have established in SAQ 1 that L^2 commutes with the Hamiltonian of the particle:

$$[L^2,H]=0 \tag{9.8a}$$

Now recall Eq. (7.32b) of Unit 7. If $[D,H]=0$ in that equation, $\frac{d\langle D\rangle}{dt}=0$, i. e., $\langle D\rangle$ is constant. Applying this result to the square of angular momentum (L^2) we obtain

$$\langle L^2\rangle=\text{constant} \tag{9.8b}$$

i.e., the square of the angular momentum is a constant of motion for a central potenti .

We can solve for $Y(\theta,\phi)$ by separating the variables θ and ϕ and writing

$$Y(\theta,\phi)=P(\vartheta)\,\Phi(\phi). \tag{9 a}$$

Substituting Eq. (9.9a) and Eq. (9.6a) in Eq. (9.7), and using separation of v.. iables, we obtain

$$\left[-\frac{\sin\theta}{P(\theta)}\frac{d}{d\theta}\left(\sin\theta\frac{dP(\theta)}{d\theta}\right)+K\sin^2\theta\right]=-\frac{1}{\Phi}\frac{d^2\Phi}{d\phi^2} \qquad (9.9b)$$

Eq. (9.9b) shows that both sides are equal to the same constant, say m_l^2. Hence, we can write

$$\frac{1}{\Phi}\frac{d^2\Phi}{d\phi^2}=-m_l^2 \qquad (9.10)$$

and its solution is

$$\Phi(\phi)=e^{im_l\phi} \qquad (9.11)$$

Now $\Phi(\phi)$ has to be single valued. Hence we must have

$$e^{im_l\phi}=e^{im_l(\phi+2\pi)} \qquad (9.12)$$

because angles $\phi = 0$ and $\phi = 2\pi$ are actually the same. Therefore m_l must be an integer For the operator l_z (Eq. 9.6b) operating upon $e^{im_l\phi}$, we obtain

$$L_z\,e^{im_l\phi}=m_l\hbar\,e^{im_l\phi} \qquad (9.13)$$

Thus, $e^{im_l\phi}$ is an eigenfunction of the operator L_z with the eigenvalue $m_l\hbar$.

The differential equation for $P(\theta)$ is given by

$$\sin\theta\frac{d}{d\theta}\left(\sin\theta\frac{dP(\theta)}{d\theta}\right)+K\,P(\theta)\sin^2\theta=m_l^2P(\theta) \qquad (9.14)$$

Eq. (9.14) can be solved analytically. However, the procedure is lengthy and out of place here. Hence here we only quote the results and discuss them qualitatively. If we take the constant K equal to $l\,(l+1)$ then it is found that quantum mechanically acceptable solutions of Eq.(9.14) are obtained only if the constant l is equal to one of the integers

$$l=|m_l|,\ |m_l|+1,\ |m_l|+2,\ \ldots \qquad (9.15)$$

Alternatively, we can also say that for a given integer l there will be following $(2l+1)$ values of m_l

$$-l,\ -l+1,\ -l+2,\ \ldots,\ 0,\ \ldots,\ l-1,\ l \qquad (9.16)$$

Then the acceptable solutions are given by

$$P_l^{|m_l|}(\theta)=\sin^{|m_l|}(\theta)\,F_{l,|m_l|}(\cos\theta) \qquad (9.17)$$

where $F_{l\,|m_l|}(\cos\theta)$ are polynomials in $\cos\theta$ and $P_l^{|m_l|}(\theta)$ are known as the *associated Legendre polynomials*. Thus, substituting Eqs (9.11) and (9.17) in Eq. (9.9a) we obtain the eigenfunctions of L^2:

$$Y_{l\,m_l}(\theta,\phi)=P_l^{|m_l|}(\theta)\,e^{im_l\phi} \qquad (9.18)$$

The functions given by Eq. (9.18) are known as **spherical harmonics.** Putting $K = l(l+1)$ in Eq.(9.7) you can readily see that they are the eigenfunctions of the operator L^2 with eigenvalues $l(l+1)\,\hbar^2$.

$$L^2Y_{l,m_l}(\theta,\phi)=l(l+1)\hbar^2\,Y_{l,m_l}(\theta,\phi) \qquad (9.19)$$

Thus $(2l+1)$ eigenfunctions $Y_{l\,m_l}(\theta,\phi)$ corresponding to the same l but different m_l (ranging from $-l$ to l) have the same eigenvalues. Hence we can say that these eigenfunctions are $(2l+1)$-fold degenerate. These functions specify the angular part of the steady state eigenfunctions of all particles moving under spherically symmetric potentials. They form an orthonormal set (as explained in Sec. 7.3 of Unit 7) and any function of θ and ϕ can be expressed as a linear combination of $Y_{l\,m_l}(\theta,\phi)$. We give here the explicit forms of some of the lower order $Y_{l\,m_l}(\theta,\phi)$ for ready reference:

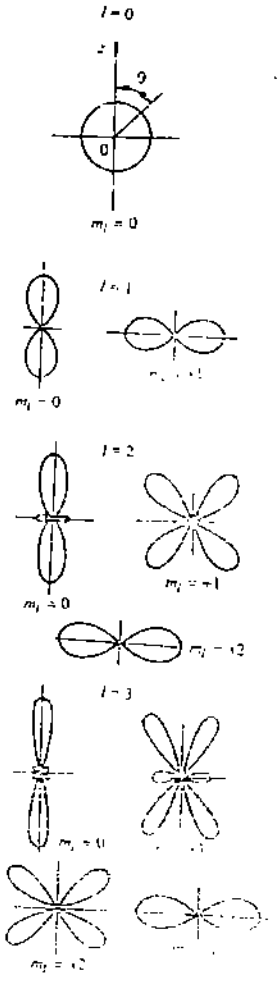


Fig. 9.1: Polar plots of $|Y_{l,m_l}(\theta,\phi)|^2$, for some values of l and m_l.

$$Y_{00} = \left(\frac{1}{4\pi}\right)^{1/2},$$

$$Y_{10} = \left(\frac{3}{4\pi}\right)^{1/2} \cos\theta$$

$$Y_{1,\pm1} = \mp\left(\frac{3}{8\pi}\right)^{1/2} \sin\theta\ e^{\pm i\phi}$$

$$Y_{20} = \left(\frac{5}{16\pi}\right)^{1/2} (3\cos^2\theta - 1) \qquad (9.20)$$

$$Y_{2,\pm1} = \mp\left(\frac{15}{8\pi}\right)^{1/2} \sin\theta\cos\theta\ e^{\pm i\phi}$$

and $$Y_{2,\pm2} = \left(\frac{15}{32\pi}\right)^{1/2} \sin^2\theta\ e^{\pm i2\phi}$$

The squares of some of these functions (which also represent the angular part of the wave function of the hydrogen atom) are shown in Fig. 9.1. You may now like to perform an exercise based on the ideas discussed so far.

Spend 10 min

SAQ 2

(a) Show that $Y_{l,m_l}(\theta, \phi)$ is an eigenfunction of L_z. Determine its eigenvalues.

(b) Show that $Y_{2,2}$ is normalised and is orthogonal to $Y_{2,-2}$.

Before studying further, it would do us well to examine the parity of the spherical harmonics. For this, we reflect $Y_{l,m_l}(\theta, \phi)$ about the origin. In such a reflection θ changes to $\pi - \theta$ and ϕ changes to $\pi + \phi$.

Now

$$e^{im_l(\pi+\phi)} = (-1)^{m_l} e^{im_l\phi} \qquad (9.21a)$$

Further $\sin(\pi - \theta) = \sin\theta$ and $\cos(\pi - \theta) = -\cos\theta$. Hence, it can be shown that

$$P_l^{|m_l|}(\pi-\theta) = (-1)^{l-|m_l|}\ P_l^{|m_l|}(\phi) \qquad (9.21b)$$

Therefore, the parity of $Y_{l,m_l}(\theta, \phi)$ is given by $(-1)^{l-|m_l|+m_l}$ which is to equal $(-1)^l$. We shall make use of these concepts in the later sections of the Unit.

Spend 2 min

SAQ 3

Use the expression given by Eq. (9.20) for $Y_{2,1}(\theta, \phi)$ and verify that it is of even parity.

To sum up, so far we have obtained the solutions of the angular part of the stationary states of a particle moving in a spherically symmetric potential. These are nothing but the eigenfunctions of the angular momentum operator L^2. Their exact functional dependence on θ and ϕ is given by the spherical harmonics. The eigenvalues of the operator L^2 are $l(l+1)\hbar^2$ where l takes discrete integral values given by Eq. (9.15).

Let us now try to understand what these results mean physically, in terms of what is called space **quantization**.

9.2.2 Space Quantization

You have shown in SAQ 2(a) that the spherical harmonics $Y_{l,m_l}(\theta, \phi)$ are eigenfunctions of L_z with eigenvalues $m_l\hbar$. Thus we can determine exact values of L^2 and L_z simultaneously. However, since L_x, L_y, L_z cannot be determined simultaneously

according to the uncertainty principle, L_x and L_y will be uncertain. Thus, we are confronted with some surprises about the quantum mechanical angular momentum as compared with its classical counterpart.

Classically, for the same magnitude of the angular momentum, we can obtain an infinite number of states by changing the direction of the angular momentum vector. But quantum mechanically, for each value of angular momentum, there are only a finite number of states characterised by l and m_l. Moreover, in quantum mechanics, the components of **L** in two of the three directions being uncertain, we do not describe a state by specifying the directions of the angular momentum vector. Instead, we give the component of the angular momentum along a specific direction. We conveniently choose this direction to be along the z-axis. So how do we visualise this situation?

There is a useful pictorial way to communicate these quantum mechanical results — this is the so called **vector model** for the angular momentum. In this model we represent the angular momentum of the particle in motion by a vector **L** of length $[l(l+1)]^{1/2}$. The angular momentum vector precesses around the z-axis in such a way that the magnitude of **L** (hence L^2) and L_z (projection of **L** on z-axis) are constants (see Fig. 9.2a).

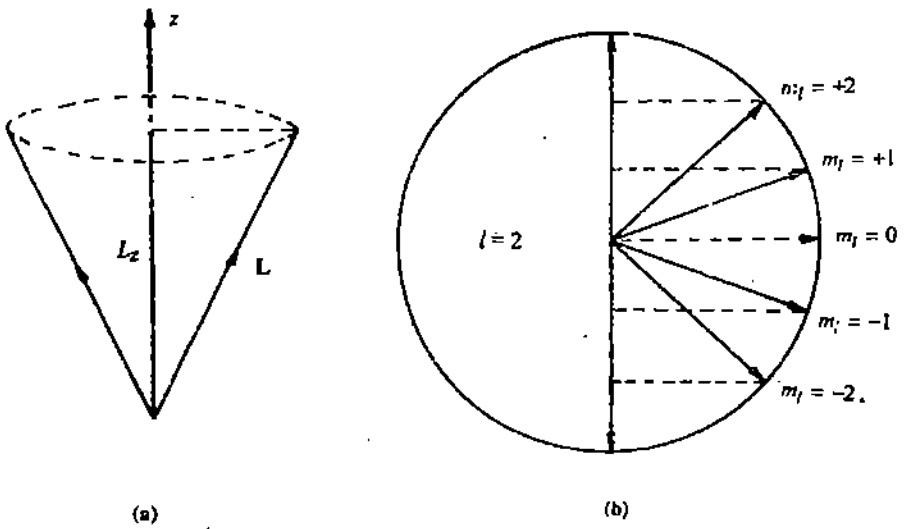


Fig. 9.2: (a) Precession of L about z-axis; (b) space quantization for $l = 2$. The radius of the circle is $[2(2+1)]^{1/2}$. The multiplicity of states is 5.

Since for a given value of l, the eigenvalues of L_z are $m_l\hbar$ with integer values of m_l (ranging from $-l$ to $+l$), the component of **L** along the z-axis is quantized. A measurement of L_z will yield only the $2l + 1$ quantized values, with a maximum value l less than the magnitude of the vector **L** for $l \neq 0$. Further, the vector **L** can make only certain quantized angles with the z-axis; the angle θ, between **L** and L_z can take only discrete values given by

$$\cos\theta = \frac{m_l}{[l(l+1)]^{1/2}} \tag{9.22}$$

This phenomenon of the quantization of the direction of **L** with respect to one of the coordinate axes, is known as **space quantization**. Since $|m_l|$ is always less than $l(l+1)$ (except for $l = 0$), the vector **L** can never be along z-axis. For $l = 2$, the values of m_l are 2, 1, 0, −1 and −2 as shown in Fig. 9.2b. Furthermore, although L_x and L_y are uncertain, $L_x^2 + L_y^2$ being equal to $L^2 - L_z^2$ have definite non-zero values unless $l = 0$; however, the values of L_x and L_y are not quantized. Thus we can visualise the angular momentum vector sweeping out in all possible directions in the xy plane.

Having analysed the angular part of the wavefunction and some of its implications, let us now consider the radial part of the eigenfunctions for a spherically symmetric potential.

9.2.3 Radial Eigenfunctions

Putting $K = l(l+1)$ in Eq. (9.4), we obtain the differential equation for the radial function $R(r)$. This is given by

$$-\frac{\hbar^2}{2\mu}\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR(r)}{dr}\right)+\left[V(r)+\frac{\hbar^2 l(l+1)}{2\mu r^2}\right]R(r)=ER(r) \qquad (9.23a)$$

This is a one-dimensional eigenvalue equation for the radial eigenfunction $R(r)$. The actual solution depends upon the form of potential energy function $V(r)$. However, the **effective potential energy** of the particle is

$$V_{eff}(r)=V(r)+\frac{\hbar^2 l(l+1)}{2\mu r^2} \qquad (9.23b)$$

Thus, there is an extra term in the form of repulsive potential energy ($l(l+1)\hbar^2/2\mu r^2$) which increases with l (see Fig. 9.3). You can see that this term decreases the probability of finding the particle near the centre of force. This term is also known as the **centrifugal potential energy**, or the **centrifugal barrier.**

The origin of the centrifugal term can be understood in the following manner using classical correspondence. For a particle of mass μ moving in a circular orbit of radius r, classically, there is a centrifugal force directed radially outward. The magnitude of the force is $\mu v^2/r = L^2/\mu r^3$, where $L = \mu vr$ for a circular orbit. The potential corresponding to such a force is $L^2/2\mu r^2$ (since $F = \partial V/\partial r$). In quantum mechanics we must replace L^2 by its eigenvalue $l(l+1)\hbar^2$, hence we obtain the quantum mechanical expression for the centrifugal potential.

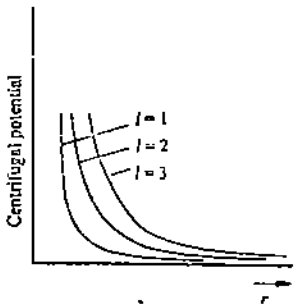


Fig. 9.3: The centrifugal barrier for some values of l.

For bound particles (such as a simple harmonic oscillator), the values of E (eigenvalues) are discrete. Otherwise E varies in a continuous manner. But whatever be the form of $V(r)$, as long as it is spherically symmetric, the angular part of the eigenfunction of the particle is given by the spherical harmonics $Y_{l,m_l}(\theta, \phi)$.

In the next section we shall take $V(r)$ to be the Coulomb potential energy, appropriate to a hydrogen atom and shall obtain eigefunctions and eigenvalues of a hydrogen atom.

9.3 THE HYDROGEN ATOM

Let us consider the hydrogen atom as an example of a three-dimensional quantum mechanical system. As you know, a hydrogen atom consists of a proton and an electron. Thus it is a two-particle system. The Hamiltonian for two-body motion in a central force field is given as

$$H=\frac{p_1^2}{2M}+\frac{p_2^2}{2m}+V(\mathbf{r}_1,\mathbf{r}_2) \qquad (9.24a)$$

Thus the stationary states of the hydrogen atom are the solutions of the following time independent Schrödinger equation

$$\left[-\frac{\hbar^2}{2M}\nabla_1^2-\frac{\hbar^2}{2m}\nabla_2^2-\frac{e^2}{4\pi\varepsilon_0|\mathbf{r}_1-\mathbf{r}_2|}\right]\psi(\mathbf{r}_1,\mathbf{r}_2)=E_T\psi(\mathbf{r}_1,\mathbf{r}_2) \qquad (9.24b)$$

where M and m are the masses of the proton and the electron, respectively; $\mathbf{r}_1$ and $\mathbf{r}_2$ are the coordinates of the proton and the electron, respectively, with respect to an origin O. E_T is the total energy of the system and ε_0 is the permittivity constant. Thus we are required to solve a six-dimensional differential equation to obtain the eigenfunction $\psi(\mathbf{r}_1, \mathbf{r}_2)$ and eigenvalue E_T. However, we can reduce the above equation into two three-dimensional equations in the following manner.

Let $\mathbf{R}$ be the coordinate of the centre of mass of the atom. Then

$$\mathbf{R}=\frac{M\mathbf{r}_1+m\mathbf{r}_2}{M+m} \qquad (9.25)$$

The separation between the proton and the electron is given by

$$\mathbf{r}=\mathbf{r}_1-\mathbf{r}_2 \qquad (9.26)$$

Solving for $\mathbf{r}_1$ and $\mathbf{r}_2$ in terms of $\mathbf{R}$ and $\mathbf{r}$ we get

$$\mathbf{r}_1 = \mathbf{R} + \frac{m}{M+m}\ \mathbf{r} \tag{9.27}$$

and

$$\mathbf{r}_2 = \mathbf{R} - \frac{M}{M+m}\ \mathbf{r} \tag{9.28}$$

Now you know that

$$\frac{\partial \psi}{\partial x_1} = \frac{\partial \psi}{\partial X}\frac{\partial X}{\partial x_1} + \frac{\partial \psi}{\partial x}\frac{\partial x}{\partial x_1}$$

where x_1, X and x are the x components of $\mathbf{r}_1$, $\mathbf{R}$ and $\mathbf{r}$, respectively. Hence Eqs. (9.25) to (9.27) yield for the x component

$$\frac{\partial \psi}{\partial x_1} = \frac{M}{m+M}\frac{\partial \psi}{\partial X} + \frac{\partial \psi}{\partial x} \tag{9.29}$$

Hence, in three-dimensions

$$\nabla_1 = \frac{M}{m+M}\nabla_{\mathbf{R}} + \nabla \tag{9.30}$$

where

$$\nabla = \hat{\mathbf{i}}\frac{\partial}{\partial x} + \hat{\mathbf{j}}\frac{\partial}{\partial y} + \hat{\mathbf{k}}\frac{\partial}{\partial z} \tag{9.31}$$

and (x_1, y_1, z_1), (x, y, z) and (X, Y, Z) are the components of $\mathbf{r}_1$, $\mathbf{r}$ and $\mathbf{R}$, respectively. Similarly, from Eq. (9.28) we obtain

$$\nabla_2 = \frac{m}{m+M}\nabla_{\mathbf{R}} - \nabla \tag{9.32}$$

Eqs.(9.30) and (9.32) yield

$$\nabla_1^2 = \left(\frac{M}{m+M}\right)^2 \nabla_{\mathbf{R}}^2 + 2\left(\frac{M}{m+M}\right)\nabla_{\mathbf{R}}.\nabla + \nabla^2 \tag{9.33a}$$

and

$$\nabla_2^2 = \left(\frac{m}{m+M}\right)^2 \nabla_{\mathbf{R}}^2 - 2\left(\frac{m}{m+M}\right)\nabla_{\mathbf{R}}.\nabla + \nabla^2 \tag{9.33b}$$

Putting the expressions for ∇_1^2 and ∇_2^2 from Eqs. (9.33a and b) into Eq. (9.24b) we get

$$\left[-\frac{\hbar^2}{2(M+m)}\nabla_{\mathbf{R}}^2 - \frac{\hbar^2}{2}\left(\frac{1}{m}+\frac{1}{M}\right)\nabla^2 - \frac{e^2}{r}\right]\psi(\mathbf{R}, \mathbf{r}) = E_T\psi(\mathbf{R}, \mathbf{r}) \tag{9.34}$$

where for our convenience we have replaced $e^2/4\pi\varepsilon_0$ by only e^2 with

$$e^2 = 2.31 \times 10^{-28}\ \text{Joule.metre.}$$

Eq. (9.34) is separable in the coordinates $\mathbf{R}$ and $\mathbf{r}$. Taking

$$\psi(\mathbf{R}, \mathbf{r}) = \phi(\mathbf{R})\ \psi(\mathbf{r}) \tag{9.35}$$

we find that $\phi(\mathbf{R})$ and $\psi(\mathbf{r})$ are the respective solutions of the following three-dimensional differential equations

$$-\frac{\hbar^2}{2(M+m)}\nabla_{\mathbf{R}}^2\,\phi(\mathbf{R}) = E_H\phi(\mathbf{R}) \tag{9.36}$$

and

$$-\frac{\hbar^2}{2\mu}\nabla^2\,\psi(\mathbf{r}) - \frac{e^2}{r}\psi(\mathbf{r}) = E\,\psi(\mathbf{r}) \tag{9.37a}$$

where

$$\mu = \left(\frac{1}{m}+\frac{1}{M}\right)^{-1} \tag{9.37b}$$

and

$$E_T = E + E_H \tag{9.37c}$$

As you know, μ is the *reduced mass* of the system.

Eq. (9.36) shows that a particle of mass $(m + M)$, which is the total mass of the hydrogen atom, is moving freely in a three-dimensional space and its total energy is E_H (with zero potential energy). This is a problem that you have already solved in Sec. 8.2. Its eigenfunctions are given by plane waves

$$\phi(\mathbf{R}) = e^{i\mathbf{K}.\mathbf{R}} \tag{9.38}$$

with $$\frac{\hbar^2 K^2}{2(m+M)} = E_H \tag{9.39}$$

The eigenvalue E_H and the corresponding quantum number K vary in a continuous manner.

On the other hand, Eq. (9.37a) describes the motion of a particle of mass μ having potential energy $-e^2/r$ with respect to a fixed centre. Thus by the above procedure we have reduced a two-body system into two one-body systems, one of mass $(m+M)$ which moves freely in space and other of mass μ and charge e which moves under an attractive potential $-e^2/r$. You should note that in the present model the relative motion of electron and proton with respect to each other has been replaced by the motion of a particle of mass μ with respect to a fixed centre of force.

Now Eq. (9.37a) is exactly the same as Eq. (9.1) with $V(r) = -e^2/r$. Hence the eigenfunctions of the particle of mass μ, which are also known as the eigenfunctions of the hydrogen atom, are given by

$$\psi(r, \theta, \phi) = R(r)\, Y_{l, m_l}(\theta, \phi) \tag{9.40}$$

The radial function $R(r)$ is the solution of the following one-dimensional differential equation

$$-\frac{\hbar^2}{2\mu}\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR(r)}{dr}\right) + \left\{-\frac{e^2}{r} + \frac{\hbar^2}{2\mu}\frac{l(l+1)}{r^2}\right\} R(r) = E\,R(r). \tag{9.41}$$

The above equation has been obtained from Eq. (9.23a) by taking $V(r) = -e^2/r$. The effective potential energy in this case is

$$V_{eff}(r) = -\frac{e^2}{r} + \frac{\hbar^2 l(l+1)}{2\mu r^2}$$

It is shown in Fig. 9.4.

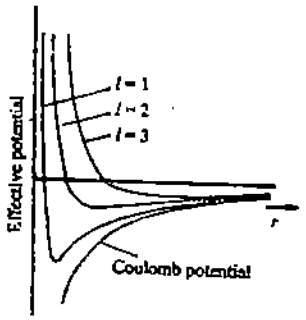


Fig 9.4: The Coulomb potential and the effective potential for some values of l. Compare this with Fig. 9.3.

You can see from Fig. 9.4 that near the origin $\frac{\hbar^2 l(l+1)}{2\mu r^2}$ is much larger than $-e^2/r$. We will now present some results here without going into the detailed mathematical solution of Eq. (9.41). We find that for finite values of E, the solution of Eq. (9.41) near the origin is given by $R(r) = cr^l$, where c is a constant. Further, at large values of r, $V(r)$ tends to zero and the differential equation reduces to

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR(r)}{dr}\right) = -\frac{2\mu E}{\hbar^2}R(r) \tag{9.42}$$

Let us take the zero (or the reference level) of the energy E to be the energy of that state where the hydrogen atom is ionised but the free electron of the hydrogen atom has zero kinetic energy. Then the bound eigenstates of the hydrogen atom have negative total energy E (the positive kinetic energy of the particle of mass μ being less than the magnitude of negative potential energy). On the other hand, the free eigenstates of the hydrogen atom have positive energy E. For the bound states $E < 0$ and we put $E = -|E|$ so that, the solution of Eq. (9.42) is given by

$$R(r) = c_2 \exp\left[-\left(-\frac{2\mu E}{\hbar^2}\right)^{1/2} r\right] \tag{9.43}$$

E being negative, the radial wave function decreases exponentially at large values of r. At the intermediate values of r, the solution for $R(r)$ can be obtained by employing power series method. However, here we only state the final result. As you have noticed

for the bound states of other systems (say simple harmonic oscillator), the acceptable solutions of $R(r)$ exist only for discrete values of E and these discrete values are given by

$$E_n = -\frac{\mu}{2\hbar^2}\frac{e^4}{n^2} \tag{9.44}$$

where n is a positive integer and for a given value of l ranges from

$$n = l+1,\ l+2,\ l+3,\ \ldots \tag{9.45}$$

Thus, we find that the radial eigenfunctions depend upon n as well as l and are given by

$$R_{n\,l}(r) = N_{n\,l}\exp\left(-\frac{r}{na_0}\right)\left(\frac{r}{a_0}\right)^l G_{n\,l}\,(r/a_0) \tag{9.46}$$

where $G_{n\,l}(r/a_0)$ are the *associated Laguerre polynomials* and $N_{n\,l}$ is the normalisation constant. The parameter a_0 is given by

$$a_0 = \frac{\hbar^2}{\mu e^2} \tag{9.47}$$

It is interesting to note that a_0 is equal to the radius of the first orbit of the electron in a hydrogen atom for the model proposed by Bohr, provided μ is replaced by the rest mass of the electron m. Since the ratio of m/μ is very close to unity (1.0005), we shall take a_0 to be equal to the first Bohr radius with 0.529×10^{-10} metre as its value. Under the same approximation the eigenenergy E_n is given by

$$E_n = -\frac{R}{n^2} \tag{9.48}$$

where

$$R = \frac{me^4}{2\hbar^2}$$

As you know, R is the Rydberg constant.

You may now like to take a break. Attempt a simple exercise.

SAQ 4

Spend 5 min

Obtain the value of the Rydberg constant in the units of electron volts and m^{-1}.

A few of the lower radial eigenfunctions of the hydrogen atom are given by

$$R_{10}(r) = \frac{2}{a_0^{3/2}}\,e^{-r/a_0} \tag{9.49a}$$

$$R_{20}(r) = \frac{1}{(2a_0)^{3/2}}\left(2-\frac{r}{a_0}\right)e^{-r/2a_0} \tag{9.49b}$$

$$R_{21}(r) = \frac{1}{(2a_0)^{3/2}}\,\frac{r}{a_0\sqrt{3}}\,e^{-r/2a_0} \tag{9.49c}$$

$$R_{30}(r) = \left(\frac{1}{3a_0}\right)^{3/2} 2\left[1-\frac{2r}{3a_0}+\frac{2r^2}{27a_0^2}\right]e^{-r/3a_0} \tag{9.49d}$$

$$R_{31}(r) = \left(\frac{1}{3a_0}\right)^{3/2}\frac{4\sqrt{2}}{3}\,\frac{r}{a_0}\left(1-\frac{r}{6a_0}\right)e^{-r/3a_0} \tag{9.49e}$$

$$R_{32}(r) = \left(\frac{1}{3a_0}\right)^{3/2}\frac{2\sqrt{2}}{27\sqrt{5}}\,\frac{r^2}{a_0^2}\,e^{-r/3a_0} \tag{9.49f}$$

SAQ 5

Spend 10 min

Show that $R_{20}(r)$ is normalised and is orthogonal to $R_{10}(r)$.

Finally, the eigenfunctions of a hydrogen atom are given by

$$\psi_{nlm_l}(r, \theta, \phi) = R_{nl}(r) Y_{l, m_l}(\theta, \phi) \tag{9.50}$$

where $R_{n\,l}(r)$ and $Y_{l, m_l}(\theta, \phi)$ are given by Eq. (9.46) and Eq. (9.18), respectively. These eigenfunctions form an orthonormal set, i.e.,

$$\int \psi^*_{n\,l\,m_l}(\mathbf{r})\, \psi_{n'l'm_l'}(\mathbf{r})\, r^2\, dr \sin\theta\, d\theta\, d\phi = \delta_{nn'}\, \delta_{ll'}\, \delta_{m_l m_l'} \tag{9.51}$$

where $\delta_{jj'} = 1$ for $j = j'$ and zero otherwise.

The radial part of some eigenfunctions are shown in Fig. 9.5.

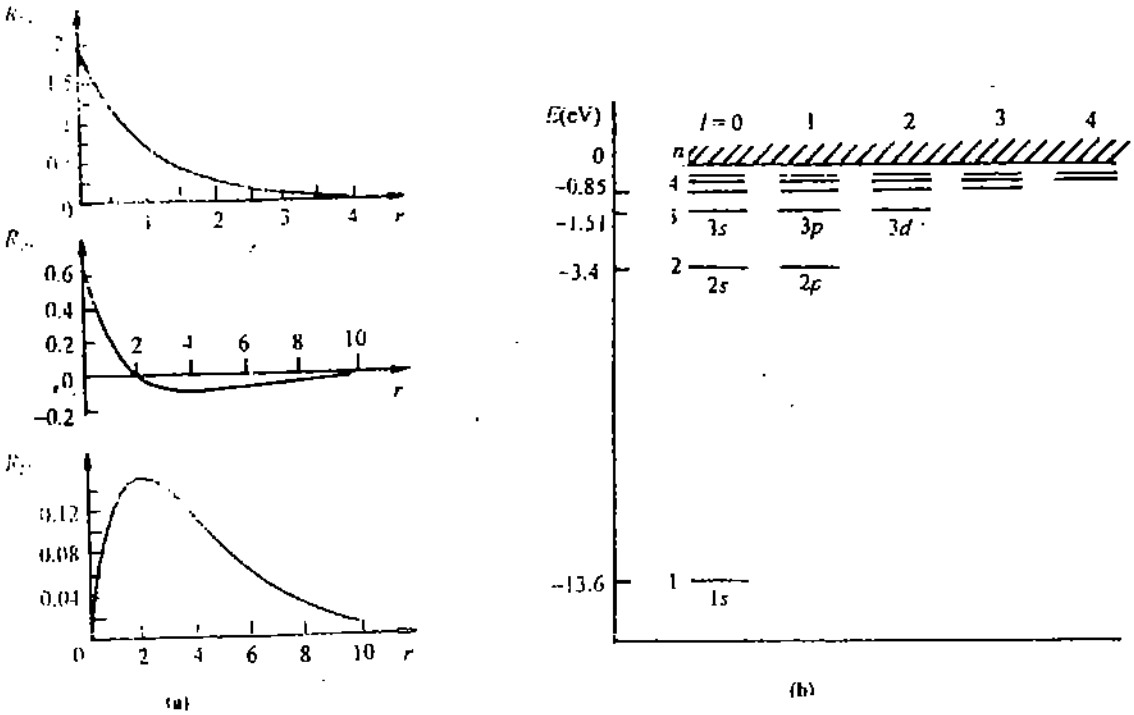


Fig. 9.5: (a) A few radial eigenfunctions; (b) eigenenergies of bound and continuum states of a hydrogen atom.

We can now make some observations about the Coulomb potential energy problem in relation to the hydrogen atom. Eq. (9.46) tells us that the bound state eigenfunctions ($E < 0$) for the Coulomb potential go to zero as r goes to infinity. Notice that this potential gives an infinite number of bound states starting at energy $-\mu e^4/2\hbar^2$ and ending at 0. The eigenenergy given by Eq. (9.44) varies in a discrete manner. The difference between the energies of the two consecutive energy states decreases as n increases (see Fig. 9.5b). For large n, the energy difference becomes quite small. The states with high n are called *Rydberg states*. Finally at $n = \infty$, the eigenenergy becomes zero and the hydrogen atom is ionised into a proton and an electron with zero total energy. The eigen states with $E > 0$ are *continuum states*. They are shown by the shaded portion of Fig. 9.5b. The eigenfunctions of such states do not go to zero as r goes to infinity and E varies in a continuous manner. The eigenfunctions of continuum states of a hydrogen atom are Coulomb waves.

The atomic electrons having $l = 0, 1, 2, 3 \ldots$ are known as $s, p, d, f, \ldots$ electrons, respectively. It is evident from Eqs. (9.46) and (9.50) that only for s electrons, for which $l = 0$, the eigenfunction ψ_{nlm_l} is finite at $r = 0$, which is practically the position of the nucleus. Hence only s electrons have a finite probability of their existence at the nucleus whereas for the electrons having non-zero angular momentum ($l > 0$) the probability is zero. Such a behaviour can be understood from Eq. (9.41). The centrifugal

potential energy $(\hbar^2/2\mu)\, l(l+1)/r^2$) for $l > 0$ does not allow p, d, f, ... electrons to come very close to the nucleus. $n = 1$ state is known as the ground state of the hydrogen atom while $n = \infty$ corresponds to its lowest ionised state. Thus we require one Rydberg energy to ionise a hydrogen atom.

There is another interesting feature worth commenting upon: we have a degeneracy in the spectrum, the l-degeneracy. The energy does not depend on l, but only on n; yet for a fixed n, possible l values are $l = 0, 1, 2, ..., n-1$. In addition to the l-degeneracy, there is also the m_l-degeneracy, the result of spherical symmetry. For each l, m_l goes from $-l$ to $+l$, giving us $2l + 1$ degenerate levels. For any n, the total degeneracy, then, is

$$\sum_{l=0}^{n-1} (2l + 1) = n^2$$

And if we take into account the two-valuedness called spin, which you will study in Unit 10, the total degeneracy is $2n^2$.

The discussion so far helps us to beautifully explain the spectra of the hydrogen atom.

9.3.1 Spectra of the Hydrogen Atom

When the electron in a hydrogen atom makes a transition from its excited states ($n > 1$) to a lower excited state or to the ground state ($n = 1$), it emits electromagnetic radiations of characteristic frequencies of the hydrogen atom. The energy difference between two eigenstates of the hydrogen atom is given by

$$\Delta E = R\left(\frac{1}{n_2^2} - \frac{1}{n_1^2}\right) \tag{9.52}$$

If we take R in the units of m^{-1} the wave number of emitted radiation is given by

$$\bar{\nu} = 1.097 \times 10^7 \left(\frac{1}{n_2^2} - \frac{1}{n_1^2}\right) \text{ m}^{-1} \tag{9.53}$$

For $n_2 = 1$ and $n_1 = 2, 3, 4, ...$ we get a series of electromagnetic radiation of different wavelengths. This series is known as **Lyman series** and lies in the ultraviolet region of the electromagnetic spectrum. **The Balmer series** correspond to the transitions from $n_1 = 3, 4, 5, ...$ to $n_2 = 2$. Similarly, **Paschen and Brackett series** are produced due to transitions from $n_1 = 4, 5, 6, ..$ to $n_2 = 3$ and $n_1 = 5, 6, 7, ..$ to $n_2 = 4$, respectively. Thus, the theoretically obtained eigenenergy spectrum successfully explains the observed line spectra of the hydrogen atom.

However, when we compare the energy spectrum with very accurate experimental data, we find some discrepancies. This is because in the real hydrogen atom there are other interactions that we have neglected here. You will study about these interactions in higher level courses. We will now discuss an important concept regarding such systems — the concept of quantum numbers that are constants of motion and characterise the state of such a system.

9.3.2 Quantum Numbers and Constants of Motion

In the study of the hydrogen atom you have come across three integers namely n, l and m_l. These integers are known as **quantum numbers.** Since n is connected with the eigenenergy of the system (see Eq. 9.48) it is known as energy **quantum number** or **principal quantum number.** Its existence is due to the fact that the energy is a constant of motion, i.e., the states are stationary states. The energy quantum number exists for continuum states also with the difference that now it varies in a continuous manner. In the continuum, it is usually denoted by k.

Due to spherically symmetric potential the angular momentum **L** of the object becomes a constant of motion. However, due to non-commutability of L_x, L_y and L_z the angular momentum vector **L** is not a constant of motion in quantum mechanics but as you have seen, L^2 is a constant of motion. This gives rise to another quantum number l, which varies in a discrete manner and is a positive integer. Since l is connected with the

orbital motion of the object it is called **orbital quantum number.** From Eq. (9.45) we can show that l is less than n and for a given value of n it takes the following values:

$$0, 1, 2, \ldots n-1 \tag{9.54}$$

Since a spherically symmetric potential is also axially symmetric, the z component of **L** also becomes a constant of motion and gives rise to a third quantum number m_l. It can take negative as well as positive integer values. The energy of a free hydrogen atom depends only upon the quantum number n. However, it can be shown that if the atom is placed under a magnetic field, its energy depends upon m_l. Hence m_l is known as **magnetic quantum number.** For a given value of l, the permissible values of m_l are

$$-l, -l+1, \ldots 0, 1, 2, \ldots l-1, l \tag{9.55}$$

The existence of three quantum numbers is also a consequence of the fact that the time independent Schrödinger equation contains three independent variables r, θ and ϕ. We have one quantum number for each space coordinate.

Spend 5 min

SAQ 6

Show that for $n = 3$ there are 9 degenerate eigenfunctions for a hydrogen atom.

Let us now summarise what you have studied in this unit.

9.4 SUMMARY

- In this unit we have discussed the quantum mechanical behaviour of a particle having *constant total energy* and moving under a **three-dimensional, spherically symmetric potential.**
- A **spherically symmetric potential** depends only upon the radial coordinate r and is independent of the polar coordinates θ and ϕ.
- In classical mechanics, the angular momentum **L** of such a particle is a constant of motion. However, in quantum mechanics all the three components L_x, L_y and L_z of the vector **L** cannot be constants of motion simultaneously due to the fact that these three components do not commute among themselves. However, the magnitude of **L** or L^2 and any one of the components of **L** (which we have taken as L_z) can be a constant of motion.
- The **three-dimensional time independent Schrödinger** equation of a particle moving under central potential can be separated into three one-dimensional differential equations, each one of them being a function of only one coordinate, r, θ or ϕ. The solution of the differential equation in ϕ is $e^{im_l\phi}$ whereas the solution of the differential equation in θ is the associated legendre polynomial $P_l^{|m_l|}(\theta)$. For quantum mechanically acceptable solutions l and m_l are to be only integers. Thus, l takes only positive integer values and for a given l there are $(2l+1)$ values of m_l given by

  $$m_l = -l, -l+1, -l+2, \ldots 0, 1, \ldots l-1, l$$

  The quantum numbers l and m_l are known as **orbital** and **magnetic quantum numbers**, respectively. The product of the two solutions in θ and ϕ is known as **spherical harmonics** and it is represented by $Y_{l,m_l}(\theta, \phi)$.
- The vector **L** is never stationary in space but precesses around the z-axis. The angle between the z-axis and **L** can take only discrete values given by $\cos^{-1}(m_l/\sqrt{l(l+1)})$. This quantization of the orientation of **L** with respect to one coordinate axis is known as **space quantization.**

- The nature of the radial wave function $R(r)$ depends upon whether the state is a bound state or a continuum state. For bound states the eigenfunction varies as $\exp\left[-\left(\frac{2\mu}{\hbar^2}E\right)^{1/2}\frac{r}{a_0}\right]$ at large values of r. Hence the probability of finding the particle goes to zero as r increases to ∞. On the other hand, for continuum states the probability remains finite even as r goes to infinity. At small and intermediate values of r, different spherically symmetric potentials give rise to different radial functions. In this unit we have considered Coulomb potential, appropriate to a hydrogen atom.

- A hydrogen atom is a two-particle system consisting of a proton and an electron and its stationary state Schrödinger equation is a six-dimensional differential equation. However, it can be separated into two three-dimensional differential equations: One corresponding to the motion of a free particle having mass $(m + M)$. Its solutions are plane waves $e^{iK.R}$, where the energy quantum number K varies in a continuous manner and is related to E_H by Eq. (9.39). The second three-dimensional differential equation describes the motion of a particle of mass μ (reduced mass of the system) having a charge $-e$ in a spherically symmetric potential (Coulomb potential) due to a fixed centre of force having charge $+e$. The latter differential equation again separates into three one-dimensional differential equations, one for each polar coordinate r, θ and ϕ. These three degrees of freedom give rise to **three quantum numbers** n, l and m_l.

- The **angular eigenfunctions of the hydrogen atom** are again **spherical harmonics** and the **radial eigenfunctions** for bound states are given in terms of the **associated Laguerre polynomials.** The quantum number n takes only positive integer values excluding zero. For a given value of n, the quantum number l takes n values given by

$$l = 0,\ 1,\ \ldots\ n-1$$

For each value of l the magnetic quantum number m_l takes the values

$$m_l = -l, -l+1, \ldots 0, l, \ldots l-1, l$$

- The energy E corresponding to different eigenfunctions $\psi_{nlm_l}(r, \theta, \phi)$ depends only upon the quantum number n. Hence for a given n there are n^2 degenerate eigenfuncions corresponding to different permissible values of l and m_l. For states with $E > 0$, the radial eigenfunctions are Coulomb waves and the energy quantum number k varies in a continuous manner. Now energy states are no longer discrete and we talk in terms of energy states per unit energy range.

- The quantum mechanical treatment of the hydrogen atom explains the production of various electromagnetic series experimentally observed in the spectrum of the atom.

9.5 TERMINAL QUESTIONS

Spend 30 min

1. Use the expressions given by Eq. (9.20) to show that

$$\sum_{m_l=-2}^{2} |Y_{2,m_l}(\theta, \phi)|^2 = \frac{5}{4\pi}$$

You may note that in general

$$\sum_{m_l=-l}^{2} |Y_{l,m_l}(\theta, \phi)|^2 = \frac{(2l+1)}{4\pi}$$

2. According to the virial theorem the average value of the potential energy of a particle subjected to a Coulomb potential in any stationary bound state is two times its total energy. Verify the above theorem for the ground state of the hydrogen atom. Further show that the average kinetic energy is equal to the magnitude of total energy.

3. Use the uncertainty relation to show that the dimension of the most stable ground state hydrogen atom is of the order of the first Bohr radius a_0.

4. Obtain the most probable value and expectation value of r for the ground state of a hydrogen atom.

9.6 SOLUTIONS AND ANSWERS

Self-Assessment Questions

1. We know that

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} = r\,[\hat{\mathbf{e}}_r \times (-i\hbar\nabla)]$$

In spherical polar coordinates

$$\nabla = \hat{\mathbf{e}}_r \frac{\partial}{\partial r} + \hat{\mathbf{e}}_\theta \frac{1}{r} \frac{\partial}{\partial \theta} + \hat{\mathbf{e}}_\phi \frac{1}{r \sin\theta} \frac{\partial}{\partial \phi}$$

Therefore

$$\mathbf{L} = -i\hbar r \hat{\mathbf{e}}_r \times \left[\hat{\mathbf{e}}_\theta \frac{1}{r} \frac{\partial}{\partial \theta} + \hat{\mathbf{e}}_\phi \frac{1}{r \sin\theta} \frac{\partial}{\partial \phi}\right] \quad (\because \hat{\mathbf{e}}_r \times \hat{\mathbf{e}}_r = 0)$$

$$= -i\hbar r \left[\hat{\mathbf{e}}_\phi \frac{1}{r} \frac{\partial}{\partial \theta} - \hat{\mathbf{e}}_\theta \frac{1}{r \sin\theta} \frac{\partial}{\partial \phi}\right]$$

$$= -i\hbar \left[\hat{\mathbf{e}}_\phi \frac{\partial}{\partial \theta} - \hat{\mathbf{e}}_\theta \frac{1}{\sin\theta} \frac{\partial}{\partial \phi}\right]$$

Hence

$$L^2 = \mathbf{L} \cdot \mathbf{L} = -\hbar^2 \left(\hat{\mathbf{e}}_\phi \frac{\partial}{\partial \theta} - \hat{\mathbf{e}}_\theta \frac{1}{\sin\theta} \frac{\partial}{\partial \phi}\right) \cdot \left(\hat{\mathbf{e}}_\phi \frac{\partial}{\partial \theta} - \hat{\mathbf{e}}_\theta \frac{1}{\sin\theta} \frac{\partial}{\partial \phi}\right)$$

The derivatives of the unit vectors are given as

$$\frac{\partial}{\partial \theta}\,\hat{\mathbf{e}}_\theta = -\hat{\mathbf{e}}_r, \qquad \frac{\partial}{\partial \phi}\,\hat{\mathbf{e}}_\theta = \hat{\mathbf{e}}_\phi \cos\theta$$

$$\frac{\partial}{\partial \theta}\,\hat{\mathbf{e}}_\phi = 0, \qquad \frac{\partial}{\partial \phi}\,\hat{\mathbf{e}}_\phi = -(\hat{\mathbf{e}}_r \sin\theta + \hat{\mathbf{e}}_\theta \cos\theta)$$

Substituting in the equation for L^2 above, we obtain after some algebra

$$L^2 = -\hbar^2 \left[\frac{1}{\sin\theta} \frac{\partial}{\partial \theta} \left(\sin\theta \frac{\partial}{\partial \theta}\right) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial \phi^2}\right]$$

$$L_z = \hat{\mathbf{e}}_z \cdot \mathbf{L} = (\hat{\mathbf{e}}_r \cos\theta - \hat{\mathbf{e}}_\theta \sin\theta) \cdot \left(-i\hbar\,\hat{\mathbf{e}}_\phi \frac{\partial}{\partial \theta} + i\hbar\,\hat{\mathbf{e}}_\theta \frac{1}{\sin\theta} \frac{\partial}{\partial \phi}\right)$$

$$= -i\hbar \frac{\partial}{\partial \phi}$$

Now

$$[H, L^2] = \left[-\frac{\hbar^2 \nabla^2}{2\mu} + V(r), L^2\right]$$

$$= \left[-\frac{\hbar^2}{2\mu} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r}\right) + \frac{L^2}{2\mu r^2} + V(r), L^2\right] \quad \text{(Using Eq. 9.6a)}$$

$$= 0 \quad [\because r, \theta \text{ and } \phi \text{ are independent variables and } L^2 \text{ commutes with itself}]$$

Similarly

$$[H, L_z] = \left[-\frac{\hbar^2}{2\mu} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r}\right) + \frac{L^2}{2\mu r^2} + V(r), L_z\right] = 0 \quad (\because [L^2, L_z] = 0)$$

2.a) Using Eq. (9.6b) we can write

$$L_z Y_{l,m_l}(\theta,\phi) = c\, Y_{l,m_l}(\theta,\phi)$$

or

$$-\frac{ih}{2\pi}\frac{\partial}{\partial\phi} Y_{l,m_l}(\theta,\phi) = c\, Y_{l,m_l}(\theta,\phi)$$

where c is the eigenvalue of L_z, corresponding to $Y_{l,m_l}(\theta, \phi)$. Now L_z involves operation on ϕ alone and the ϕ component of $Y_{l,m_l}(\theta, \phi)$ is just $e^{im_l\phi}$. Hence, using Eq. (9.18), we get

$$-\frac{ih}{2\pi}\, im_l\, Y_{l,m_l}(\theta,\phi) = c\, Y_{l,m_l}(\theta,\phi)$$

$$\therefore\ c = m_l h/2\pi = m_l\, \hbar$$

(b) We have to show that

$$\int Y^*_{2,2}(\theta, \phi)\, Y_{2,2}(\theta, \phi)\, d\Omega = 1$$

or

$$\int_0^{\pi}\int_0^{2\pi} Y^*_{2,2}\, Y_{2,2} \sin\theta\, d\theta\, d\phi = 1$$

Let $I_1 = \int_0^{\pi}\int_0^{2\pi} Y^*_{2,2}\, Y_{2,2} \sin\theta\, d\theta\, d\phi$

$$= \left(\frac{15}{32\pi}\right) 2\pi \int_0^{\pi} \sin^4\theta \sin\theta\, d\theta$$

$$= \frac{15}{16}\int_{-1}^{1} (1 - 2\mu^2 + \mu^4)\, d\mu \text{ where } \mu = \cos\theta$$

or

$$I_1 = \frac{15}{16}\left(2 - \frac{4}{3} + \frac{2}{5}\right)$$

$$= \frac{15}{8}\frac{(15 - 10 + 3)}{15} = 1$$

Hence proved.

Next we have to show that $\int Y^*_{2,2}(\theta, \phi)\, Y_{2,-2}(\theta, \phi)\, d\Omega = 0$

Let $I_2 = \int_0^{\pi}\int_0^{2\pi} Y^*_{2,2}\, Y_{2,-2} \sin\theta\, d\theta\, d\phi$

$$= \frac{15}{32\pi}\int_0^{\pi}\int_0^{2\pi} \sin^4\theta\, e^{-4i\phi}\, d\theta\, d\phi$$

Now $\int_0^{2\pi} e^{-4i\phi}\, d\phi = 0$, $\therefore I_2 = 0$. Hence proved.

3. Replacing θ by $\pi - \theta$ and ϕ by $\pi + \phi$ we obtain

$$Y_{2,1}(\pi - \theta, \pi + \phi) = -\left(\frac{15}{8\pi}\right)^{1/2} \sin(\pi - \theta)\cos(\pi - \theta)\, e^{i(\pi+\phi)}$$

$$= -\left(\frac{15}{8\pi}\right)^{1/2} \sin\theta\cos\theta\, e^{i\phi} = Y_{2,1}(\theta, \phi)$$

Hence the parity of $Y_{2,1}(\theta, \phi)$ is even.

4. $R = \frac{me^4}{2\hbar^2} = 2.18 \times 10^{-18}$ Joules $= 13.6$ eV

also in units of m^{-1}, $R = \frac{me^4}{2\hbar^2}\frac{1}{ch} = 1.10 \times 10^7$ m^{-1}

An accurate value of the Rydberg constant is

$R = 1.09737373 \times 10^7$ m^{-1}.

5. We have to show that $\int_0^\infty R_{20}^*(r)\, R_{20}(r)\, r^2 dr = 1$

Let $I = \int_0^\infty R_{20}^2(r)\, r^2 dr.$

$$I = \frac{1}{(2a_0)^3} \int_0^\infty \left(4 - 4\frac{r}{a_0} + \frac{r^2}{a_0^2}\right) r^2\, e^{-r/a_0}\, dr.$$

Now we know that

$$\int_0^\infty r^p e^{-\beta r}\, dr = \frac{p!}{\beta^{p+1}}$$

hence $$I = \frac{1}{8a_0^3}\left[4\,\frac{2!}{\left(\frac{1}{a_0}\right)^3} - \frac{4}{a_0}\,\frac{3!}{\left(\frac{1}{a_0}\right)^4} + \frac{1}{a_0^2}\,\frac{4!}{\left(\frac{1}{a_0}\right)^5}\right]$$

$$= \frac{1}{8}[8 - 24 + 24]$$

$$= 1.$$

Furthermore,

$$\int_0^\infty R_{20}^*(r)\, R_{10}(r)\, r^2 dr = \frac{1}{\left(2a_0^6\right)^{1/2}} \int_0^\infty \left(2 - \frac{r}{a_0}\right) r^2 e^{-(3r/2a_0)}\, dr$$

$$= \frac{1}{\sqrt{2}\, a_0^3}\left[2\,\frac{2!}{\left(\frac{3}{2a_0}\right)^3} - \frac{1}{a_0}\,\frac{3!}{\left(\frac{3}{2a_0}\right)^4}\right] = 0.$$

6. Since $n = 3$, the permissible values of l are 0, 1 and 2. Hence, there are nine permissible values of m_l given by 0 (for $l = 0$) –1, 0 and 1 (for $l = 1$) and –2, –1, 0, 1 and +2 (for $l = 2$). Each combination of n, l and m_l gives rise to an eigenfunction. However the energy depends only upon n. Hence all these nine eigenfunctions are degenerate.

Terminal Questions

1. $$\sum_{m_l=-2}^{+2} = |Y_{2,m_l}(\theta, \phi)|^2 = 2\,\frac{15}{8\pi}\sin^2\theta\cos^2\theta + 2\,\frac{15}{32\pi}\sin^4\theta + \frac{5}{16\pi}(3\cos^2\theta - 1)^2$$

$$= \frac{5}{16\pi}[6\cos^4\theta - 6\cos^2\theta + 4 + 6\sin^2\theta\cos^2\theta]$$

$$= \frac{5}{4\pi}.$$

You may note that the sum $\sum_{m_l=-l}^{+l} |Y_{l,m_l}(\theta, \phi)|^2$ is always spherically symmetric.

2. $V(r) = -\frac{e^2}{r}$

$$\langle V(r)\rangle = \int |\psi_{100}(r)|^2 \left(-\frac{e^2}{r}\right) r^2\, dr \sin\theta\, d\theta\, d\phi$$

$$= \frac{1}{\pi a_0^3}\, 4\pi(-e^2)\int_0^\infty e^{-2r/a_0}\, r\, dr$$

$$= -\frac{4e^2}{a_0^3}\,\frac{1}{\left(\frac{2}{a_0}\right)^2} = -\frac{e^2}{a_0} = 2E_0. \qquad \left(\because a_0 = \frac{\hbar^2}{\mu e^2} \text{ and } E_0 = -\frac{\mu e^4}{2\hbar^2}\right)$$

Since kinetic energy + potential energy = total energy

$$\langle \text{K.E.}\rangle = E_0 - \langle V(r)\rangle = -\frac{e^2}{2a_0} + \frac{e^2}{a_0} = \frac{e^2}{2a_0} = |E_0|$$

3. Let the size of the atom be R. Since the electron is inside the atom, the uncertainty in the momentum is $p = \hbar/R$. The linear momentum of magnitude p can be in any direction so its components can have values from $-p$ to p. Hence the uncertainty in momentum is also approximately p. Hence we take $\Delta p = p$. Now we take

$$\langle \text{K.E.} \rangle = \frac{(\Delta p)^2}{2\mu} = \frac{\hbar^2}{2\mu R^2} = \text{K.E.}$$

$$E = \text{K.E.} + V = \frac{\hbar^2}{2\mu R^2} - \frac{e^2}{R} \text{ at } R.$$

Now $\frac{dE}{dR} = -\frac{\hbar^2}{\mu R^3} + \frac{e^2}{R^2} = 0$ for a stable atom,

hence $$R = \frac{\hbar^2}{\mu e^2} = a_0 .$$

Hence the size of the most stable atom is the first Bohr radius itself.

4. The probability of finding the electron between r and $r + dr$ is given by

$$|\psi_{100}(r)|\, r^2\, 4\pi dr = \frac{1}{\pi a_0^3} e^{-2r/a_0} r^2\, 4\pi dr$$

$$= \frac{4}{a_0^3} r^2 e^{-2r/a_0}\, dr.$$

Hence to determine the most probable value of r we differentiate $r^2\, e^{-2r/a_0}$ with respect to r and equate the result to zero. Thus we get

$$\left(r^2 \left(-\frac{2}{a_0}\right) + 2r\right) e^{-2r/a_0} = 0$$

or

$$r = a_0.$$

However the average value of r is given by

$$\langle r \rangle = \int |\psi_{100}(r)|^2\, r\, r^2\, dr \sin\theta\, d\theta\, d\phi$$

$$= \frac{1}{\pi a_0^3}\, 4\pi \int_0^\infty e^{-2r/a_0} r^3 dr$$

$$= \frac{4}{a_0^3} \frac{6}{\left(\frac{2}{a_0}\right)^4}$$

$$= \frac{3}{2} a_0.$$

UNIT 10 ATOMIC SPECTRA

Structure

10.1 Introduction
Objectives
10.2 Stern–Gerlach Experiment
10.3 Spin Angular Momentum
10.4 Total Angular Momentum
10.5 Spectral Terms, Optical Spectra of Hydrogen-like Atoms and Selection Rules
10.6 Multielectron Atoms
Life Time of Excited States and Line Broadening
10.7 Summary
10.8 Terminal Questions
10.9 Solutions and Answers

10.1 INTRODUCTION

In the previous unit you have learnt about the vector model of the atom and space quantization, in connection with the concept of angular momentum. Now since angular momentum is an observable, it must be measurable. You know that it is related to the magnetic moment of an atom. Therefore, if we are able to measure magnetic moment, we can measure angular momentum. One such experiment was devised in 1922 by O. Stern and W. Gerlach to measure angular momentum. In this unit we start by describing the Stern–Gerlach experiment which also verified the concept of space quantization.

However, the quantitative results obtained from the Stern–Gerlach experiment could not be explained with the help of only three quantum numbers, namely, n, l and m_l. Besides, there were a good number of spectroscopic data which could not be explained by the vector model of the atom.

To break the above deadlock, two research students S. A. Goudsmit and G.E. Uhlenbeck hypothesised that every electron possesses a spin angular momentum S. The spin angular momentum combines with the orbital angular momentum L to yield total angular momentum J. Like L, the angular momenta S and J also show space quantization and yield quantum numbers (s, m_s) and (j, m_j), respectively. You will study about these concepts in Secs. 10.3 and 10.4. The introduction of the spin angular momentum proved very valuable for the qualitative explanation of the Stern–Gerlach experimental results. It was also used to explain a large number of the existing spectroscopic data, especially for hydrogen-like atoms. In Sec. 10.5, we discuss the optical spectra of such atoms.

Finally, we discuss the spectra of multielectron atoms in Sec. 10.6. In such atoms, every electron moves in a field produced by the nucleus of the atom and the remaining electrons. Such a potential is *not* spherically symmetric. Hence strictly speaking, for a multielectron atom, the orbital angular momentum of the atomic electrons is not a constant of motion.

Nevertheless, it is a good approximation to assume that every atomic electron moves in a spherically symmetric potential and its energy state is characterized by four quantum numbers n, l, m_l and m_s or n, l, j and m_j. In this section, we will also discuss Pauli exclusion principle which has provided shell structure to the atoms. The above principle enables us to arrange the atomic electrons according to their energy states (**electronic configuration**). The electronic configuration is utilised to obtain L, S and J quantum numbers of the whole atom and thus the spectroscopic terms (defined by L, S and J) of the ground as well as excited states of the atoms.

You will learn the selection rules which are obeyed when an atom makes a transition from one state to another. A transition from an excited state to a lower one produces a spectral line of characteristic frequency. This gives rise to atomic spectra. Such

transitions take finite time for their completion. Thus every excited state has a finite life time τ. Hence according to the Heisenberg uncertainty principle, ($\Delta E\,\Delta t \sim \hbar$), every spectral line has a frequency width. The transition of inner electrons from one energy state to another give rise to spectra in a higher frequency region, namely, the *X*-ray spectra. In the next unit, which is also the last in this block, we shall briefly discuss *X*-ray spectra.

Objectives

After studying this unit you should be able to

- describe Stern–Gerlach experiment,
- explain the concept of spin angular momentum,
- calculate the total angular momentum,
- compute the spectral terms for hydrogen-like and multielectron atoms,
- distinguish between allowed and forbidden transitions.

10.2 STERN–GERLACH EXPERIMENT

In the previous unit you have seen that the angle between the angular momentum vector **L** and the z-axis takes only discrete values. The angle is given by $\cos^{-1}(m_l/\sqrt{l(l+1)})$ where for a given value of l, the magnetic quantum number m_l takes only integer values given by $-l, -l+1, \ldots\ldots l-1, l$. The phenomenon, known as *space quantization*, was verified by an experiment performed by Stern and Gerlach. In this experiment, a highly inhomogeneous magnetic field having magnetic induction **B** was applied along z-axis and an atomic beam travelling along x direction was passed through this field (see Fig. 10.1). It was found that a single atomic beam produced more than one trace on the screen S. This clearly showed that the inhomogeneous magnetic field resolved one single beam of atoms into more than one discrete component

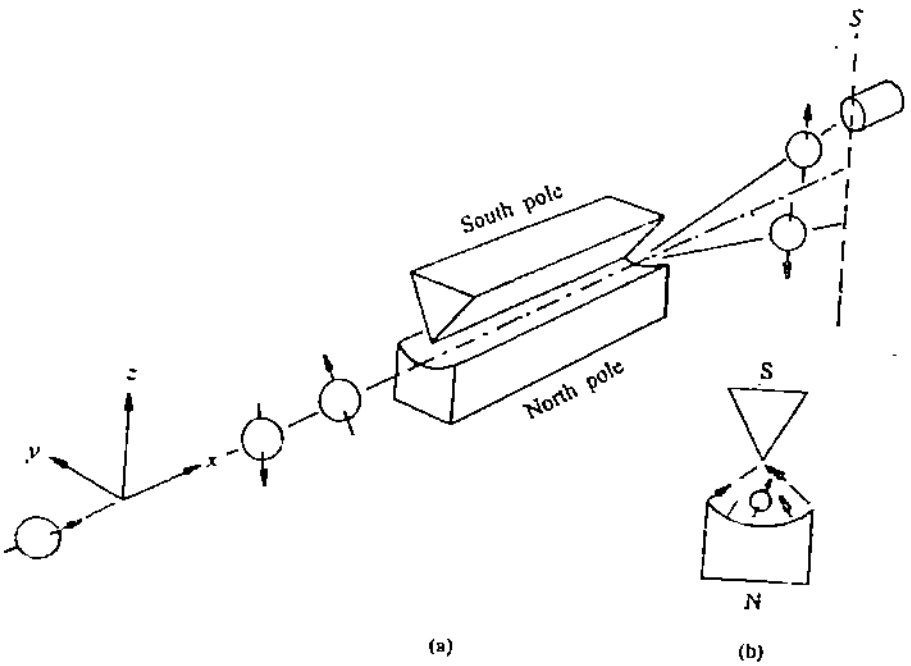


Fig. 10.1: Stern–Gerlach Experiment. Silver atoms produce two traces on the screen S. (a) The incoming beam along the x-direction is split into two as it passes through an inhomogeneous magnetic field; (b) how the inhomogeneity of the magnetic field is created. The field depends sharply on z, increasing with increasing z.

To understand this experimental result, let us assume that an electron of mass μ of the atom has an angular momentum **L**. Since this electron has a charge e and is moving inside the atom, according to classical physics, the electron also has a magnetic moment $\boldsymbol{\mu}_L$ given by

$$\boldsymbol{\mu}_L = -\frac{e}{2\mu}\,\mathbf{L} \tag{10.1}$$

Classically, the deflecting force on the atom due to the inhomogeneous magnetic field is given by

$$\mathbf{F} = \nabla\,(\boldsymbol{\mu}_L \cdot \mathbf{B}) \tag{10.2}$$

Since **B** is along z-axis, the force will also be along z-axis and we have

$$F_z = \mu_z \frac{\partial B}{\partial z} = -\frac{e}{2\mu}\, L_z \frac{\partial B}{\partial z} \tag{10.3}$$

Thus the particles in the beam should be deflected up or down in direct proportion to the z component of the magnetic dipole moment. By looking at the trace of the deflected particle on a screen, therefore, we can measure the component of its magnetic moment in the direction of the magnetic field.

In this way, the Stern–Gerlach apparatus does to a particle beam of magnetic dipoles what a prism does to white light; it refracts the magnetic dipoles and displays the spectrum of magnetic moment that the particles of the beam possess.

What would we expect classically if a beam of atomic dipoles is sent through the Stern–Gerlach prism? Classically, the spectrum of the z-component of magnetic moment — its allowed values — is continuous (ranging from $-\mu_z$ to μ_z). What did Stern and Gerlach find when they first did their experiment? A discrete line spectrum, of course! The spectrum of magnetic moment is quantized; it should be obvious; μ_z is proportional to L_z, and L_z is quantized.

To understand these results, let us convert the classical expression of F_z into the corresponding quantum mechanical expression. For this we regard F_z as an operator. Then the average value of F_z is given by

$$\langle F_z \rangle = \int \psi^*_{nlm_l}(\mathbf{r})\, F_z\, \psi_{nlm_l}(\mathbf{r})\, d\tau \tag{10.4}$$

Now $\psi_{nlm_l}(\mathbf{r})$ as given by Eq. (9.50) is an eigenfunction of L_z with the eigenvalue m_l. Hence the force acting on the atom in the z direction is

$$\langle F_z \rangle = -(e\hbar/2\mu)\, m_l \frac{\partial B}{\partial z} \tag{10.5}$$

The quantity $e\hbar/2\mu$ is termed the **Bohr magneton** μ_B.

Stern and Gerlach also measured the distance between the traces and concluded that m_l changes by one. Thus the concept of space quantization was experimentally verified.

However, as we have said above, the number of traces obtained in the Stern–Gerlach experiment could not be explained on the basis of the quantum numbers l and m_l. It is evident from Eq. (10.5) that due to discrete values of m_l a single beam of atoms will break into $(2l + 1)$ beams and each will produce its own trace on the screen S. Hence the number of traces must always be odd. But unexpectedly, when a beam of silver atoms was sent through the inhomogeneous magnetic field, it produced only two (an even number) traces. It is well known that the valency of silver is one. Hence it has only one active electron and in the ground state this electron is an s ($l = 0$) electron. Hence the only possible value of m_l is zero. Therefore, there should not have been any deflection and the screen S should have registered only one trace and not two. On the other hand if one assumes that the silver atoms were in the p ($l = 1$) state then the number of traces should have been three. Hence, it became evident that the results of Stern–Gerlach experiment could not be explained on the basis of l and m_l quantum numbers. Something was found to be missing. This discrepancy was resolved by

Goudsmit and Unlenbeck who introduced the concept of **spin angular momentum**. You will need to study this concept carefully.

But before that, here's an exercise for you.

SAQ 1

Spend 5 min

Show that the value of the Bohr magneton in SI unit is 9.27×10^{-24} joule tesla^{-1} (or amp meter^2). Take μ as the rest mass of the electron.

10.3 SPIN ANGULAR MOMENTUM

Goudsmit and Uhlenbeck analysed a large number of spectroscopic lines and found that these also could not be explained with the help of only three quantum numbers, namely, n, l and m_l. One of the familiar example is the observation of two closely spaced lines D_1 and D_2 in the spectrum of a sodium lamp. On the basis of the above three quantum numbers there should have been only one line corresponding to $3p$ to $3s$ transition. To explain such discrepancies, Goudsmit and Uhlenbeck advanced a new hypothesis according to which **every electron has an intrinsic angular momentum S and consequently an intrinsic magnetic moment $\boldsymbol{\mu}_s$ also.** However, the ratio of $\boldsymbol{\mu}_s$ and S is $-e/\mu$ and not $-e/2\mu$. Their argument was based on this logic: a free electron does not have any orbital angular momentum. Thus the two-valuedness evident in the traces must be due to an intrinsic angular momentum, which is a purely quantum mechanical attribute of particles. They called it the spin angular momentum (S), or simply spin. Since the multiplicity $(2s + 1)$ is 2, we can identify the angular momentum quantum numbers associated with spin as 1/2. Thus electrons have spin 1/2.

What about the vector S? Just like the orbital angular momentum L, the spin angular momentum S also precesses about z-axis, The angle between S and z-axis is quantized, i.e., it also shows space quantization. However, since the value of the quantum number s associated with S is 1/2, m_s has only two values equal to $\pm s$, i.e., $\pm$ 1/2. Consequently, there are only two spin functions, one corresponding to m_s = 1/2 (spin up) and the other to $m_s = -$ 1/2 (spin down). These two spin functions are the eigenfunctions of the operators S^2 and S_z with eigenvalues $s\ (s + 1)\hbar^2$ and m_s, respectively. Hence, the quantized angles, obtained by taking s = 1/2 and $m_s = \pm$ 1/2, in place of l and m_l respectively, in Eq. (9.22), are $\pm \cos^{-1}(1/\sqrt{3})$.

With the help of the above hypothesis it is easy to explain the production of only two traces by the ground state of the silver atoms in the Stern–Gerlach experiment. Since $l = 0$ and s = 1/2 we get $m_l = 0$ and $m_s = \pm$ 1/2. Each m_s produces its own trace and one silver atomic beam splits into two.

The concepts introduced above will take you some time to get used to. Here is a note of warning; you should take spin angular momentum and the corresponding magnetic moment as the intrinsic properties of an electron, just like its charge and mass. **Do not picture the electron as some sort of a spinning top spinning about its axis.** Such a description is wrong and leads to absurd results as you will discover in SAQ 2. Spin is one more strangeness of the quantum world that you will have to get accustomed to!

You may wonder; why does the electron or any other particle possess a spin angular momentum? An attempt to understand the reason for the occurrence of spin angular momentum will take us much beyond the scope of the present course. You will be able to learn about this in your higher degree studies.

SAQ 2

Spend 5 min

Show that if we assign the intrinsic angular momentum of an electron due to the spinning of the electron about its own axis, the velocity of the electron will be greater than the velocity of light.

The spin angular momentum vector S does not depend upon the space coordinates hence its origin is purely quantum mechanical. Further, since S^2 and S_z are constants of motion, the eigenfunction of an atomic electron is characterised by four quantum numbers n, l, m_l and m_s (the value of s is always 1/2). Thus spin provides a fourth degree of freedom to an electron — the electron is a four-dimensional particle. We need four measurements — so we have four dimensions!

The introduction of spin angular momentum, leads us to the concept of total angular momentum.

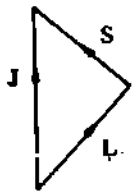


Fig. 10.2: Vector addition of L and S vectors.

10.4 TOTAL ANGULAR MOMENTUM

The addition of the vectors L and S gives rise to total angular momentum vector J (see Fig. 10.2).

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \tag{10.6}$$

The values of the quantum number j are given by

$$j = l+s,\ l+s-1,\ \ldots\ldots\ |l-s|+1,\ |l-s| \tag{10.7}$$

Thus vector J also precesses around z-axis and the angle between J and z-axis is given by $\cos^{-1}\left(\frac{m_j}{\sqrt{j(j+1)}}\right)$, where j is the total angular momentum quantum number and m_j is its component along z-axis (see Fig. 10.3). The wave functions of an electron, including its spin part, are eigenfunctions of J^2 and J_z with eigenvalues $j\,(j+1)\hbar^2$ and $m_j\hbar$, respectively. Like l, the quantum number j is also always positive but it can have integral as well as half-integral values. For example, if $l = 2$ and $s = 1/2$, the values of j are 5/2 and 3/2. Corresponding to $j = 5/2$, the values of m_j are 5/2, 3/2, 1/2, −1/2, −3/2 and −5/2 (see Fig. 10.4).

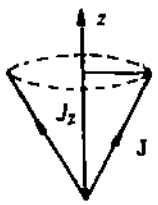


Fig. 10.3: Precession of J around z-axis.

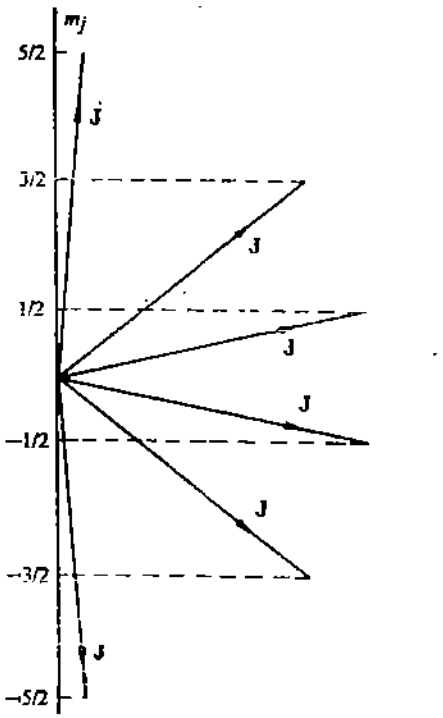


Fig. 10.4: Different discrete orientations of J vector and values of m_j for $|J| = 5/2$. The tip of J lies on a circle.

You should go slow, read carefully and absorb these ideas before studying further. The following exercise should further help you understand and concretise these ideas in the context of what you have learnt so far about the quantum world.

SAQ 3

Spend 10 min

The spin up and spin down wavefunctions of an electron are given by

$$\alpha = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \beta = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

and the spin angular momentum operator is given by

$$\mathbf{S} = \frac{\hbar}{2}\sigma$$

where σ is Pauli spin matrix and its three components are given by

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}$$

Show that α and β are eigenfunctions of S^2 and S_z and are orthogonal to each other.

Thus, through SAQ 3, you have essentially solved the eigenvalue problem of the operators S^2 and S_z for the electron. You have come to know of the spin wave functions of an electron and their eigenvalues. We have also briefly introduced you to the Pauli spin matrix σ and its three components. The concepts developed so far help us explain atomic spectra in the optical region of the electromagnetic spectrum.

10.5 SPECTRAL TERMS, OPTICAL SPECTRA OF HYDROGEN-LIKE ATOMS AND SELECTION RULES

All atomic systems having only one electron are classified as hydrogen-like atoms. Thus singly ionized helium (He^+), doubly ionized lithium (Li^{++}) and triply ionized beryllium (Be^{+++}) etc. are examples of hydrogen-like atoms. In general, a hydrogen-like atom has Z positive charges in its nucleus and one electron outside the nucleus. Hence the Hamiltonian of the relative motion is given by

$$H = -\frac{\hbar^2 \nabla^2}{2\mu_A} - \frac{Ze^2}{r} \tag{10.8}$$

where μ_A is the reduced mass of the system. Thus the time independent Schrödinger equation of a hydrogen-like atom will be the same as that of a hydrogen atom with the difference that e^2 and μ of Eq. (9.37a) will be replaced by Ze^2 and μ_A, respectively. Consequently, the eigenenergies and eigenfunctions of the above system will be given by Eqs. (9.44) and (9.50), respectively but e^2 and μ will have to be replaced by Ze^2 and μ_A at all the places.

The inclusion of spin gives two eigenfunctions for the same n, l and m_l: one for $m_s = +1/2$ and the other for $m_s = -1/2$. These are

$$\begin{aligned} \psi_{nlm_l}(\mathbf{r}, s) &= \psi_{nlm_l}(\mathbf{r})\,\alpha \qquad \text{for spin up} \\ &= \psi_{nlm_l}(\mathbf{r})\,\beta \qquad \text{for spin down} \end{aligned} \tag{10.9}$$

Now for a hydrogen-like atom $s = 1/2$. Hence for a given value of l, according to Eq. (10.7), there will be two values of j given by $l \pm 1/2$. You can now learn to write the spectral terms for hydrogen-like atoms.

In spectroscopy, the spectral term of an atom is given by

$$\boxed{{}^{2S+1}L_J}$$

where S is the total spin angular momentum, L is the total orbital angular momentum and J is the total angular momentum. The quantity $2S + 1$ is known as **multiplicity**. For $L > S$, we obtain $2S + 1$ values of J. However, for $L < S$ the values of J are $2L + 1$. Sometimes the numerical value of n is also attached in the spectral term, which is then written as $n\,{}^{2S+1}L_J$.

Now let us consider the state of the hydrogenic atom for which $l = 0$. For this case $S = \frac{1}{2}$, $L = 0$, $J = \frac{1}{2}$. Thus we may write for a hydrogen-like atom that

$$J = \frac{1}{2} \text{ for any } l = 0 \text{ state}$$

And so we have only one term for a hydrogenic (or hydrogen-like) atom given by $^2S_{1/2}$. Notice that in the spectral term we have used the symbol S instead of the numerical value 0 to show that $L = 0$ for this case. Similarly, for $L = 1, 2, 3, \ldots$ we use the symbols $P, D, F \ldots$ etc, in the spectral terms instead of the numerical values of L. So we use the capital letter to denote the orbital quantum number according to the following letter code:

$$S \text{ for } L = 0, \quad P \text{ for } L = 1, \quad D \text{ for } L = 2, \quad F \text{ for } L = 3, \ldots$$

Now let us consider the states for which $l \neq 0$.

For $l = 1$ and $s = \frac{1}{2}$, $L = 1$, $S = \frac{1}{2}$ and $J = \frac{3}{2}, \frac{1}{2}$

which yields two terms: $^2P_{1/2}$ and $^2P_{3/2}$.

Similarly, when $l = 2$ and $s = \frac{1}{2}$, $L = 2$, $S = \frac{1}{2}$ and $J = \frac{5}{2}, \frac{3}{2}$

This again yields two terms $^2D_{3/2}$ and $^2D_{5/2}$. In general, you can see that for a hydrogen-like atom only two possible values of J will result:

$$J = L + \frac{1}{2} \text{ or } L - \frac{1}{2} \text{ for any } l \neq 0 \text{ state}$$

Thus all the states with $l \geq 1$ are doublets, i.e., they have two values of J. Have you noticed from these results that for $n = 1$ we have only one $^2S_{1/2}$ state but $n = 2$ gives rise to three excited states? Our advice to you is, don't just read through what we have said in these lines. Work these numbers out. And to familiarise yourself further with the calculation of spectral terms work out the following SAQ.

Spend 10 min

SAQ 4

Determine all the spectral terms for $n = 2$ and $n = 3$ for a hydrogen-like atom.

According to Eq. (9.44) the eigenenergy depends only upon n hence the inclusion of spin increases the degeneracy to $2n^2$. However, if we include relativistic effects and the interactions between orbital angular momentum and spin angular momentum in our quantum mechanical treatment, it can be shown that the eigenenergy depends upon the quantum number j in addition to the quantum number n. It is given by

$$E_{nj} = -\frac{RZ^2}{n^2}\left[1 + \frac{\alpha^2 Z^2}{n}\left(\frac{1}{j+1/2} - \frac{3}{4n}\right)\right] \tag{10.10}$$

where α is called the **fine structure constant** and is equal to $e^2/\hbar c$ (=1/137). Hence the energy level diagram of a hydrogen atom as shown by Fig. 9.5b is modified to Fig. 10.5.

Notice that although $^2S_{1/2}$ and $^2P_{1/2}$ states are still degenerate, the state $^2P_{3/2}$ lies higher than $^2P_{1/2}$. Similarly, $^2D_{5/2}$ lies higher than $^2D_{3/2}$ in the energy level diagram. Thus, the spectrum of hydrogen atom and hydrogen-like atoms has a fine structure. We will come to this point again after you have done this exercise.

SAQ 5

Spend 5 min

Use Eq. (10.10) to obtain the energy difference between $3^2P_{3/2}$ and $3^2P_{1/2}$ states for the hydrogen atom in electron volts.

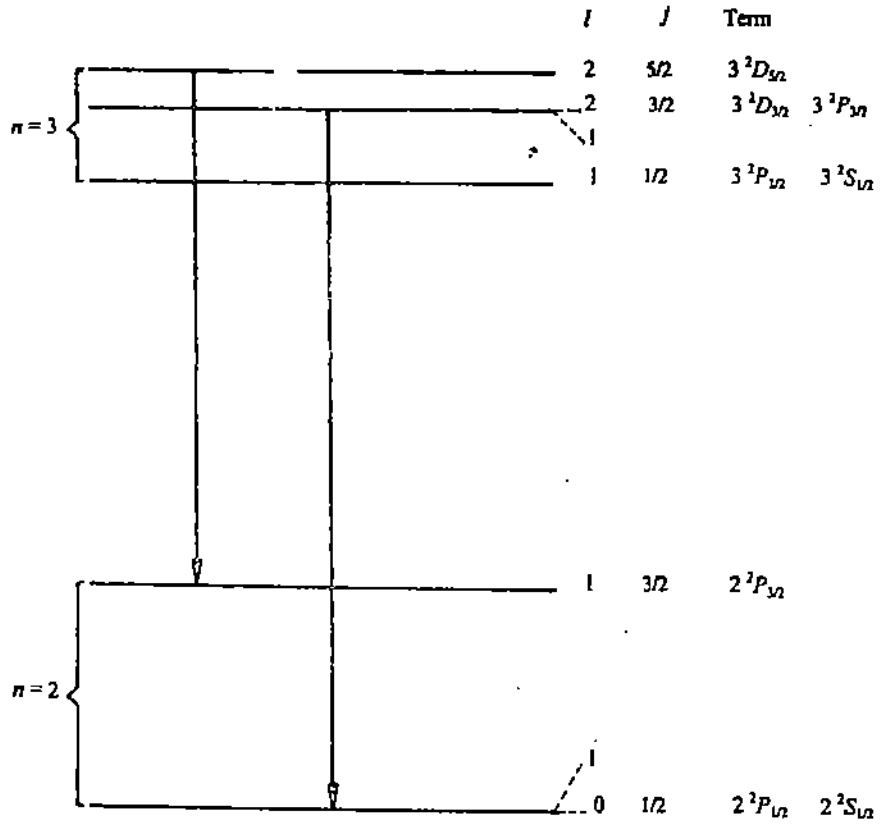


Fig. 10.5: Doublet structure of H_α line in hydrogen atom.

Since an atom in an excited state has more energy than its ground state, it has a natural tendency to make a transition to lower excited states or ground state and thereby reduce its energy by emitting electromagnetic radiations. An atom can also absorb electromagnetic energy and make a transition from a lower state to higher excited states. However all transitions are not allowed. Quantum mechanics gives certain selection rules for **allowed transitions** which we will now state. The remaining transitions are forbidden – they cannot occur.

Selection Rules

Suppose a hydrogen-like atom makes a transition from a state characterised by nlj to $n'l'j'$. Then the **selection rules** for the allowed transitions are as follows:

Selection rules for hydrogen-like atoms

(i) $j' - j = \pm 1, 0$ or $\Delta j = \pm 1, 0$

but both j' and j cannot be zero simultaneously. (10.11)

(ii) $l' - l = \pm 1,$ or $\Delta l = \pm 1.$

Notice that there is no restriction on the values of $n' - n$. Since the parity of a state is given by $(-1)^l$ it changes in an allowed transition. According to the above rules a transition from $3\,^2P_{3/2}$ to $2\,^2S_{1/2}$ is allowed since in this case the change in the value of l, i.e., $\Delta l = 1 - 0 = 1$. Is the transition from $3^2D_{5/2}$ to $1\,^2S_{1/2}$ allowed? No, because $\Delta l = 2 - 0 = 2$ in this case.

All those transitions which do not obey the above selection rules are said to be **forbidden transitions**. Some of the forbidden transitions do take place but with intensities which are about 10^4 times smaller than those of the allowed transitions. Using the energy level diagram shown in Fig. 10.5, we can easily explain the doublet structure, observed experimentally, of the first member of the Balmer series. They are produced by the following two transitions: $3\,^2D_{5/2}$ to $2\,^2P_{3/2}$ and $3\,^2P_{3/2}$ to $2\,^2S_{1/2}$. Similarly D_1 and D_2 lines in a sodium lamp are produced by the transitions $3\,^2P_{3/2}$ to $3\,^2S_{1/2}$ and $3\,^2P_{1/2}$ to $3\,^2S_{1/2}$. Not only the above structures but many other spectral features can be explained by associating a spin angular momentum to every electron and using the selection rules given by Eq. (10.11).

Let us now extend these ideas to multielectrons atoms.

10.6 MULTIELECTRON ATOMS

Let us consider atoms having more than one electron. If $\mathbf{l}_i$ is the orbital angular momentum vector of the i^{th} atomic electron, the total angular momentum vector $\mathbf{L}$ of the whole system is given by

$$\mathbf{L} = \sum_i \mathbf{l}_i \tag{10.12}$$

To obtain the values of the quantum number L we first combine l_1 and l_2 to obtain their resultant l_R. Then this l_R is combined with l_3 to obtain new resultants l_R'. This process is repeated until all electrons are taken into consideration. According to the vector model of the atom the resultant l_R has following values

$$l_R = l_1 + l_2,\ l_1 + l_2 - 1,\ \ldots,\ |l_1 - l_2| + 1,\ |l_1 - l_2| \tag{10.13}$$

Following the above rule each l_R will combine with l_3 to yield

$$l_R' = l_R + l_3,\ l_R + l_3 - 1,\ \ldots\ |l_3 - l_R| + 1,\ |l_3 - l_R| \tag{10.14}$$

etc. For example if $l_1 = 1$, $l_2 = 1$ and $l_3 = 2$ the values of l_R will be 2, 1 and 0 and the values of l_R' will be (4, 3, 2, 1, 0), (3, 2, 1) and 2. Similarly, for the spin angular momentum we have

$$\mathbf{S} = \sum_i \mathbf{s}_i \tag{10.15}$$

and the values of the resultant S are obtained by following Eqs. (10.12) to (10.14) by replacing l_i with s_i. Since for each electron $s_i = 1/2$, therefore, for two electrons $S = 1$ and 0 and for three electrons $S = (3/2, 1/2)$ and $1/2$. You can easily verify that for an atom having even number of electrons the resultant S will have integer values but for atoms having odd number of electrons the values of the resultant S will be half integers.

Finally $\mathbf{L}$ and $\mathbf{S}$ combine according to (10.6) to give the values of the total angular momentum quantum number J and the spectral terms given by $^{2S+1}L_J$ are obtained. The values of J are obtained from (10.13) by replacing l_R, l_1 and l_2 with J, L and S, respectively. Thus

$$J = L + S,\ L + S - 1,\ \ldots,\ |L - S| + 1,\ |L - S| \tag{10.16}$$

This kind of addition of angular momenta is called the *LS coupling*. There is another kind of addition of angular momenta called the *JJ coupling* — we are not going into its details here.

Let us now consider the assignment of four quantum number (n, l, m_l, m_s) to individual electrons in a multielectron atom. For the stability of the atom its energy should be minimum. It follows then that for all the electrons in an atom we should have $n = 1$, $l = 0$, $m_l = 0$ and $m_s = +1/2$ or $-1/2$. However, the above assignment is not correct. The distribution of atomic electrons according to their quantum numbers n and l is known as the **electronic configuration of an atom**. It follows the **Pauli exclusion principle**. We will discuss it in brief here.

According to this principle **no two electrons in an atom can have the same four quantum numbers** (n, l, m_l, m_s). For example, in the ground state of a helium atom, the four quantum numbers associated with two atomic electrons are $(1, 0, 0, +1/2)$ and $(1, 0, 0, -1/2)$. Thus, the first three quantum numbers n, l and m_l are the same, but the fourth quantum number m_s is different for the two electrons. However, in the excited states of helium one electron may be in $(1, 0, 0 + 1/2)$ state while the other electron may have any other value of the quantum numbers n, l, m_l and m_s. Now if $s_1 = \frac{1}{2}$ and $s_2 = \frac{1}{2}$, the resultant value of S can be 1 as well as 0 and we have two types of terms given by ${}^3L_J$ and ${}^1L_J$. For the first term, known as a **triplet**, the three values of J are $L+1$, L, $L-1$ with $L \neq 1$. On the other hand for the second term, known as **singlet**, J has only one value equal to L.

With the help of Pauli's exclusion principle, we can now describe the **shell structure** of an atom.

As you have studied earlier in this unit, the central field in a multielectron atom is not spherically symmetric. Therefore, the energy of the system depends upon L as well as S. A quantum mechanical calculation shows that for a helium atom, the triplet terms have lower energies in comparison with the corresponding singlet terms. But to a good approximation, the energy of **each atomic electron** depends upon its principal quantum number n and orbital angular momentum quantum number l, unlike the situation for the hydrogen atom. These energies must not depend on m_l and m_s because the potential energy of each electron is spherically symmetric and spin independent. Hence, there are $2(2l + 1)$ degenerate states with the same energy $E_{n\,l}$, corresponding to the two possible values of m_s and $(2l + 1)$ possible values of m_l. All of them will have nearly the same energy. This group of $2(2l + 1)$ spin-orbitals constitutes **an atomic subshell** at energy $E_{n,l}$ for each pair of quantum numbers n and l. They are said to form an **$(n\,l)$ energy subshell**. Thus in an atom, for $n = 2$ and $l = 1$, there will be $2(2 \times 1 + 1) = 6$ electrons in the $(2p)$ subshell.

Each value of n determines an **electron shell** for a given atom. The number of electrons in the nth shell is $2n^2$. Thus, in the shell with $n = 1$, there will be 2 electrons; in the $n = 2$ shell there will be 8 electrons, and so on. Further, a shell consists of n subshells labelled by n and l, as l ranges from 0 to $n - 1$. And every $n\,l$ subshell contains $2\,(2l + 1)$ spin-orbital states. For example, for $n = 2$, $l = 0$ and 1, there will be two subshells $(2s)$ and $(2p)$. In the $2s$ subshell there will be $2(2 \times 0 + 1) = 2$ electrons. In the $2p$ subshell there will be 6 electrons, and so on.

These shells are also labelled K, L, M, N, ... according to whether $n = 1, 2, 3, 4, \ldots$ In accordance with the Pauli exclusion principle, all electrons in a shell are characterised by different set of four quantum numbers and different eigenfunctions.

To sum up, the **ground state** of an atom with atomic number Z is the minimum energy configuration for the bound system of Z electrons organised in a shell structure. Starting with the inner-most electrons, the order of the shells and subshells with increasing energy is given as

$$1s,\ 2s,\ 2p,\ 3s,\ 3p,\ [4s,\ 3d],\ 4p,\ [5s,\ 4d],\ 5p,\ [6s,\ 4f,\ 5d],\ \ldots \qquad (10.17)$$

The numbers 1, 2, 3, ... represent the values of n and the letters, s, p, d, f ... correspond to $l = 0, 1, 2, 3$ .... The bracket in Eq. (10.17) enclose levels which have very nearly the same value of energy. Although we have said that electrons having same values of n and l have nearly the same energy but a detailed study shows that the rules for the filling of subshells that hold throughout the periodic table of elements are as follows:

1. Subshells are grouped under like values of $n + l$.
2. Groups are filled in the order of increasing $n + l$.
3. Within each $n + l$ group, subshells are filled in the order of decreasing l values.

Let us now apply these ideas to determine the electronic structure of elements in the periodic table, i.e., to write down the ground state configuration for any atom in the periodic table.

The Periodic System of the Elements

The ground state configuration of a neutral atom with Z electrons is obtained by distributing them according to the rules given above.

The first element ($Z = 1$) is atomic hydrogen, which has the ground state configuration $1s$. For $Z = 2$ (helium), both electrons occupy the $1s$ level and the configuration is $1s^2$. Thus, we may write their electron configuration as follows:

H	He
$1s$	$1s^2$

For the lithium atom, there are three electrons in its ground state and its electronic configuration is $1s^2\ 2s^1$, because $1s^3$ is forbidden by Pauli's exclusion principle. For $Z = 4$ (beryllium), the configuration is $1s^2\ 2s^2$. Thus we have

Li	Be
$1s^2 2s^1$	$1s^2 2s^2$

The next element is boron ($Z = 5$). Since the K shell and $2s$ subshell are full, the fifth electron occupies the $2p$ subshell. As Z increases from 5 to 10, the electrons fill the $2p$ subshell progressively, so that we have for $Z = 5$ to 10:

B	C	N	O	F	Ne
$1s^2\ 2s^2\ 2p^1$	$\cdots 2s^2\ 2p^2$	$\cdots 2s^2\ 2p^3$	$\cdots 2s^2\ 2p^4$	$\cdots 2s^2\ 2p^5$	$\cdots 2s^2\ 2p^6$

Then, the electronic configuration of a sodium atom having 11 electrons is $1s^2\ 2s^2\ 2p^6\ 3s^1$. It has four subshells having different values of $(n\,l)$. Out of these four subshells, the first three have maximum permissible number of electrons. Such subshells are said to be **closed subshells**. The last one is called an **open subshell.**

From $Z = 11$ to $Z = 18$ (argon), the $3p$ levels fill progressively. For $Z = 19$ (potassium), you might expect the nineteenth electron to go to $3d$ level but the $4s$ level has a lower energy than $3d$ level. Therefore, the ground state electronic configuration of a potassium atom is $1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^6\ 4s^1$ **and not** $1s^2\ 2s^2\ 2p^6\ 3s^6\ 3p^6\ 3d^1$. Similarly, the ground state electronic configuration of Scandium atom ($Z = 21$) is $1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^6\ 4s^2\ 3d^1$. These configurations agree with the experimental observations. You should concretise these ideas in your mind before studying further. The following SAQ will help you do this.

Spend 10 min

SAQ 6

Determine the electronic configurations for atoms with $Z = 20$, $Z = 25$, $Z = 31$ and $Z = 37$.

Spectral Terms of a Multielectron Atom

How do we determine the spectral terms of a multielectron atom? For such atoms, only the open subshells contribute towards total L, S and J values.

Let us consider the carbon atom ($Z = 6$) as an example. Its electronic configuration is $1s^2\ 2s^2\ 2p^2$. Since only the open subshell matters, let us consider the two electrons in the $2p$ subshell. For these electrons:

$$n_1 = 2,\ l_1 = 1,\ s_1 = \frac{1}{2} \text{ and } n_2 = 2,\ l_2 = 1,\ s_2 = \frac{1}{2}$$

Thus, the total spin $S = \frac{1}{2} + \frac{1}{2}, \left|\frac{1}{2} - \frac{1}{2}\right| = 1, 0$

and the total orbital angular momentum

$$L = 1 + 1,\ 1 + 1 - 1,\ |1 - 1|$$
$$= 2,\ 1,\ 0.$$

Thus for $S = 0$ the possible J values are $J = 0, 1, 2$ and for $S = 1$, $J = 1$, (2, 1, 0), (3, 2, 1).

Now the Pauli exclusion principle also tells us that $S = 0$ *should correspond to even values of* L and $S = 1$ *goes with states of odd* L. Therefore, we can only have the following L, S, J combinations:

$S = 0$, $L = 0$ whence $J = 0$

$S = 0$, $L = 2$ whence $J = 2$

$S = 1$, $L = 1$ whence $J = 2, 1, 0$

Thus, the spectral terms are

$$^1S_0,\ ^3P_0,\ ^3P_1,\ ^3P_2 \text{ and } ^1D_2$$

Now the question is: *which of these has the lowest energy*? We can determine this using the three **Hund's rules** given as follows:

1. The state with the largest spin has lowest energy.
2. If the incomplete subshell is less than half full, $J = |L - S|$ is the ground state: if it is more than half full $J = L + S$ is the ground state.
3. Among the levels with a given value of S, the state with the largest value of L has the smallest energy.

Let us apply these rules to Carbon atom. Rule 1 tells us that the lowest state is one of the ^3P states. Rule 2 tells us that the lowest energy state is the one with $J = 0$. Thus, the ground state of carbon atom is the $^3P_0$ state. Using Rule 3, we can tell that $^1D_2$ has lower energy than $^1S_0$, although neither is the ground state. Incidentally, Hund's rules are not adhoc — they agree with quantum mechanical calculations.

The spectral terms are very important in spectroscopy where the L, S, J quantum numbers are part and parcel of the selection rules that govern transitions between atomic states. We will briefly state them. But before that you should fix the ideas presented here by doing the following exercise.

SAQ 7

Spend 10 min

Obtain the ground state terms of He, Li, Si, and Sc.

Atoms can also exist in excited states. However, to minimise their energies they make transitions from one excited state to lower excited states and ground state. The **selection rules** for allowed transitions in multielectrons atoms are as follows:

$$\text{(i)}\quad \Delta J = 0, \pm 1 \quad (J = 0 \longrightarrow J' = 0 \text{ is not allowed})$$
$$\text{(ii)}\quad \Delta L = 0, \pm 1 \qquad (10.18)$$
$$\text{(iii)}\quad \Delta S = 0$$
$$\text{(iv)}\quad \Delta l = \pm 1$$

where l is the orbital angular momentum quantum number of the atomic electron which takes part in the transition. The same rules are followed when an atom absorbs electromagnetic radiation and makes a transition from a lower state to a higher excited state. One of the features of atomic line spectra is that the width of the emitted line is finite. This feature can be explained with the help of the concepts you have studied so far — it arises due to the finite life time of excited states. We will now discuss it briefly.

10.6.1 Life Time of Excited States and Line Broadening

Let us consider two stationary states of an atom having energies E_1 and E_2 ($> E_1$). If the atom is in the upper state it will have a natural tendency to make a transition to the lower state and during the transition it will radiate electromagnetic radiations. At the start of the transition the atomic energy was E_2 and at the end the atomic energy is E_1. Hence the excess energy decreases with time. The reduction of the excess energy with time is found to be exponential. At time t the excess energy is given by

$$E = E_0 \exp(-\gamma t) \tag{10.19a}$$

where $E_0 = E_2 - E_1$. Now the life time τ of an excited state is defined to be that time in which E reduces to a value E_0/e. You can readily verify that the above definition yields

$$\tau = 1/\gamma \tag{10.19b}$$

Due to the decrease of the excess energy E with time, the radiation emitted by the atom is not strictly monochromatic. A plot of $I(\nu)$ (energy emitted per unit frequency range) versus the frequency ν of the emitted radiation gives a curve as shown in Fig. 10.6.

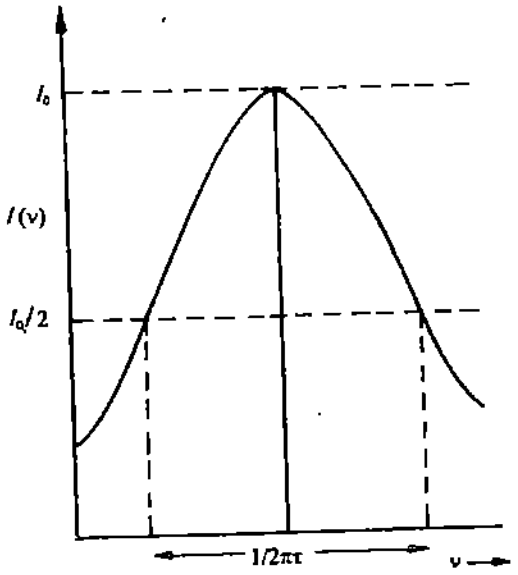


Fig. 10.6: Spectral distribution of emitted radiation.

We find that the frequency width of the curve at $I_0/2$, where I_0 is the maximum value of $I(\lambda)$, is $1/2\pi\tau$. Hence the energy width of the spectral line is given by

$$\Delta E = h\,\Delta\nu = h/2\pi\tau$$

or

$$\tau\,\Delta E = \hbar \tag{10.19c}$$

Thus the smaller is the life time, larger is the energy width of the emitted line. You may recall that the width of a line is a consequence of the uncertainty principle and τ is known as the **natural life time**. For excited states which are connected to the lower states by allowed transitions, the life time is of the order of 10^{-8} s. For other excited states, the life time is much longer and can be equal to several seconds. According to Eq. (10.19c), the increase of τ reduces the width of the spectral line.

Let us now summarise what you have studied in this unit.

10.7 SUMMARY

- In this unit you have studied the **Stern–Gerlach experiment** which measured the magnetic moment of an atom and also gave a direct verification of space quantization. The quantitative results of the Stern–Gerlach experiment could not be

explained by the help of three quantum numbers n, l and m_l. Goudsmit and Uhlenbeck introduced the concept of an intrinsic angular momentum known as spin angular momentum **S** associated with an electron. The quantum number associated with **S** is always 1/2 giving rise to $m_s = \pm 1/2$.

- The **orbital and spin angular momenta** couple to produce **total angular momentum J** and quantum numbers j and m_j. Due to **spin-orbit coupling** (combination of **L** and **S**) and relativistic effects, the energy of hydrogen-like atoms depends upon n as well as j. This explains the doublet structure of the first member of Balmer series and the two close spectral lines produced in a sodium lamp.
- Certain **selection rules** are followed when a hydrogen-like atom or a multielectron atom makes a transition from one stationary state to another.
- Every excited state has a **finite life time** τ and when an atom makes a transition from one state to another the spectral line so produced has a **finite width**: $\Delta E = \hbar/\tau$.

10.8 TERMINAL QUESTIONS

Spend 30 min

1. Use expressions given for **S** in SAQ 3 to show that it satisfies the following operator equation

$$\mathbf{S} \times \mathbf{S} = i\hbar\mathbf{S}$$

2. The number of allowed values of J for two different atoms in P and D states is the same and equal to three. Determine the spin angular momentum of the atoms in these states.

3. State with reasons whether following transitions for a multielectron atom are allowed:

(i) $^3P_0 \longrightarrow {}^3S_1$

(ii) $^3S_1 \longrightarrow {}^1S_0$

(iii) $^1S_{1/2} \longrightarrow {}^1P_{1/2}$

(iv) $^1S_{1/2} \longrightarrow {}^1D_{3/2}$

10.9 SOLUTIONS AND ANSWERS

Self-Assessment Questions

1. $$\mu_B = e\hbar/2\mu = \frac{1.60 \times 10^{-19}\ \text{C} \times 1.054 \times 10^{-34}\ \text{Js}}{2 \times 9.109 \times 10^{-31}\ \text{kg}}$$

$$= 9.27 \times 10^{-24}\ \text{Joule Tesla}^{-1}\ \text{(or amp. meter}^2\text{)}$$

2. Let the radius and the linear velocity of the spinning electron be r and v, respectively. The spinning of the electron about its own axis produces a current of the order of magnitude $ev/2\pi r$. The corresponding magnetic moment is given by

$$\mu_B = IA = \frac{ev}{2\pi r}\ \pi r^2$$

Equating the above quantity to the experimental value $e\hbar/2\mu$ we get

$$v = \hbar/\mu r$$

Now the classical radius of the electron is obtained by equating its self electrostatic energy e^2/r to its rest mass energy μc^2. Thus we get

$$\frac{v}{c} = \frac{\hbar c}{e^2} = 137.$$

Hence v is greater than c and it violates the special theory of relativity.

3. You can easily show by matrix multiplication (see margin remark) that

$$\sigma_x\alpha = \beta, \quad \sigma_x\beta = \alpha, \quad \sigma_y\alpha = i\beta, \quad \sigma_y\beta = -i\alpha$$

$$\sigma_z\alpha = \alpha, \quad \sigma_z\beta = -\beta$$

Hence $S^2\alpha = (S_x^2 + S_y^2 + S_z^2)\,\alpha$

$$= \frac{\hbar^2}{4}(\sigma_x^2 + \sigma_y^2 + \sigma_z^2)\,\alpha$$

$$= \frac{\hbar^2}{4}(\sigma_x\,\sigma_x + \sigma_y\,\sigma_y + \sigma_z\,\sigma_z)\,\alpha$$

$$= \frac{\hbar^2}{4}(\sigma_x\beta + i\sigma_y\beta + \sigma_z\alpha)$$

$$= \frac{\hbar^2}{4}[\alpha + \alpha + \alpha]$$

$$= \frac{3\hbar^2}{4}\alpha$$

and $S_z\alpha = \frac{\hbar}{2}\sigma_z\alpha = \frac{\hbar}{2}\alpha$

Similarly, $S^2\beta = (3/4)\,\hbar^2\beta$ and $S_z\beta = -(1/2)\hbar\beta$.

Furthermore

$$\alpha\beta = (1\ 0)\begin{pmatrix}0\\1\end{pmatrix} = 0$$

Hence α and β are orthogonal.

$\sigma_x\alpha$

$= \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}$

$= \begin{pmatrix}0\\1\end{pmatrix} = \beta$

$\sigma_x\beta$

$= \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}$

$= \begin{pmatrix}1\\0\end{pmatrix} = \alpha$

$\sigma_y\alpha = \begin{pmatrix}0 & -i\\i & 0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}$

$= \begin{pmatrix}0\\i\end{pmatrix} = i\beta$

$\sigma_y\beta = \begin{pmatrix}0 & -i\\i & 0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}$

$= \begin{pmatrix}-i\\0\end{pmatrix} = -i\alpha$

$\sigma_z\alpha = \begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}$

$= \begin{pmatrix}1\\0\end{pmatrix} = \alpha$

$\sigma_z\beta = \begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}$

$= \begin{pmatrix}0\\-1\end{pmatrix} = -\beta$

4. For $n = 2$, $l = 0$, 1 and $s = \frac{1}{2}$. Thus, for

$L = 0, \quad J = 1/2$

$L = 1, \quad J = 1/2, 3/2$

So the terms are $2\,^2S_{1/2}$, $2\,^2P_{1/2}$, $2\,^2P_{3/2}$.

For $n = 3$, $l = 0, 1, 2$ and $s = \frac{1}{2}$. Thus, for

$L = 0, \quad J = 1/2$

$L = 1, \quad J = 1/2, 3/2$

$L = 2, \quad J = 3/2, 5/2$

So the terms are $3\,^2S_{1/2}$, $3\,^2P_{1/2}$, $3\,^2P_{3/2}$, $3\,^2D_{3/2}$, $3\,^2D_{5/2}$.

5. $$\Delta E = -\frac{RZ^4\alpha^2}{n^3}[1/2 - 1/1]$$

$$= 13.6 \times \left(\frac{1}{137}\right)^2 \times \frac{1}{27} \times \frac{1}{2}\ \text{eV}$$

$$= 1.34 \times 10^{-5}\ \text{eV}$$

6. These are

$Z = 20$ (Calcium) $1s^2\ 2s^2\ 2p^6\ 3s^2\ 3s^6\ 4s^2$

$Z = 25$ (Manganese) $1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^6\ 4s^2\ 3d^5$

$Z = 31$ (Gallium) $1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^6\ 4s^2\ 3d^{10}\ 4p$

$Z = 37$ (Rubidium) $1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^6\ 4s^2\ 3d^{10}\ 4p^6\ 5s$

7. For He ($Z = 2$), the configuration is $1s^2$.

Since $n_1 = 1$, $l_1 = 0$, $s_1 = \frac{1}{2}$, $n_2 = 1$, $l_2 = 0$, $s_2 = \frac{1}{2}$

Therefore

$$L = 0$$

$S = 1, 0$ whence $J = 0$ for $L = 0$ and $S = 0$ (since $S = 1$ goes only with odd values of L which are non-existent in this case.)

Thus, the spectral term for the ground state of He atom ($1s^2$) is 1^1S_0.

For Li, $Z = 3$ and the configuration is $1s^2\ 2s^1$. Thus, for the electron in the incomplete shell,

$$n = 2,\ l = 0,\ s = \frac{1}{2} \text{ and } J = \frac{1}{2}$$

The ground state spectral term is $2\,^2S_{1/2}$.

For Si ($Z = 14$) the configuration is $1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^2$. For the 2 electrons in the incomplete subshell ($3p$), we have

$$n_1 = 3,\ l_1 = 1,\ s_1 = \frac{1}{2},\ n_2 = 3,\ \ l_2 = 1,\ s_2 = \frac{1}{2}$$

Once again $S = 1, 0$

$$L = 2, 1, 0$$

For $S = 0,\ L = 0,\ J = 0$

$S = 1,\ L = 1,\ J = 2, 1, 0$

$S = 0,\ L = 2,\ J = 2$

Hence the spectral terms are $^1S_0$, $^3P_0$, $^3P_1$, $^3P_2$ and $^1D_2$. The ground state is $3\,^3P_0$.

For Sc ($Z = 21$), the configuration is $1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^6\ 4s^2\ 3d^1$. For the valence electron $n = 3$, $l = 2$, $s = \frac{1}{2}$ and $J = \frac{3}{2}, \frac{5}{2}$. The spectral terms are $3\,^2D_{5/2}$, $3^2D_{3/2}$. According to Hund's rule 2, the ground state is $3\,^2D_{3/2}$ since the subshell is less than half full.

Terminal Questions

1. We have

$$(\mathbf{S} \times \mathbf{S})_x = S_y S_z - S_z S_y$$

$$= \frac{1}{4}\hbar^2 (\sigma_y \sigma_z - \sigma_z \sigma_y)$$

$$= \frac{1}{4}\hbar^2 (2i\ \sigma_x) = i\frac{\hbar^2}{2}\sigma_x = i\hbar S_x. \ (\because \sigma_y \sigma_z = i\sigma_x \text{ and } \sigma_z \sigma_y = -i\sigma_x)$$

Similarly, we can obtain the values for y and z components of $\mathbf{S} \times \mathbf{S}$.

2. For the P state $L = 1$ and it is given that $J = 3$. Hence $S = J - L = 2$.

   For the D state $L = 2$ and $J = 3$. Therefore $S = 3 - 2 = 1$.

3. (i) Allowed because it is according to the selection rules given by Eq. (10.18).

   (ii) Forbidden because $\Delta S \neq 0$.

   (iii) Allowed.

   (iv) Forbidden because $\Delta L = 2$.

UNIT 11 X-RAY SPECTRA

Structure

11.1 Introduction
Objectives
11.2 X-ray Spectra and Selection Rules
11.3 Moseley's Law
11.4 Applications of X-rays
11.5 Summary
11.6 Terminal Questions
11.7 Solutions and Answers

11.1 INTRODUCTION

In Unit 10, you have studied the optical spectra of atoms. You know that optical spectra results when electrons in the outer open subshells make a transition from excited states to the ground state; the photons emitted in this process have wavelengths in the visible region. In this sense, observation of optical spectra supports the theory of atomic shell structure. But optical spectra are not the only source of information about the shell structure of atoms. As we have said in Unit 10, the transitions of electrons in inner shells result in *X*-ray spectra.

Therefore, in the last unit of the block we focus our attention on *X*-rays. Now, in order to generate *X*-rays an anticathode in a vacuum tube is bombarded by high energy electrons. Such a bombardment produces two types of *X*-rays. One of them has a continuous spectrum and is produced due to the deceleration of the charged electrons inside the anticathode. The highest frequency of such *X*-rays is given by E/h, where E is the kinetic energy of the bombarding electrons. The intensity distribution of these *X*-rays as a function of the frequency depends little on the material of the anticathode. This phenomenon is known as **Bremsstrahlung**.

Simultaneously, a second type of *X*-rays are also produced. Their frequencies are characteristic of the material of the anticathode. Hence they are known as **characteristic** *X*-rays. It is the study of characteristic *X*-rays that leads to the determination of the atomic structure. In this unit we are interested in **characteristic** *X*-rays which are discrete in nature and are produced by transitions involving the inner shells of atoms. You may know that the *X*-ray part of the electromagnetic spectrum extends from wavelengths of about 10^{-9} m to wavelengths about 6×10^{-14} m corresponding to frequencies between about 3×10^{17} Hz and 5×10^{23} Hz. The energy of *X*-ray photons lies in the range of 1.2×10^{3} eV to about 2.4×10^{7} eV. These energies correspond to differences in inner shell electron energies. So, in Sec. 11.2 we shall discuss the atomic transitions responsible for *X*-ray spectra and the relevant selection rules.

Moseley used the shell model to analyze *X*-ray spectra of many elements and demonstrated the connection between the atomic number and the frequencies emitted. You will study this relationship known as Moseley's law in Sec. 11.3.

Finally, in the last section of this unit we present a brief discussion of the applications of *X*-rays, in medicine, materials science, astronomy and industry.

Objectives

After studying this unit you should be able to

- determine *X*-ray terms and the allowed atomic transitions which produce characteristic *X*-rays,

- apply Moseley's law,
- discuss applications of *X*-rays.

11.2 X-RAY SPECTRA AND SELECTION RULES

X-ray spectra are associated with complex atoms containing many electrons. Characteristic *X*-rays (Fig. 11.1) are produced when electrons in the inner shells of the atoms make transitions from one state to another. In order to facilitate our study let us first learn the *X*-ray terms. In *X*-ray nomenclature, the inner-most shell of an atom ($n = 1$) is known as the *K* shell. The next shell, i.e., $n = 2$ is termed the *L* shell. However, you know that for $n = 2$, l has two values 0 and 1 and therefore, for $s = 1/2$, we have $j = 1/2$ and 3/2. Thus we have three terms given by $2\,^2S_{1/2}$, $2\,^2P_{1/2}$ and $2\,^2P_{3/2}$. All the three terms have slightly different energies and in *X*-ray nomenclature they are known as L_I, L_{II} and L_{III} subshells. Similarly for $n = 3$ shell, l has three values 0, 1 and 2 and correspondingly $j = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}$. Thus, this shell has five subshells ($3\,^2S_{1/2}$, $3\,^2P_{1/2}$, $3\,^2P_{3/2}$, $3\,^2D_{3/2}$, $3\,^2D_{5/2}$) denoted by M_I to M_V. You may like to determine certain *X*-ray terms yourself.

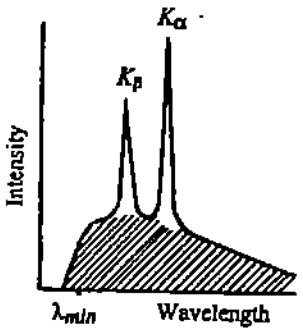


Fig. 11.1: X-ray spectrum.

SAQ 1

Spend 5 min

Show that there will be seven subshells for $n = 4$ shell and give all the spectroscopic terms and their *X*-ray nomenclature.

We have listed some of the *X*-ray subshells with the corresponding values of n, l and j and spectroscopic terms in Table 11.1.

Table 11.1: *X*-ray terms

Subshell	n	l	j	Term
K	1	0	1/2	$1\,^2S_{1/2}$
L_I	2	0	1/2	$2\,^2S_{1/2}$
L_{II}	2	1	1/2	$2\,^2P_{1/2}$
L_{III}	2	1	3/2	$2\,^2P_{3/2}$
M_I	3	0	1/2	$3\,^2S_{1/2}$
M_{II}	3	1	1/2	$3\,^2P_{1/2}$
M_{III}	3	1	3/2	$3\,^2P_{3/2}$
M_{IV}	3	2	3/2	$3\,^2D_{3/2}$
M_V	3	2	5/2	$3\,^2D_{5/2}$

The corresponding energy level diagram is shown in Fig. 11.2. Study both Table 11.1 and Fig. 11.2 before proceeding further.

You may now ask: How are *X*-rays produced? Are all the transitions between the inner energy levels allowed or do there exist certain selection rules as in the case of optical spectra? Let us find the answer to these questions.

In its normal state an atom has two electrons in its *K* shell. Suppose one of the *K* shell electron is taken out of the atom by some process, such as the bombardment of the target in an *X*-ray tube. The collision causes the ejection of an atomic electron from the *K* shell. The atom is singly ionized and has one hole (vacancy) in its *K* shell. Such an ion is in a highly excited state. The ion deexcites when one of the remaining electrons makes a quantum jump from an outer energy state and fills the vacancy left by the

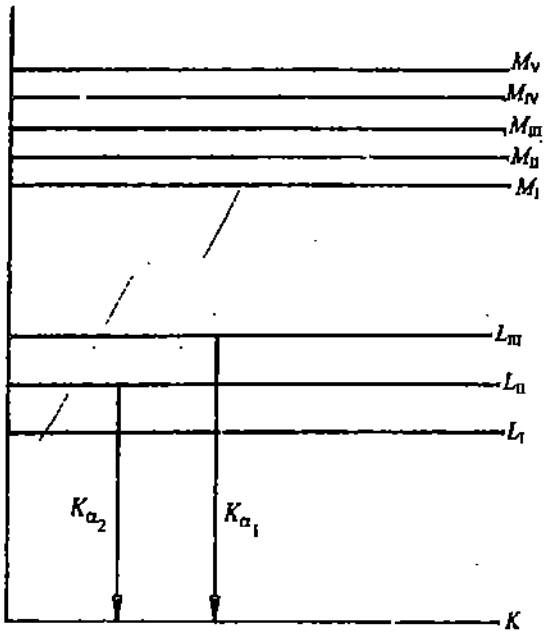


Fig. 11.2: Some X-ray terms and K_α lines.

ejected electron. In the process of stabilisation, a number of transitions of the atomic electrons from the upper (outer) states to lower (inner) states take place till the vacancy is transferred to the upper (outer) most level. Every transition produces characteristic emission lines and some of the lines lie in the X-ray region. Thus X-ray spectra is produced.

The so called K and L series of X-rays result from all such electron transitions to the $n = 1$ and $n = 2$ inner shells of the singly ionized atom. A vacancy in L shell or M shell of heavy atoms also produces X-ray spectra. Like optical spectra, the X-ray spectra is also subjected to the following **selection rules.**

$$\boxed{\Delta l = \pm 1, \qquad \Delta j = 0, \pm 1} \tag{11.1}$$

Hence an L_I shell electron cannot make a transition to K shell but L_{II} to K and L_{III} to K transitions are allowed. The above two transitions give rise to K_{α_1} and K_{α_2} lines, respectively. It is easy to see that the intensity ratio of K_{α_1} and K_{α_2} lines is 2:1. A measurement of the wavelengths of these lines can identify the atom. Similar transitions between K and M shells or between L and M shells also produce characteristic X-ray lines.

You may now like to apply the selection rules to X-ray spectra.

Spend 5 min

SAQ 2

Draw approximate energy levels for L and M shells and show all the allowed transitions.

The first comprehensive study of characteristic X-rays was done by H.G.J. Moseley. He investigated the K and L spectra of many elements in the periodic table. His survey of the elements revealed a pattern in the relationship between emitted frequencies and the atomic number of the atoms. These empirical observations are encapsulated in the form of **Moseley's law.** Let us now study this law.

11.3 MOSELEY'S LAW

As shown in Sec. 11.2, the energy of an X-ray subshell depends upon the quantum numbers n, l, and j.

However, in a crude approximation we may represent the energy value of a shell by the hydrogenic formula after replacing the atomic number Z by $Z-\sigma$, where σ is known as **screening constant.** Thus for n shell we take

$$E_n = -\frac{R\,(Z-\sigma)^2}{n^2} \qquad (11.2)$$

Now a transition between n_2 to n_1 will produce an X-ray line of frequency

$$\nu_{n_2,\,n_1} = \frac{E_{n_2} - E_{n_1}}{h}$$

$$= \frac{R}{h}\left[\frac{(Z-\sigma_{n_2})^2}{n_2^2} + \frac{(Z-\sigma_{n_1})^2}{n_1^2}\right] \qquad (11.3)$$

On the assumption that the screening constants σ_{n_1} and σ_{n_2} have the same value σ we obtain

$$\nu_{n_2,\,n_1} = \frac{R}{h}(Z-\sigma)^2\left[\frac{1}{n_2^2} + \frac{1}{n_1^2}\right] \qquad (11.4)$$

The above equation shows that ν is directly proportional to the square of the atomic number Z. Such a relationship between the frequency ν and the atomic number Z is known as Moseley's law (Fig. 11.3). Moseley was able to change targets in his X-ray tube and observe the frequencies of X-rays for more than 40 elements between aluminium and gold in the periodic table. His experimental results were in agreement with Eq. (11.4). However, you should note that taking σ to be independent of n is not a good approximation. Hence Moseley's law has only a limited validity.

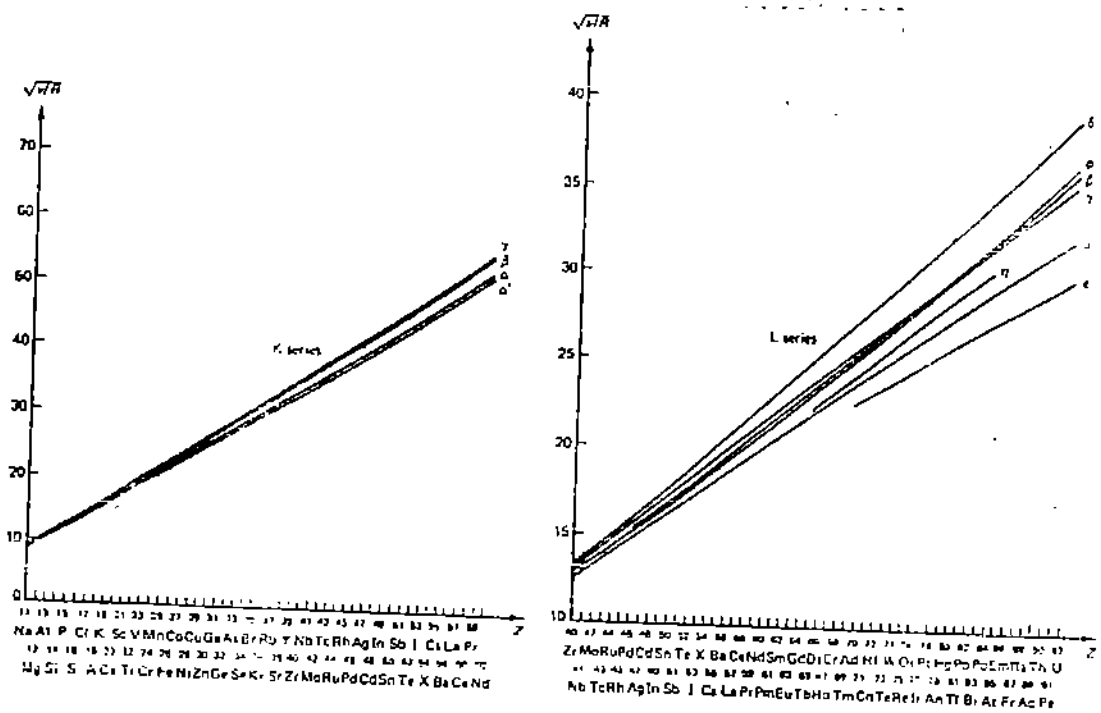


Fig. 11.3 : Moseley's Law.

SAQ 3

Spend 5 min

Use Moseley's law to obtain the frequency of an X-ray line when an L to K transition takes place in a silver atom. Take $\sigma = 3$.

Let us now briefly consider some applications of X-rays, mainly in medicine, industry, materials science and astronomy.

11.4 APPLICATIONS OF X-RAYS

Due to their greater energy, X-rays ionize or dissociate the atoms and molecules of substances through which they pass. The phenomenon of X-ray absorption is an example of the familiar photoelectric effect — the absorption of an X-ray photon excites the atom above its ionization level and ejects a bound electron. A quantum mechanical probability can be introduced to describe the photon-atom interaction. And an absorption cross section can be defined to account for the behaviour of a beam of X-rays incident on the atoms in a sample of matter. We measure absorption in the laboratory by observing the attenuation of an X-ray beam in its passage through a thickness of material. The fractional decrease in intensity $-dI/I$ is related to the element of thickness dx by the proportionality

$$-\frac{dI}{I} = \mu_x\, dx,$$

where the constant μ_x defines the absorption coefficient of the material. This expression is easily integrated to give the intensity as a function of distance x through the sample, starting with incident intensity I_0:

$$\int_{I_0}^{I} \frac{dI}{I} = -\int_0^x \mu_x\, dx$$

or
$$\ln \frac{I}{I_0} = -\mu_x x$$

or
$$I = I_0 e^{-\mu_x x}$$

The absorption coefficient varies with the material and depends on the wavelength of the X-rays. We can use measurements of the attenuation to determine this dependence, and we can then infer the related behaviour of the absorption cross section for the given element.

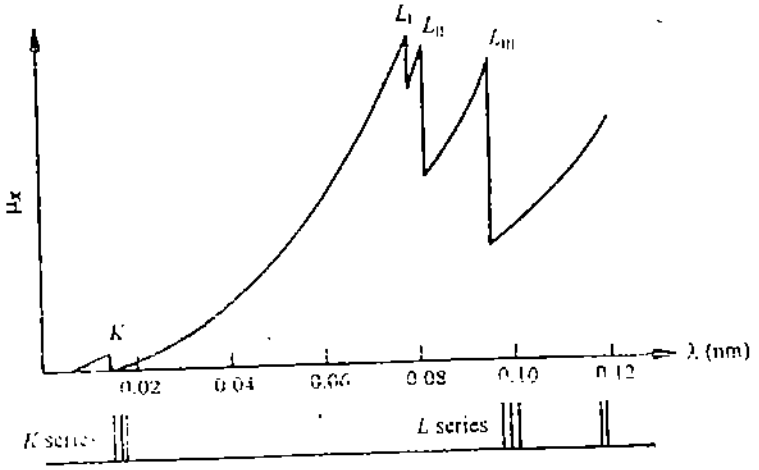


Fig. 11.4: K and L absorption edges of lead. The wavelength thresholds occur where the X-ray photon energy becomes insufficient to eject a K or L shell electron. Emission lines of lead in the K and L series are also shown.

Fig. 11.4 shows a typical graph of μ_x as a function of the wavelength λ. We observe zero absorption in the limit $\lambda \to 0$. This observation tells us that the absorbing medium is transparent to X-rays when the beam energy is very large. We then observe a steady

growth in absorption as the photon energy decreases from large values and as λ increases from zero, until we reach a sharp value of λ where the medium suddenly becomes transparent again. This feature of the graph is called an **absorption edge**, the first of several to appear with increasing λ in the figure. The indicated K absorption edge occurs at wavelength λ_K, where the photon energy is the minimum needed to ionize the atom and leave a vacancy in the K shell. When λ becomes larger than λ_K, the X-ray photon energy becomes too small to free a K-shell electron but remains large enough to eject an electron from an L (or higher) shell. We again observe a steady growth in absorption as the wavelength continues to increase until we reach one of the indicated L absorption edges. The various absorption thresholds along with the characteristic X-ray emission lines provide a signature of the particular atom, and both give an indication of the energy levels of the system. We include the emission lines of the K and L series in the figure so that we can note the positions of these spectral lines relative to the absorption edges.

The property of penetration of materials by X-rays makes them very useful for various applications, particularly in **medical diagnosis**. The relatively greater absorption of X-radiation by bones as compared with tissue results in a fairly 'well-defined' photograph of the bone structure. You must surely have seen such X-ray plates. X-rays are also used for treatment of cancer since they seem to have a tendency to destroy diseased tissue more readily than healthy tissue. But you must remember that X-radiation (in any amount, small or large) does destroy some good tissue. Hence, extreme care must be taken to protect oneself when handling X-rays or during exposure to X-rays.

Medicine

X-rays are used to produce a photographic image of an opaque specimen and provide information about the gross internal structure of any object that they can penetrate. This technique called **radiography** is widely used in diverse areas ranging from medicine to industry. Whether it is to examine the chest of a patient for evidence of tuberculosis, silicosis, heart pathology or embedded foreign objects, of bones in cases of fractures or of arthritis or other bone diseases, X-rays, as you know, are the most handy tool used in medical applications.

X-ray radiography is also used in detecting internal flaws in metal casting or welded joints. A defective casting or welded joints inserted into a bridge or a building can lead to disastrous results. Such metal parts and welds in a pipe are routinely examined by X-rays to observe cracks, inclusions and voids before they are used. X-ray radiography also helps in detecting any crack in the body of ships, cars and aeroplanes. Industrial radiography enables detection of internal physical imperfections in materials such as flaws, segregations, porosities etc. It is often used to visualize inaccessible internal parts of industrial systems to check their location or condition, e.g., in the foundry industry to guarantee the soundness of castings; in the welding of pressure vessels, pipelines, ships and reactor components to guarantee the soundness of welds; in the manufacture of fuel elements for reactors to guarantee their size and soundness; in the solid-propellant and high explosives industry to guarantee the purity of the material being used; in the automotive, aircraft, nuclear, space, oceanic and guided-missile industries, whenever internal soundness is required.

Industry

Among the many objects now examined by radiography are coal, minerals, rubber tyres, golf balls, fabricated objects with internal seals, electrical equipment, printed circuits, fibers, plastics, containers of all kinds, grain, fruit, meats, battery, plates, suitcases, postal packages, and paintings.

Computerized tomography (CT) scan use X-rays to produce images of internal organs of the body and has made a strong impact on medical diagnosis and industrial inspection.

X-rays also find application in **materials science**. You have briefly studied Bragg's diffraction law in the context of wave-particle duality. It was first discovered for diffraction of X-rays from the surface layers of a crystal. With a known crystal of lattice spacing d, we can measure the wavelength of the radiation; and with a known wavelength we can measure the lattice spacing d. X-ray diffraction has been developed into a standard technique for analysing crystal structure and its defects. X-ray diffraction and crystallography has led to extensive study of crystal structures, their atomic arrangements and electron distribution, etc. Chemical elemental analysis of solids, liquids and thin films uses X-rays as a non-destructive physical method. X-ray microscopy is used to obtain quantitative chemical information about samples as small

Material Science

Astronomy

X-rays are being widely used in astronomy for exploring the universe. All types of astronomical objects, from stars to galaxies and quasars emit X-rays which can be detected by specially designed *X-ray telescopes* placed in rockets, satellites and space probes above the Earth's atmosphere. These have led to the discovery of new stellar objects and yielded information about the distribution of stellar objects in the sky, time-evolution of galaxies and supernova remnants, and later stages of a star's life when it metamorphoses into a collapsed object like a white dwarf, neutron star or black hole.

With this brief discussion of X-ray applications, we come to the end of this unit. Let us now summarise what you have studied in this unit.

11.5 SUMMARY

- X-rays are produced in two ways

  (i) when high speed electrons penetrate atoms, they decelerate as they pass close to the atomic nuclei and produce continuous X-radiation spectrum. This is often referred to as **"Bremsstrahlung"**.

  (ii) In another process, these electrons remove electrons from the inner shells by collision. The transitions of atomic electrons from outer shells to vacant inner shells result in **characteristic X-rays.**

- The **selection rules** for atomic transitions that yield a characteristic X-ray spectrum are

  $$\Delta l = \pm 1,\ \Delta j = 0, \pm 1$$

- The relationship between the characteristic X-ray frequencies emitted by an atom and its atomic number are given by **Moseley's law.**

- X-rays find several applications in medicine, industry, astronomy and materials science.

11.6 TERMINAL QUESTIONS

Spend 15 min

1. At what potential difference must an X-ray tube operate to produce X-rays with a minimum wavelength of 1Å?

2. X-rays from a certain cobalt target tube are composed of the strong K-series of cobalt and weak K lines due to impurities. The wavelengths of the K_α lines are 1.785 Å for cobalt and 1.537 Å and 2.285 Å for the impurities. Using Moseley's law, calculate the atomic numbers of the impurities and identify the elements.

11.7 SOLUTIONS AND ANSWERS

Self-Assessment Questions

1. For $n = 4$, $l = 0, 1, 2, 3$, and $s = 1/2$

   Thus for $l = 0$, $j = 1/2$

   $l = 1$, $j = 1/2, 3/2$

   $l = 2$, $j = 3/2, 5/2$

   $l = 3$, $j = 5/2, 7/2$

Hence for $n = 4$, the subshells and corresponding terms are

N_I	N_{II}	N_{III}	N_{IV}	N_V	N_{VI}	N_{VII}
$4\,^2S_{1/2}$	$4\,^2P_{1/2}$	$4\,^2P_{3/2}$	$4\,^2D_{3/2}$	$4\,^2D_{5/2}$	$4\,^2F_{5/2}$	$4\,^2F_{7/2}$

2. See Fig. 11.5.

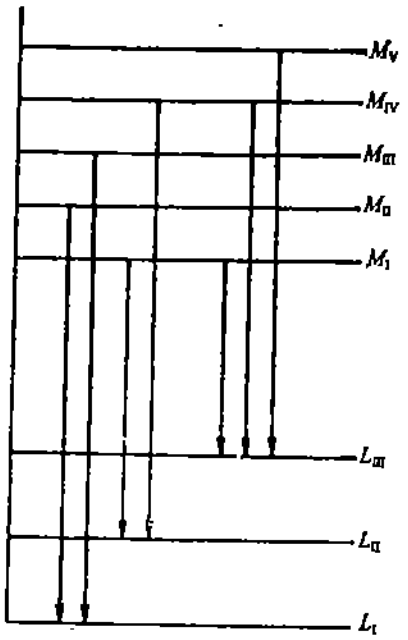


Fig.11.5: Energy levels for L and M shells, and allowed transitions between L and M shells.

3. $\nu = \dfrac{13.6 \times 1.602 \times 10^{-19}\ \text{J}}{6.626 \times 10^{-34}\ \text{Js}} \times (47-3)^2 \left(-\dfrac{1}{4}+1\right)$ (∵ for silver $Z = 47$) $n_1 = 1$ and $n_2 = 2$

$= 4.7 \times 10^{18}$ Hz

Terminal Questions

1. The energy is given by

$$E = h\nu = \frac{hc}{\lambda}$$

$$= \frac{6.626 \times 10^{-34}\ \text{Js} \times 3 \times 10^{8}\ \text{ms}^{-1}}{10^{-10}\ \text{m}}$$

$$= 2.0 \times 10^{-15}\ \text{J}$$

And the potential difference is

$$V = \frac{E}{e} = \frac{2.0 \times 10^{-15}\ \text{J}}{1.6 \times 10^{-19}\ \text{C}}$$

$$= 1.25 \times 10^{4}\ \text{V}$$

2. Using Moseley's law we can write

$$\nu = \frac{c}{\lambda} = \frac{R}{h}(Z-\sigma)^2 \left(-\frac{1}{n_2^2}+\frac{1}{n_1^2}\right)$$

For K series $n_1 = 1$, $n_2 = 2$. For cobalt $Z = 27$ and applying Moseley's law we can write

$$\frac{3 \times 10^8 \text{ ms}^{-1}}{1.785 \times 10^{-10} \text{ m}} = \frac{13.6 \times 1.6 \times 10^{-19} \text{J}}{6.626 \times 10^{-34} \text{ Js}} \times (27 - \sigma)^2 \times \left(1 - \frac{1}{4}\right)$$

or

$$(27 - \sigma)^2 = \frac{3 \times 10^8 \times 6.626 \times 10^{-34} \times 4}{1.785 \times 10^{-10} \times 13.6 \times 1.6 \times 10^{-19} \times 3}$$

$$= 680$$

or

$$(27 - \sigma) \simeq 26$$

$$\therefore \sigma = 1$$

(i) Now for the first impurity, $\hat{\lambda} = 1.537$ Å. Therefore, from Moseley's law we have

$$\frac{3 \times 10^8 \text{ ms}^{-1}}{1.537 \times 10^{-10} \text{ m}} = \frac{13.6 \times 1.602 \times 10^{-19} \text{J}}{6.626 \times 10^{-34} \text{ Js}} (Z - 1)^2 \times \frac{3}{4}$$

or

$$(Z - 1)^2 = 790$$

or

$$(Z - 1) \simeq 28$$

and

$Z = 29$, so the impurity is copper.

(iii) for the second impurity, we have $\lambda = 2.285$ Å. Thus

$$(Z - 1)^2 = \frac{3 \times 10^8 \text{ ms}^{-1} \times 6.626 \times 10^{-34} \text{ Js}}{2.285 \times 10^{-10} \times 13.6 \times 1.6 \times 10^{-19} \text{ J}} \times \frac{4}{3}$$

$$= 530$$

or

$$(Z - 1) = 23$$

and

$Z = 24$, so the impurity is chromium.

Table of fundamental constants

Quantity	*Symbol*	*Value*
Planck's constant	h	6.62618×10^{-34} J s
	$\hbar = \frac{h}{2\pi}$	1.05459×10^{-34} J s
Velocity of light in vacuum	c	2.99792×10^{8} m s^{-1}
Elementary charge (absolute value of electron charge)	e	1.60219×10^{-19} C
Permeability of free space	μ_0	$4\pi \times 10^{-7}$ H m^{-1} $= 1.256\,64 \times 10^{-6}$ H m^{-1}
Permittivity of free space	$\varepsilon_0 = \frac{1}{\mu_0 c^2}$	$8.854\,19 \times 10^{-12}$ F m^{-1}
Gravitational constant	G	6.672×10^{-11} N m^2 kg^{-2}
Fine structure constant	$\alpha = \frac{e^2}{4\pi\varepsilon_0 \hbar c}$	$\frac{1}{137.036} = 7.297\,35 \times 10^{-3}$
Avogadro's number	N_A	$6.022\,05 \times 10^{23}$ mol^{-1}
Faraday's constant	$F = N_A e$	$9.648\,46 \times 10^{4}$ C mol^{-1}
Boltzmann's constant	k	$1.380\,66 \times 10^{-23}$ J K^{-1}
Gas constant	$R = N_A k$	$8.314\,41$ J mol^{-1} K^{-1}
Atomic mass unit	a.m.u. $= \frac{1}{12} M_{^{12}C}$	$1.660\,57 \times 10^{-27}$ kg
Electron mass	m or m_e	$9.109\,53 \times 10^{-31}$ kg $= 5.485\,80 \times 10^{-4}$ a.m.u.
Proton mass	M_p	$1.672\,65 \times 10^{-27}$ kg $= 1.007\,276$ a.m.u.
Neutron mass	M_n	$1.674\,92 \times 10^{-27}$ kg $= 1.008\,665$ a.m.u.
Ratio of proton to electron mass	M_p/m_e	1836.15
Electron charge to mass ratio	$\lvert e \rvert / m_e$	$1.758\,80 \times 10^{11}$ C kg^{-1}
Classical radius of electron	$r_0 = \frac{e^2}{4\pi\varepsilon_0 m c^2}$	$2.817\,84 \times 10^{-15}$ m
Bohr radius for atomic hydrogen (with infinite nuclear mass)	$a_0 = \frac{4\pi\varepsilon_0 \hbar^2}{m e^2}$	5.29177×10^{-11} m
Rydberg's constant for infinite nuclear mass	$R_\infty = \frac{m c^2}{8\varepsilon_0^2 h^3 c} = \frac{\alpha}{4\pi a_0}$	$1.097\,37 \times 10^{7}$ m^{-1}
Rydberg's constant for atomic hydrogen	R_H	$1.096\,78 \times 10^{7}$ m^{-1}
Bohr magneton	$\mu_B = \frac{e\hbar}{2m}$	9.27408×10^{-24} J T^{-1}
Nuclear magneton	$\mu_N = \frac{e\hbar}{2M_p}$	5.05082×10^{-27} J T^{-1}

UTTAR PRADESH
RAJARSHI TANDON OPEN UNIVERSITY

Block

4

NUCLEAR PHYSICS

UNIT 12

Radioactivity 5

UNIT 13

The Nucleus 23

UNIT 14

Applied Nuclear Science 43

UNIT 15

Elementary Particles 63

BLOCK INTRODUCTION

In Block 3, you have studied about the atom and its structure mainly in terms of the dynamics of electrons surrounding the atomic nucleus, which is manifested in atomic spectra. It is now time for us to delve deeper into the secrets of the atom and enter the domain of the atomic nucleus. This forms the subject matter of this block wherein we intend to provide you with a bird's eye view of the physics of the nucleus.

The march towards our understanding the physics of nucleus can be said to have begun about a hundred years ago with the chance discovery of the phenomenon of radioactivity. This discovery provided researchers convenient tools to probe the structure of matter and led to discoveries like artificial transmutation production of transuranic elements and radio isotopes which find wide use in medical diagnosis and therapy, agriculture, radiocarbon dating etc. You will study the phenomenon of radioactivity and the physics underlying it in Unit 12.

The alpha particles from radioactive nuclei were used by Rutherford to 'see' the atomic structure. As you are aware, these classic scattering experiments of alpha particles by atoms led to the discovery of the nuclear model of atom. This opened the flood gate of investigations into the atomic nucleus – its composition and structure, its properties and the forces of interaction between its constituents. In this process, answer to the question of nuclear stability was also sought. In Unit 13, we present the current understanding covering these aspects.

Science and its applications are bound together as the fruit to the tree which bears it. This statement of Louis Pastem's beautifully captures the spirit in which we teach physics in our courses. Thus having discussed the fundamentals of nuclear physics in the first two units of this block, in Unit 14 we focus our attention on the peaceful uses of nucleus – its applications in the field of nuclear energy, hydrology, medicine, agriculture and industry.

Investigations into the nuclear structure have led not only to the discovery of newer, more elementary particles but have also provided deep insights into fundamental questions confronting human consciousness today – in particular, the questions of the origin of universe and its evolution. Therefore, in the last unit (Unit 15) of this block (and this course), we take you on a discovery trip into the world of elementary particles. In the process we also familiarise you with the tools and techniques (viz. particle accelerators and detectors) which have made these insights in the nuclear domain a reality.

One final word about how to study the block. At the risk of repeating ourselves, we would like to remind you that while studying the units you should answer the SAQs and terminal questions on your own. Donot give into the temptation of looking up the answers before attempting the exercises. The study time envisaged by us on the basis of our experiences is about 5h, 6h, 7h, and 8h for Units 12, 13, 14, 15 respectively. The actual time you take may vary depending on your previous knowledge

We hope that you will enjoy this journey into the world of atomic nucleus as much as we have.

We wish you good luck!

UNIT 12 RADIOACTIVITY

Structure

12.1 Introduction
Objectives
12.2 Discovery of Radioactivity and Preliminary Studies
12.3 Radioactive Decay
12.4 Growth and Decay of Radioactivity
12.5 Successive Radioactive Transformations
Radioactive Equilibrium
12.6 Summary
12.7 Terminal Questions
12.8 Solutions and Answers

12.1 INTRODUCTION

Towards the end of nineteenth century, physicists thought that the era of exciting discoveries in physics was over. However, the chance discovery of radioactivity by Becquerel in 1896 overcame this pessimism and opened the flood gates for new discoveries. This became possible because spontaneously emitted radiations — alpha, beta and gamma-rays — could be used as convenient tools to probe matter. For instance, the alpha- particles from radioactive nuclei were used by Rutherford to propose the nuclear model of atom. (You will learn about it in the next unit.) It further led to discoveries of artificial transmutation and production of transuranic elements as well as radioisotopes, which find wide use in medical diagnosis and therapy, research, agriculture, carbon dating of archaeological specimen etc. The studies of beta decay led to the discovery of neutrino. In short, discovery of radioactivity acted as precursor of fundamental developments in nuclear physics in the early part of this century.

In Sec. 12.3 we have discussed the theory of radioactive decay. We have applied this theory to explain the growth and decay of radioactivity in a given radioactive sample in Sec. 12.4. Successive radioactive disintegrations and the condition of radioactive equilibrium amongst the different members of a radioactive series is discussed in Sec. 12.5.

Objectives

After going through this unit, you should be able to

- identify the three types of radioactive radiations
- formulate the laws of radioactive transformation
- compute the half-life and disintegration constant of a radioactive substance
- explain the growth and decay of radioactivity in a given sample
- explain the radioactive equilibrium amongst the different members of a radioactive series and
- list different elements belonging to naturally occurring radioactive series.

12.2 DISCOVERY OF RADIOACTIVITY AND PRELIMINARY STUDIES

The story of discovery of radioactivity is very interesting. In 1896, Henri Becquerel was working on the phenomenon of fluorescence in which certain substances emit visible light when they are exposed to ultra-violet radiations, say from the sun. In one of the drawers of his desk, Henry Becquerel had kept a collection of various minerals, which also included uranium salts, along with several cardboard boxes of photographic plates wrapped with thick black paper. A few days later he used one of the boxes of photographic plates. When he developed the plate, he was amazed to observe that it was heavily fogged. He tried other plates and found them also to be exposed. This puzzled him because all boxes were unused. Can you guess what had affected these photographic plates?

Becquerel conjectured very rightly that uranium salts, placed beside the photographic plates must have emitted some new type of radiation(s). After extensive experimentation, Becquerel also proved that these radiations, if passed through some gas, make it conducting by virtue of their ionising power. The emission of ionising and penetrating radiation(s) from uranium was named radioactivity.

Following an exhaustive study, Madame Marie Curie found evidence of radioactivity in elements like thorium, polonium and radium as well. She also obtained convincing evidences that radioactivity is a nuclear property of the element concerned and remains unaffected by physical or chemical changes.

By studying the ionising and penetrating power of these radiations, Rutherford established the existence of two distinct components, α-rays and β-rays. α -rays are more easily absorbed in matter compared to β-rays but have comrelatively greater ionising power. However, the penetrating power of β-rays is about 100 times more than that of α-rays. The existence of the third component, called gamma rays, which are much more penetrating than the other two, was established by P. Villars. By subjecting these radiations to a magnetic field, it was also established that the gamma rays are electrically neutral, whereas the alpha-rays are positively charged and the beta-rays are negatively charged particles. We now know that α- rays can be identified with helium nuclei and β-rays with electrons. Moreover, α-particles affect photographic films and excite flourescence in many substances. Also, they produce phosphorescence and their speed is of the order of 1/100th of the speed of light. On the other hand, the speed of β-particles is about 1/10th of the speed of light. Gamma rays (γ) are very short wavelength electromagnetic radiations emitted from nuclei of radioactive substances and travel with speed of light. Further studies revealed that radioactive emanations have energies in the MeV range. To enable you to comprehend these concepts, we give below some examples of radioactive decays which involve emission of α and β-rays:

$$^{238}U \rightarrow {}^{234}Th + \alpha \qquad (4.20\ MeV)$$

$$^{226}Ra \rightarrow {}^{222}Rn + \alpha \qquad (4.78\ MeV)$$

$$^{60}Co \rightarrow {}^{60}Ni + \beta^- \qquad (0.318\ MeV)$$

$$^{90}Sr \rightarrow {}^{90}Y + \beta^- \qquad (0.546\ MeV)$$

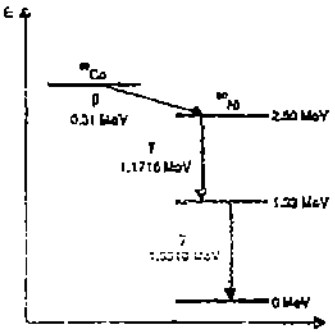


Fig. 12.1 : Energy level diagram for the ^{60}Ni nucleus formed in the decay $^{60}Co \rightarrow {}^{60}Ni + \beta^-$

(The β-decay is accompanied by the emission of neutrinos. But these are very difficult to detect and we will discuss about them later.) You will agree that all these reactions involve transmutation of one element to another: uranium decays to thorium, radium to ruthenium, cobalt to nickel and strontium to ytrium. In many cases of α and β decays, the daughter nucleus emerges in an excited state and subsequently undergoes a transition to a lower/ground state by emitting a γ-ray. In Fig. 12.1 we have shown the energy-level diagram for the ^{60}Ni nuclei formed in the β-decay of ^{60}Co. A nucleus in an excited state can also make a transition to a lower state by transferring its excitation energy to an atomic electron. Such an energy transfer is called **internal conversion**. The electron gaining this energy gets ejected from the atom as a β-ray.

12.3 RADIOACTIVE DECAY

You now know that when a radioactive element disintegrates spontaneously, its nucleus emits either an α or a β-particle. The nucleus of new element formed could also be unstable. The first decay is then succeeded by another and another, ... , resulting in a sequential series. Refer to the decay of ^{238}U shown below:

$$^{238}U \xrightarrow{\alpha} {}^{234}Th \xrightarrow{\beta} {}^{234}Pa \xrightarrow{\beta} {}^{234}U \xrightarrow{\alpha} {}^{230}Th \xrightarrow{\alpha} {}^{226}Ra \ldots$$

The decays continue until a stable isotope is reached. In this series, the final stable decay product is lead. ^{238}U is called the **parent nucleus** and the resultant nucleus ^{234}Th is called the **daughter nucleus**. You may now like to know other such series. We have four such series in all: thorium series, neptunium series, uranium series and actinium series. The mass numbers of the members of these series are given by $4n$, $4n+1$, $4n+2$, $4n+3$ respectively, where n is an integer.

In all the radioactive series, a parent radioactive element of large atomic number and very long half-life gives rise to a series of radioactive elements as a result of successive emissions of α or β-particles. We can express the transformations by expressions of the form:

$$^{A}_{Z}X \rightarrow {}^{A-4}_{Z-2}Y + \alpha$$

$$^{A}_{Z}P \rightarrow {}^{A}_{Z+1}Q + \beta^{-}$$

The decay products may themselves be radioactive and decay by emitting α or β-particles. The successive radioactive transformations continue, until we reach a stable isotope of lead.

The members of the uranium series are listed in Table 12.1, together with the half-life and mode of disintegration. Table 12.2 gives the corresponding information about the actinium series, while Table 12.3 refers to the thorium series. In the uranium series, RaA, RaC' and RaF all have atomic number 84 and are isotopes of polonium. Similarly, RaB, RaD and RaG all have atomic number 82 and are isotopes of lead.

Table 12.1: The Uranium Series ($A = 4n + 2$)

Radioactive species	Chemical symbol	Z	A	Half-life	Particles emitted
Uranium I	UI	92	238	4.5×10^{9}yr	α
Uranium X_1	UX_1	90	234	24.1d	β
Uranium X_2	UX_2	91	234	1.18 m	β
Uranium Z	UZ	91	234	6.7 h	β
Uranium II	UII	92	234	2.5×10^{5}yr	α
Ionium	Io	90	230	8.0×10^{4}yr	α
Radium	Ra	88	226	1620 yr	α
Radon	Rn	86	222	3.82 d	α
Radium A	RaA	84	218	3.05 min	α, β
Radium B	RaB	82	214	26.8 min	β
Astatine-218	At-218	85	218	2 s	α
Radium C	RaC	83	214	19.7 min	β, α
Radium C'	RaC'	84	214	1.64×10^{-4}s	α
Radium C''	RaC''	81	210	1.32 min	β
Radium D	RaD	82	210	19.4 yr	β
Radium E	RaE	83	210	5.0 d	β
Radium F	RaF	84	210	138.3 d	α
Radium G	RaG	82	206	—	stable

—UX_1 exhibits a branching effect; 99.65 percent of UX_1 atoms emit β-particles to form UX_2, and 0.35 percent of the UX_1 atoms emit β-particles to form UZ. Both UX_2 and UZ have the same mass number 234 and the same atomic number 91, but their nuclear energy levels are different. Such pairs of radioactive species are known as nuclear isomers.

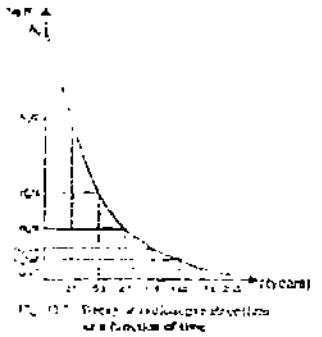
We note that a few isotopes in each of the radioactive series have two alternative modes of decay: these nuclides decay both by α-emission and by β- emission. This type of disintegration is called **branching decay** and is characterised by a definite branching ratio for each mode of decay. For example, we find branching decays for RaA and RaC in the uranium series, for Ac, AcA and AcC in the actinium series, and for ThA and ThC in the thorium series. In most cases, one mode of decay is more probable than the other. Thus RaA and AcA decay almost entirely by α-emission (branching ratio > 99%); only a small fraction (< 1%) of the atoms disintegrate by β-emission. On the other hand, RaC, Ac and AcC decay almost entirely by β-emission, with < 1% of the atoms decaying by α-emission. The only exception is ThC, which has a branching ratio 66.3% for α -disintegration and 33.7% for β-disintegration. The neptunium series starts with plutonium (Pu) and the final stable end product is an isotope of bismuth (Bi).

Table 12.3 : The Actinium Series ($A = 4n + 3$)

Radioactive species	Symbol	Z	A	Half-life	Particles emitted
Actinouranium	AcU	92	235	7.1×10^8y	α
Uranium	Y UY	90	231	25.6 h	β
Protoactinium	Pa	91	231	3.4×10^4yr	α
Actinium	Ac	89	227	22yr	β, α
Radioactinium	RdAc	90	227	18.2d	α
Actinium K	AcK	87	223	22min	β, α
Actinium X	AcX	88	223	11.68 d	α
Actatin 219	At 219	85	219	0.9 min	α, β
Actinon	An	86	219	3.92 s	α
Bismuth 215	Bi-215	83	215	8 min	β
Actinium A	AcA	84	215	1.83×10^{-3}s	α, β
Actinium B	AcB	82	211	36.1 min	β
Astatine 215	At 215	85	215	10^{-4} s	α
Actinium C	AcC	83	211	2.16 min	β, α
Actinium C′	AcC′	84	211	0.52 s	α
Actinium C″	AcC″	81	207	4.8 min	β
Actinium D	AcD	82	207	—	stable

Table 12.3 : The Thorium Series ($A = 4n$)

Radioactive species	Symbol	Z	A	Half-life	Particles emitted
Thorium	Th	90	232	1.39×10^{10}yr	α
Mesothorium I	$MsTh_1$	88	228	6.7 yr	β
Mesothorium II	$MsTh_2$	89	228	6.13 h	β
Radiothorium	RdTh	90	228	1.9 yr	α
Thorium X	ThX	88	224	3.64 d	α
Thoron	Tn	86	220	51.5 s	α
Thorium A	ThA	84	216	0.16 s	α, β
Thorium B	ThB	82	212	10.6 h	β
Astatine-216	At-216	85	216	3×10^{-4} s	α
Thorium C	ThC	83	212	60.5 min	β, α
Thorium C′	ThC′	84	212	3.0×10^{-7} s	α
Thorium C″	ThC″	81	208	3.1 min	β
Thorium D	ThD	82	208	—	stable



We now know that if we have a given amount of radioisotope, it will gradually decrease with time. Measurements show that the quantitative law describing the decay process is very simple. To understand this, refer to Fig. 12.2. It shows a plot of amount of radioactive strontium (on the logarithmic scale) as a function of time. You will note that

- it takes 29 years for one-half of the initial amount of strontium to decay
- during the next 29 years, one-half of the remainder strontium decays, i.e. we are left with: $\frac{1}{2}\left(\frac{N_0}{2}\right) = \frac{N_0}{4}$ nuclei.

Hence, the fraction of parent radioactive substance (left) after 29, 58, 87 years will be

$\frac{1}{2}, \frac{1}{4}, \frac{1}{8}$ of the initial amount. Do you recognise this series? It forms a geometric progression with $r = \frac{1}{2}$. So if $N(t)$ denotes the number of strontium nuclei surviving decay at time t and N_0 is the number at $t = 0$, then we can write

$$N(t) = N_0 \left(\frac{1}{2}\right)^{t/(29\ yr)} \tag{12.1}$$

The time required for one-half of the parent nuclei (material) to decay is called its **half-life**. We will denote it by the symbol $T_{1/2}$. The half-lives of some important radioisotopes are given in Table 12.4. You will note that values of $T_{1/2}$ show a very wide range; from 4.5×10^9 yr for ^{238}U to 3×10^{-9}s for ^{212}Po.

Table 12.4 : Half-lives of some radioisotopes

Radioisotope	$T_{1/2}$
^{14}C	5730 yr
^{40}K	1.3×10^9 yr
^{60}Co	5.24 yr
^{90}Sr	28.5 yr
^{131}I	8.05 d
^{212}Po	3×10^{-9}s
^{238}U	4.5×10^9 yr

In terms of the half-life, we can rewrite Eq. (12.1) as

$$N(t) = N_0 \left(\frac{1}{2}\right)^{t/T_{1/2}} \tag{12.2}$$

Let us pause for a while and ask: What is the physical implication of this result? This relation tells us how the given quantity of a radioactive sample disintegrates as time passes. You may also like to know: Does this formula hold only for times $t = 0, T_{1/2}, 2T_{1/2}$, ... etc.? The answer to this question is: Eq. (12.2) holds for all times.

It may be emphasized here that the general appearance of the decay curve is essentially the same for all radioactive elements. However, each element takes its own characteristic time to decay.

Using the identity $2 = e^{\ln 2} = e^{0.693}$, we can rewrite Eq. (12.2) into a more convenient form as

$$N(t) = N_0 \exp(-\ln 2\, t/T_{1/2})$$
$$= N_0 \exp(-t/\tau) \tag{12.3}$$

where $\tau = \dfrac{T_{1/2}}{\ln 2}$.

Eq. (12.3) expresses the **law of radioactive decay** mathematically. It shows that the number of atoms in a radioactive sample decreases exponentially with time with a characteristic time constant τ. Let us now find out the number of nuclei decaying in any time interval dt. From Eq. (12.3) it readily follows that

$$dN = -\frac{N_0}{\tau} e^{-t/\tau} dt$$

The negative sign signifies that the number of nuclei decreases with time due to their continuous disintegration.

By definition, the average life time, $\bar{t}$, is

$$\bar{t} = \frac{\int_0^\infty t\,|dN|}{\int_0^\infty |dN|} = \frac{1}{N_0}\int_0^\infty t\,|dN|$$

$$= \frac{1}{N_0}\int_0^\infty \frac{t}{\tau} N_0\, e^{-t/\tau}\, dt$$

$$= \frac{1}{\tau}\int_0^\infty t\, e^{-t/\tau}\, dt$$

Using the method of integration by parts, we obtain

$$\bar{t} = \frac{1}{\tau}\left[-\tau e^{-t/\tau}\Big|_0^\infty + \tau\int_0^\infty e^{-t/\tau}\, dt\right]$$

The first term vanishes at both the limits. Therefore, this expression simplifies to

$$\bar{t} = \int_0^\infty e^{-t/\tau}\, dt = \left[-\tau\, e^{-t/\tau}\right]_0^\infty \qquad (12.4)$$

That is, τ is mean life of radioactive nuclei. It may be pointed out here that like $T_{1/2}$, τ will also vary over a very wide range. The measurement techniques over such a wide spread of time are bound to differ vastly. But the law of radioactive decay is common to all radioactive decay processes.

We will now illustrate these concepts through solved examples.

Example 1

The half-life of radon is 3.8 days. After how many days will only 5% of radon be left over?

Solution

We know that $T_{1/2}$ = 3.8 days. Therefore

$$\tau = \frac{T_{1/2}}{\ln 2} = \frac{T_{1/2}}{0.693} = \frac{3.8 \text{ days}}{0.693} = 5.48 \text{ days}$$

We are required to calculate the number of days in which only 5% radon is left. Therefore, we can write $\frac{N}{N_0} = 0.05$.

From Eq. (12.3) it readily follows that

$$\frac{N}{N_0} = 0.05 = \exp\left(-\frac{t}{5.48 \text{ days}}\right)$$

We can write it as

$$\exp\left(\frac{t}{5.48 \text{ days}}\right) = 20$$

Taking natural logarithm of both sides, we find that

$$\frac{t}{5.48 \text{ days}} = \ln 20 = 2.303 \log_{10} 20 = 2.303 \times 1.3010 = 2.996$$

Hence

$$t = 2.996 \times 5.48 \text{ days} = 16.42 \text{ days}$$

Example 2

Due to accident in a research laboratory a radioactive element got spread inside a room. As a result, the level of radiation became 50 times the permissible level for normal occupancy of the room. After how many days the room would be safe for occupation? The half-life of the radioactive substance is 30 days.

Solution

Here $\frac{N}{N_0} = \frac{1}{50}$ and $T_{1/2} = 30$ days.

so that $\tau = \frac{T_{1/2}}{0.693} = \frac{30 \text{ days}}{0.693} = 43.3 \text{ days}$

$\frac{N}{N_0} = \frac{1}{50} = \exp(-t/43.3 \text{ days})$. This can be rearranged to give

$$\frac{t}{43.3 \text{ days}} = \ln 50 = 3.912$$

and

$$t = 3.912 \times 43.3 \text{ days} = 169.4 \text{ days}$$

Example 3

A sample of pitchblende has a lead-uranium weight ratio of 9/40. Calculate the age of the mineral. The half-life of uranium is 4.5×10^9yr. The atomic weights of lead and uranium are 206.0 and 238.4, respectively.

Solution

Since the weight ratio of lead to uranium is 9/40, we can say that if there were 9 kg of lead, the amount of uranium would be 40 kg. The number of atoms in 9 kg of lead $= \frac{9}{206} \times 6 \times 10^{26} = 0.262 \times 10^{26}$atoms. Similarly, the number of atoms in 40 kg of uranium $= \frac{40}{238.4} \times 6 \times 10^{26} = 1.007 \times 10^{26}$.

$\therefore$ Total number of uranium atoms in the beginning $= 1.269 \times 10^{26}$atoms

Since $T_{1/2} = 4.5 \times 10^9$ yr, $\tau = \frac{T_{1/2}}{\ln 2} = \frac{4.5 \times 10^9 \text{ yr}}{0.693} = 6.494 \times 10^9$ yr

From Eq. (12.3) we know that

$$N = N_0 e^{-t/\tau}$$

so that

$$\frac{t}{\tau} = \ln\left(\frac{N_0}{N}\right)$$

or

$$t = \tau \ln\left(\frac{N_0}{N}\right)$$

$$= (6.494 \times 10^9 \text{ yr}) \ln$$

$$= (6.494 \times 10^9 \text{ yr}) \times 0.2296$$

$$= 1.49 \times 10^9 \text{ yr}$$

You may now like to solve an SAQ.

SAQ 1

Spend 10 min

The mean life of a radioactive element is 14.43 months. Calculate the time required for 75% of the element to decay.

In practice, we are more interested in the decay rate of the material rather than its amount, since it determines the rate of emission of α, β or γ -rays. (Moreover, this information can also be used to estimate the age of any specimen.) To this end, we note from Eq. (12.3) that

$$\frac{dN}{dt} = -\frac{N_0}{\tau} \exp(-t/\tau) = -\frac{N}{\tau} \qquad (12.5)$$

That is the decay rate at any time is proportional to the amount of radioactive material present at that instant.

Alternatively, we can write Eq. (12.5) as

$$\frac{dN}{dt} = -\lambda N \tag{12.6}$$

where λ is constant of proportionality and is called *decay constant*. It is characteristic of a particular radioactive element or decay process. In terms of λ, the law of radioactive decay can be expressed as

$$N = N_0 \exp(-\lambda t) \tag{12.7}$$

If you compare Eqs. (12.3) and (12.7), you will find that $\lambda = 1/\tau$.

The activity A of a given radioactive substance is defined as the number of atoms disintegrating per unit time. Mathematically

$$A = \left| \frac{dN}{dt} \right| \tag{12.8}$$

From Eqs. (12.6) and (12.8), we have

$$A = \lambda N \tag{12.9}$$

i.e., the activity of a given radioactive substance is directly proportional to the number of radioactive atoms present. If A_0 is the initial activity of the source at $t = 0$, then we have

$$\frac{A}{A_0} = \frac{N}{N_0} = e^{-\lambda t} \tag{12.10}$$

The quantity $\frac{A}{A_0}$ is defined as the relative activity and is a measure of the radioactivity of a given source. From this result you will note that even the relative activity of a given radioactive substance decays exponentially with time.

The most natural way to express decay rate is in disintegrations per second. But the activities encountered in practice are usually so high that a larger unit, the curie, abbreviated as Ci, is more often used. Initially the curie was defined as the activity of 1g of radium but its value kept changing as improvements in measuring techniques were accomplished. The curie is now defined as

$$1 \text{ curie} = 1 \text{ Ci} = 3.7 \times 10^{10} \text{ disintegrations per second}$$

In the SI system, decay rate is expressed in becquerel:

$$1 \text{ becquerel} = 1 \text{ Bq} = 1 \text{ disintegration per second}$$

You may have seen children with swollen thyroid gland. Do you know that to scan the thyroid, radioisotope ^{131}I is used? Let us now calculate its decay rate.

The number of nuclei in 1 g of ^{131}I is

$$1 \text{ g} \times \frac{1 \text{ mol}}{131 \text{ g}} \times 6.023 \times 10^{23} \text{ nuclei mol}^{-1} = 4.6 \times 10^{21}$$

Hence, the decay rate

$$-\frac{dN}{dt} = \frac{N}{\tau} = \frac{\ln 2}{T_{1/2}} N$$

From Table 12.4 you would note that $T_{1/2}$ (^{131}I) = 8.05 days.

$$\therefore \quad -\frac{dN}{dt} = \frac{0.693}{305 \text{ days}} (4.6 \times 10^{21} \text{ atoms})$$

$$= \frac{3.188 \times 10^{21} \text{ atoms}}{6.955 \times 10^{5} \text{ s}}$$

$$= 4.58 \times 10^{15} \text{ disintegrations per second}$$

Expressing this in curies, we find that

$$-\frac{dN}{dt} = (4.58 \times 10^{15} \text{ disintegrations per second}) \times \frac{1 \text{ Ci}}{(3.7 \times 10^{10} \text{ disintegrations per second})}$$

$$= 1.24 \times 10^{5} \text{ Ci}$$

This is an extremely large disintegration rate. The amount of ^{131}I injected into the human body is only 10^{-9}g so that the decay rate is nearly 10^{-4} Ci, which is well below the safety limit of

We will now illustrate how a knowledge of decay rate enables us to estimate the age of specimen using the carbon dating technique.

Example 4

^{14}C isotope of carbon is used for radioactive dating of organic materials. Samples of fresh carbon from trees in equilibrium with the CO_2 of the atmosphere have an abundance of 98.89% ^{12}C, 1.11% ^{13}C and 1.3×10^{-12} %^{14}C. After a tree dies, the abundance of ^{12}C and ^{13}C in the wood does not change but the abundance of ^{14}C decreases because of radioactive decay. A piece of wood is taken from an Egyptian tomb. Each gram of carbon exhibits an activity of 3.9×10^{-12}Ci. Estimate the age of the wood.

Solution

The number of nuclei in 1g of carbon is $1\text{g} \times \frac{1 \text{ mol}}{12 \text{ g}} \times 6.02 \times 10^{23} \text{ nuclei mol}^{-1} = 5.02 \times 10^{22}$ nuclei. Thus, 1g of fresh carbon should contain $5.02 \times 10^{22} \times 1.3 \times 10^{-12} = 6.53 \times 10^{10}$ nuclei of ^{14}C. From Eq. (12.4) the activity at $t = 0$ is given by

$$A_0 = \left|\frac{dN}{dt}\right|_{t=0} = \frac{\ln 2}{T_{1/2}} N_0$$

From Table 12.4 we note that $T_{1/2}$ (^{14}C) is 5730 yr. Hence

$$A_0 = \frac{0.693}{(5730 \times 365 \times 24\ 3600 \text{ s})} \times (6.53 \times 10^{10} \text{ nuclei})$$

$$= \frac{4.53 \times 10^{10} \text{ nuclei}}{1.807 \times 10^{11} \text{ s}} = 0.251 \text{ disintegrations per second}$$

To express it in curies we note that

$$A_0 = \frac{(0.251 \text{ disintegrations per second}) \times 1 \text{ Ci}}{3.7 \times 10^{10} \text{ disintegrations per second}}$$

$$= 6.78 \times 10^{-12} \text{Ci}$$

We are told that the measured activity of the sample is 3.9×10^{-12} Ci. This obviously is smaller than the initial activity by a factor of $\frac{3.9 \times 10^{-12} \text{ Ci}}{6.78 \times 10^{-12} \text{ Ci}} = 0.575$

Since the activity is proportional to the amount of radioactive material, from Eq. (12.10) we recall that

$$\frac{A}{A_0} = \frac{N}{N_0} = e^{-(\ln 2)t/T_{1/2}} = 0.575$$

so that by taking logarithm of both sides we find that

$$t = -\frac{T_{1/2}}{\ln 2} \ln (0.575)$$

$$= -\frac{5730 \text{ yr}}{0.693} \times (-0.553)$$

$$= 4572.4 \text{ yr}$$

Spend 10 min.

SAQ 2

The half life of ^{238}U is known to be 4.51×10^9yr. Compute the disintegration constant (in s^{-1}). Also calculate the number of disintegrations per second from 1 gm of ^{238}U. Avogadro's number=6.023×10^{23}.

12.4 GROWTH AND DECAY OF RADIOACTIVITY

You now know that naturally occurring radioactive elements disintegrate continuously and their decay is governed by a characteristic decay constant. Physically it means that the number of atoms (of parent nuclei) in a radioactive sample will decrease steadily (according to Eqs. (12.3) and (12.7)). However, if the daughter nuclei happen to be radioactive, their (daughter nuclei) activity will also start to build up with time (and may compensate for the loss of parent nuclei). But this cannot go on indefinitely. Let us now find out more about the observed growth and decay of the activities of the daughter nuclei.

Let us reconsider the decay of ^{238}U. We now know that it decays by emitting an α-particle with a half- life of 4.5×10^9 yr. But ^{234}Th atoms disintegrate by emitting β-particles with a half-life of only 24.1 days. This means that in the original sample of uranium, ^{234}Th atoms will disintegrate at a much faster rate and all the apparent activity of the sample will be practically due to them. However, if ^{234}Th is separated from the uranium, its activity will decay exponentially and reach half its initial value in 24.1 days.

To know how ^{234}Th grows in a freshly separated sample of uranium, we note that if at any instant of time t, the number of atoms of ^{238}U and ^{234}Th are N_U and N_{Th}, then the rate of disintegration of the parent element dN_U

$$\frac{dN_U}{dt} = -\lambda_U N_U$$

Physically it also means that ^{234}Th atoms are produced at the rate $\lambda_U N_U$, where λ_U is the decay constant of uranium. However, ^{234}Th atoms will disintegrate at a rate $\lambda_{Th} N_{Th}$, where N_{Th} is the number of ^{234}Th atoms present at time t and λ_{Th} is their characteristic decay constant. Hence, the net rate of increase of ^{234}Th atoms in uranium is given by

$$\frac{dN_{th}}{dt} = \lambda_U N_U - \lambda_{Th} N_{Th}$$

or

$$\frac{dN_{th}}{dt} + \lambda_{Th} N_{Th} = \lambda_U N_U \qquad (12.11)$$

To solve this ODE for N_{Th}, we have to convert its lefthand side into an exact differential. We can do so by multiplying throughout by $\exp(\lambda_{Th} t)$, which acts as integrating factor. This gives

$$\exp(\lambda_{Th} t)\frac{dN_{th}}{dt} + \lambda_{Th} N_{Th} \exp(\lambda_{Th} t) = \lambda_U N_U \exp(\lambda_{Th} t)$$

so that

$$\frac{d}{dt}\{N_{Th} \exp(\lambda_{Th} t)\} = \lambda_U N_U \exp(\lambda_{Th} t)$$

Integrating we obtain

$$N_{Th} \exp(\lambda_{Th} t) = \frac{\lambda_U}{\lambda_{Th}} N_U \exp(\lambda_{Th} t) + K$$

or $$N_{Th} = \frac{\lambda_U}{\lambda_{Th}} N_U + K \exp(-\lambda_{Th} t) \quad (12.12)$$

where K is a constant of integration. To evaluate it, we note that at $t = 0$, i.e. in a freshly prepared uranium sample $N_{Th} = 0$. Hence, it readily follows that $K = -\frac{\lambda_U}{\lambda_{Th}} N_U$. Therefore Eq. (12.12) takes a compact form:

$$N_{Th} = \frac{\lambda_U}{\lambda_{Th}} N_U [1 - \exp(-\lambda_{Th} t) \quad (12.13a)$$

This result shows that in the separated uranium fraction, the number of ^{234}Th atoms ultimately reaches a constant equilibrium value of N_{Th}^{∞} $(\lambda_U / \lambda_{Th}) N_U$. On combining this result with Eq. (12.13a) we can write

$$N_{Th} = N_{Th}^{\infty} [1 - \exp(-\lambda_{Th} t)] \quad (12.13b)$$

The growth of ^{234}Th are shown in Fig.12.3.

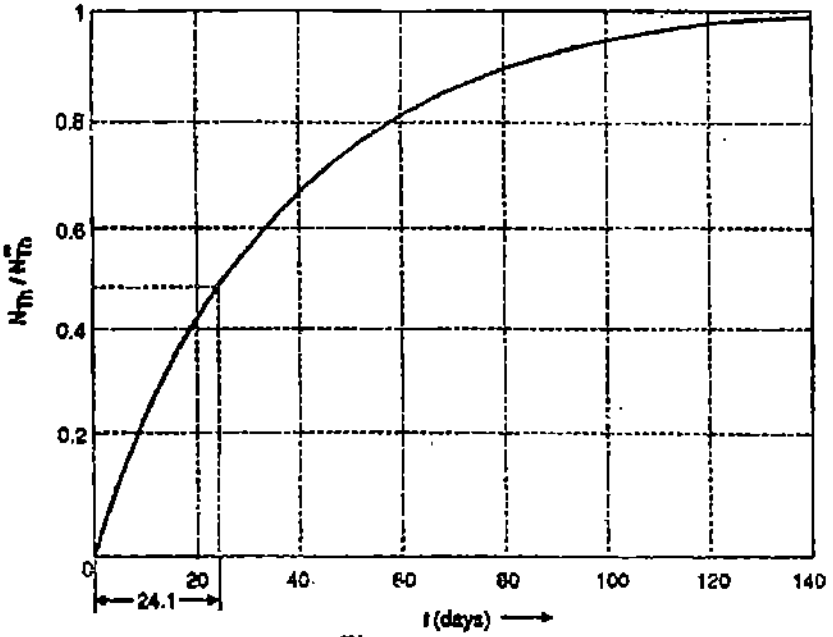


Fig. 12.3 : Growth of ^{234}Th in a freshly prepared uranium sample

12.5 SUCCESSIVE RADIOACTIVE TRANSFORMATIONS

In the preceding section we considered growth and decay of radioactivity when the parent nuclei decay into daughter nuclei, which decay in turn. You now know that naturally occurring radioactive elements undergo successive decays till a stable element is produced. While studying a particular radioactive series, we often find it necessary to work out the number of atoms of each member of the series at a specified time. We can express the problem as follows:

Suppose that initially we have N_{A0} atoms of a parent element A. It decays to an element B which in turn decays to C and so on. If λ_A, λ_B, λ_C,... are the respective disintegration constants of $A, B, C, \ldots$ let us derive expressions for the number of atoms N_A, N_B, N_C at time t. From Eq. (12.7) we note that the number of atoms of A present at time t is given by

$$N_A = N_{A0} \exp(-\lambda_A t) \quad (12.14)$$

The atoms of element B are produced at the rate $\lambda_A N_A$ and decay at the rate $\lambda_B N_B$. Hence, the net rate at which the number of atoms of B grows is given by

$$\frac{dN_B}{dt} = \lambda_A N_A - \lambda_B N_B \tag{12.15}$$

For element C, we can similarly write

$$\frac{dN_C}{dt} = \lambda_B N_B - \lambda_C N_C \tag{12.16}$$

On substituting for N_A from Eq. (12.14) in Eq. (12.15) and rearranging terms in the resultant expression, we get

$$\frac{dN_B}{dt} + \lambda_B N_B = \lambda_A N_{A0} \exp(-\lambda_A t) \tag{12.17}$$

Multiplying both sides by $\exp(\lambda_B t)$ and following the steps outlined in the previous section, you can easily show that (SAQ 3)

$$N_B = \frac{\lambda_A N_{A0}}{\lambda_B - \lambda_A}(e^{-\lambda_A t} - e^{-\lambda_B t}) \tag{12.18}$$

Spend **10 min**

SAQ 3

Starting from Eq. (12.17), derive Eq. (12.18) using the condition that initially only the parent element was present, i.e. $N_B = 0$ at $t = 0$.

If you now substitute this expression for N_B in Eq. (12.16) and multiply throughout by $\exp(\lambda_C t)$, you can rearrange the resultant expression as

$$\frac{d}{dt}\exp(\lambda_C t) = \frac{\lambda_A \lambda_B}{\lambda_B - \lambda_A} N_{A0}\left[e^{(\lambda_C - \lambda_A)t} - e^{(\lambda_C - \lambda_B)t}\right]$$

This can readily be integrated to yield

$$\exp(\lambda_C t) N_C = \frac{\lambda_A \lambda_B}{\lambda_B - \lambda_A} N_{A0}\left[\frac{e^{(\lambda_C - \lambda_A)t}}{(\lambda_C - \lambda_A)} - \frac{e^{(\lambda_C - \lambda_B)t}}{(\lambda_C - \lambda_B)}\right] + K$$

$$N_C = \frac{\lambda_A \lambda_B}{\lambda_B - \lambda_A} N_{A0}\left[\frac{e^{-\lambda_A t}}{(\lambda_C - \lambda_A)} - \frac{e^{-\lambda_B t}}{(\lambda_C - \lambda_B)}\right] + K\exp(-\lambda_C t) \tag{12.19}$$

where K is constant of integration. To evaluate it, we use the fact that at $t = 0$, $N_C = 0$. This gives

$$K = -\frac{\lambda_A \lambda_B}{\lambda_B - \lambda_A} N_{A0}\left[\frac{1}{(\lambda_C - \lambda_A)} - \frac{1}{(\lambda_C - \lambda_B)}\right]$$

$$= \frac{\lambda_A \lambda_B N_{A0}}{(\lambda_C - \lambda_A)(\lambda_C - \lambda_B)}$$

Hence

$$N_C = \lambda_A \lambda_B N_{A0}\left[\frac{\exp(-\lambda_A t)}{(\lambda_B - \lambda_A)(\lambda_C - \lambda_A)} + \frac{\exp(-\lambda_B t)}{(\lambda_C - \lambda_B)(\lambda_A - \lambda_B)} + \frac{\exp(-\lambda_C t)}{(\lambda_A - \lambda_C)(\lambda_B - \lambda_C)}\right]$$

(12.20)

In a compact form, we can rewrite it as

$$N_C = N_{A0}\left(a_1 e^{-\lambda_A t} + a_2 e^{-\lambda_B t} + a_3 e^{-\lambda_C t}\right) \tag{12.21}$$

where $$a_1 = \frac{\lambda_A \lambda_B}{(\lambda_B - \lambda_A)(\lambda_C - \lambda_A)}$$

$$a_2 = \frac{\lambda_A \lambda_B}{(\lambda_A - \lambda_B)(\lambda_C - \lambda_B)}$$

and

$$a_3 = \frac{\lambda_A \lambda_B}{(\lambda_A - \lambda_C)(\lambda_B - \lambda_C)} \qquad (12.22)$$

You can now easily extend the procedure outlined above to a chain of radioactive elements. We leave this as an exercise for you.

Let us apply these equations to the specific case of successive radioactive transformations. If a metal wire is exposed for a few seconds to radioactive radon gas, we obtain a deposit of RaA, a decay product of radon. RaA has a half-life of 3.05 min and decays to RaB, whose half-life is 27 min. RaC decays to RaD with a half-life of 20 min. RaD has a half-life of 22 years and for all practical purposes during the experiment the number of RaD atoms may be taken to be constant.

The number of RaA, RaB, RaC and RaD atoms as functions of time are shown in Fig. 12.4. The number of RaA atoms, initially assumed to be 100 decreases exponentially with time and reaches a value 50 after 3.05 min. At time $t = 0$, there are no atoms of RaB, RaC and RaD present. However, the number of RaB atoms increases with time, passes through a maximum about 11 min later and then decreases with time. The number of RaC atoms passes through a maximum after about 35 min. The number of RaD atoms increases

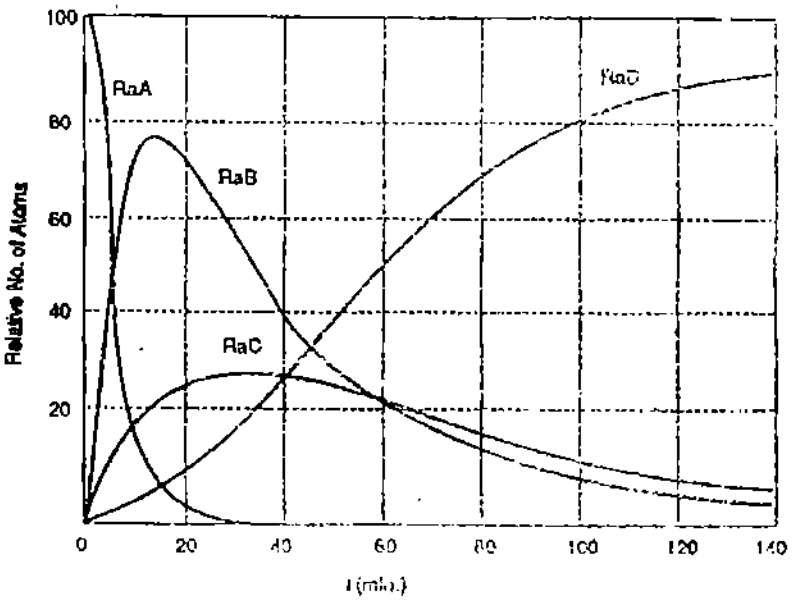


Fig. 12.4 : Variation of the relative numbers of atoms of RaA, RaB, RaC and RaD with time in a radioactive sample.

continually and reaches a maximum when RaB and RaC have practically disappeared. Eventually, the RaD will decay exponentially with a half-life of 22 years.

SAQ 4

In the above example, calculate the time at which the number of RaB atoms is a maximum. Spend 10 min

---

12.5.1 Radioactive Equilibrium

You now know that different radioactive elements, either naturally occurring or produced in a decay chain, have their characteristic half-lives. The law governing successive disintegrations deals with the quantity of any given radioactive isotope present at any time. Depending upon the relative magnitudes of the half lives of various nuclei in a decay chain, we may obtain a situation where the number of parent and/or daughter atoms either remain constant or bear a constant ratio. This is known as radioactive equilibrium. There are two possibilities. We will discuss these in turn.

Again refer to the decay of ^{235}U. You would recall that it is an extremely long-lived nuclei compared to the atoms of its daughter element ^{234}Th ($\lambda_U << \lambda_{Th}$). Let us assume that the initial sample is pure, i.e. has only ^{235}U. Then Eqs. (12.7) and (12.13) suggest that

$$N_U \equiv N_{U0} \qquad (12.23)$$

and $$N_{Th} = N_{U0}\frac{\lambda_U}{\lambda_{Th}}[1-\exp(-\lambda Th\, t)] \qquad (12.24)$$

since $\exp(-\lambda_U t) \approx 1$.

Let us pause for a while and think as to what have we achieved so far. Eqs. (12.23) and (12.24) suggest that though the number of uranium (parent) nuclei remain essentially constant, the number of daughter atoms increases exponentially with time. After a time t much greater than the half-life of daughter nuclei, $\exp(-\lambda_{Th} t)$ becomes negligibly small and N_{Th} reaches an equilibrium value of

$$N_{Th} = N_U\frac{\lambda_U}{\lambda_{Th}}$$

or $$N_{Th}\,\lambda_{Th} = N_U\,\lambda_U \qquad (12.25)$$

This result tells us that in the equilibrium state the rate of decay of the daughter atoms is equal to their rate of production implying that the number of parent and daughter atoms remains constant. Such a long-term equilibrium between the parent and daughter atoms is known as **secular equilibrium**. Such an equilibrium can be found in the formation of radon ($T_{1/2}$ = 5.5 days) from radium ($T_{1/2}$ = 2300 yr).

Spend 5 min

SAQ 5

Uranium minerals in which secular equilibrium has been obtained contain one atom of radium for every 2.8×10^6 atoms of uranium. If the half-life of radium is 1620 years, calculate the half-life of uranium.

You now know that in case of successive radioactive decays, secular equilibrium is obtained when parent nuclei are longer lived than those of the daughter element. You may now like to know: What will happen when the parent is longer-lived than the daughter ($\lambda_A < \lambda_B$), but the half-life of the parent is not very long. That is, the half-life of the parent is greater than the daughter by a small factor. You may like to identify such a situation in the uranium as well as actinium series. In such a situation we cannot use the approximation $\exp(-\lambda_A t) = 1$. If the parent and daughter atoms are initially separated, the number of their atoms are respectively given by Eqs. (12.14) and (12.18). Moreover, if $\lambda_B t \gg 1$, then in Eq. (12.18), the term involving $e^{-\lambda_B t}$ becomes negligible compared with the term involving $e^{-\lambda_A t}$. Then the number of daughter atoms is given by

$$\frac{N_B}{N_A} = \frac{\lambda_A}{\lambda_B - \lambda_A} N_{A0} \exp(-\lambda_A t) \qquad (12.26)$$

That is, the daughter nuclei eventually decay with the half-life of the parent. On combining this result with Eq. (12.14) we find that

$$\frac{N_B}{N_A} = \frac{\lambda_A}{\lambda_B - \lambda_A} = \text{constant} \qquad (12.27)$$

In words, the ratio of the number of parent atoms and the number of daughter atoms attains a constant value. This constitutes what we term as **transient equilibrium**.

When the parent has a shorter half-life than the daughter ($\lambda_A > \lambda_B$), no equilibrium is attained. If the parent and daughter are separated initially, then as the parent decays, the number of daughter atoms is given by

$$N_B = N_0\frac{\lambda_A}{\lambda_B - \lambda_A}\exp(-\lambda_B t) \qquad (12.28)$$

That is, the parent substance disappears completely and the daughter atoms eventually decay with their own half-life.

Spend 5 min

SAQ 6

For two radioactive elements A and B in transient equilibrium, show that the daughter activity is greater than the parent activity by a factor $\dfrac{\lambda_A}{\lambda_B - \lambda_A}$.

12.6 SUMMARY

- A radioactive nucleus disintegrates spontaneously by emitting either an α or a β-particle, usually accompanied by γ radiation.
- The number of atoms disintegrating per unit time is given by $N = N_0 e^{-\lambda t}$, where N_0 is the number of atoms present initially and λ is the disintegration constant of the radioactive element.
- The half-life of a radioactive element is the time taken for half of the radioactive atoms to disintegrate; this is related to the disintegration constant λ and mean life τ through the relations $T_{1/2} = 0.693/\lambda = 0.693\tau$.
- The standard unit of radioactivity, the **curie**, is defined as the quantity of any radioactive material giving 3.7×10^{10} disintegrations per second. The SI unit of radioactivity is **rutherford**. It is defined as the amount of a radioactive substance giving 10^6 ds^{-1}.
- The naturally occurring radioactive elements conform to three radioactive series, known as the **uranium, actinium and thorium series**. Each series starts with an element having an extremely long half- life and terminates in a stable isotope of lead. With the discovery of transuranic elements, a fourth radioactive series has been traced, known as the **neptunium series**; this starts with plutonium and terminates in a stable isotope of bismuth.
- If the parent atom has a half-life very long compared to any of its decay products, we get a long- term equilibrium, known as **secular equilibrium**, between the parent and the daughter atoms, when each member decays at the same rate as they are produced, i.e. we have $\lambda_A N_A = \lambda_B N_B = \lambda 3\, N_B$. However, if the parent A is longer lived than the daughter B, but $T_{1/2}$ is not very long, we obtain **transient equilibrium**, in which the ratio of the members of A and B atoms at any instant remains a constant.

12.7 TERMINAL QUESTIONS

1. Given that the half-lives of Radium and Radon are 1620 yr and 3.82 d respectively. Calculate the volume of Radon gas at N.T.P. equivalent to one curie.
2. A sample containing 0.1 mg of ^{230}Th undergoes 4.32×10^6 disintegrations per minute. What is the half- life of this nuclide?

12.8 SOLUTIONS AND ANSWERS

SAQs

1. Here $\tau = 14.43$ months

$$\therefore \quad T_{1/2} = \tau \ln 2$$
$$= (14.43 \text{ months}) \times 0.693$$
$$= 10 \text{ months}$$

Since 75% of the substance decays, only 25% remains.

$$\therefore \quad N = \frac{N_0}{4} = \left(\frac{1}{2}\right)^2 N_0$$

Using Eq. (12.2) we can write

$$\left(\frac{1}{2}\right)^2 = \left(\frac{1}{2}\right)^{t/T_{1/2}}$$

so that

$$\frac{t}{T_{1/2}} = 2$$

or

$$t = 10 \times 2 \text{ months}$$
$$= 20 \text{ months}$$

2. Half-life $T_{1/2} = 4.51 \times 10^9$ yr $= 4.51 \times 10^9 \times 365 \times 86400$ s

Disintegration constant $\lambda = \dfrac{0.693}{T_{1/2}}$

$$= \frac{0.693}{4.51 \times 10^9 \times 365 \times 86400} \text{ s}^{-1} = 4.87 \times 10^{-18} \text{ s}^{-1}$$

Number of atoms per g of uranium is

$$N = \frac{6.03 \times 10^{23}}{238}$$

$\therefore$ Rate of disintegration $\left|\dfrac{dN}{dt}\right| = \lambda N$

$$= \frac{6.03 \times 10^{23}}{238} \times 4.87 \times 10^{-18} \text{s}^{-1}$$
$$= 1.234 \times 10^4 \text{s}^{-1}$$

3. From Eq. (12.17) we know that

$$\frac{dN_B}{dt} + \lambda_B N_B = \lambda_A N_0 \exp(-\lambda_A t)$$

Multiplying throughout by $\exp(\lambda_B t)$ and re-arranging terms, you will obtain

$$\frac{d}{dt}[N_B \exp(\lambda_B t)] = \lambda_A N_0 \exp[(\lambda_B - \lambda_A)t]$$

This can readily be integrated to give

$$N_B e^{\lambda_B t} = \frac{\lambda_A}{\lambda_B - \lambda_A} N_0 \exp[(\lambda_B - \lambda_A)t] + K$$

where K is constant of integration. To evaluate it, we use the condition $N_B = 0$ at $t = 0$. This gives

$$K = \frac{\lambda_A}{\lambda_B - \lambda_A} N_0 = \frac{\lambda_A}{\lambda_A - \lambda_B} N_0$$

On inserting this value of K in the above expression we obtain the required result

$$N_B = \frac{\lambda_A}{\lambda_B - \lambda_A} N_0 [\exp(-\lambda_A t) - \exp(-\lambda_B t)]$$

4. We know from Eq. (12.18) that the number of RaB atoms is given by

$$N_B = \frac{\lambda_A}{\lambda_B - \lambda_A} N_0 [e^{-\lambda_A t} - e^{-\lambda_B t}]$$

If N_B attains maximum value at $t = t_m$, we must have $\dfrac{dN_B}{dt} = 0$ at $t = t_m$ and $\dfrac{d^2 N_B}{dt^2} < 0$.

Therefore, on differentiating the expression for N_B once with respect to t and equating it to zero, we obtain

$\therefore$ $-\lambda_A e^{-\lambda_A t_m} + \lambda_B e^{-\lambda_B t_m} = 0$

or $\lambda_B e^{-\lambda_B t_m} = \lambda_A e^{-\lambda_A t_m}$

$\therefore$ $e^{(\lambda_A - \lambda_B) t_m} = \dfrac{\lambda_A}{\lambda_B}$

and $t_m = \dfrac{\ln(\lambda_A/\lambda_B)}{\lambda_A - \lambda_B}$

Substituting $\lambda_A = 3.8 \times 10^{-3} \text{s}^{-1}$ and $\lambda_B = 4.3 \times 10^{-4} \text{s}^{-1}$, we get

$$t_m = \frac{\ln\left(\frac{3.8 \times 10^{-3}\ s^{-1}}{4.3 \times 10^{-4}\ s^{-1}}\right)}{(3.8 \times 10^{-3} - 4.3 \times 10^{-4})\ s^{-1}} = \frac{\ln (8.837)}{33.7 \times 10^{-4}\ s^{-1}} = \frac{2.179\ s}{33.7 \times 10^{-4}}$$

$$= 647s = 10m\ 47s$$

5. Since uranium is in secular equilibrium with radium, we have

$$\frac{N_1}{\tau_1} = \frac{N_2}{\tau_2}$$

or $$\tau_1 = \frac{N_1}{N_2}\tau_2 = 2.8 \times 10^6 \times 1620\ yr = 4.5 \times 10^9 yr$$

6. For elements *A* and *B* in transient equilibrium, we have from Eqn. (12.27),

$$\frac{N_B}{N_A} = \frac{\lambda_A}{\lambda_B - \lambda_A}$$

Hence the ratio of the measured activities at equilibrium is given by

$$\frac{A_B}{A_A} = \frac{\lambda_B N_B}{\lambda_A N_A} = \frac{\lambda_B}{\lambda_A} \cdot \frac{\lambda_A}{\lambda_B - \lambda_A} = \frac{\lambda_B}{\lambda_B - \lambda_A}$$

$$\therefore \quad A_B = \frac{\lambda_B}{\lambda_B - \lambda_A} A_A$$

TQs

1. One curie is equivalent to the amount of radon in equilibrium with 1 g of radium. Hence if N_{Rn} is the number of Rn atoms in equilibrium with 1 g of Ra and N_{Ra} the number of atoms in 1 g of radium, then we have $\lambda_{Rn} N_{Rn} = \lambda_{Ra} N_{Ra}$.

$$\therefore \quad N_{Rn} = N_{Ra} \frac{\lambda_{Ra}}{\lambda_{Rn}} = N_{Ra} \frac{\tau_{Rn}}{\tau_{Ra}} = \frac{3.82\ d}{1620\ yr} \times N_{Ra}$$

$$= \frac{3.82}{1620 \times 365} \times \frac{N}{226}$$

where N is Avogadro's number.

$\therefore$ Volume occupied by N_{Rn} atoms at STP

$$= \frac{N_{Rn}}{N} \times 22.4 \times 10^3\ cm^3$$

$$= \frac{3.82}{1620 \times 365 \times 226} \times 22.4 \times 10^3\ cm^3$$

$$= 6.4 \times 10^{-4}\ cm^3$$

2. No. of atoms in 0.1 mg of ^{230}Th

$$N = \frac{6.02 \times 10^{23}}{230\ g} \times (10^{-4}\ g) = 2.62 \times 10^{17}$$

Rate of disintegration $\frac{dN}{dt} = 4.32 \times 10^5\ m^{-1}$

$$= 7.2 \times 10^{-7}\ s^{-1}$$

$\therefore$ Decay constant $\lambda = \frac{1}{N}\frac{dN}{dt}$

$$= \frac{7.2 \times 10^{-7}\ s^{-1}}{2.62 \times 10^{17}} = 2.75 \times 10^{-13}\ s^{-1}$$

$$\therefore \quad T_{1/2} = \frac{0.693}{\lambda} = \frac{0.693}{2.75 \times 10^{-13}}\ s$$

$$= 7.99 \times 10^4\ yr$$

We have from Eq. (12.13 b)

$$\frac{N_2(t)}{N_2(\infty)} = 1 - e^{-\lambda_2 t} = \frac{90}{100} \text{ (given)}$$

$$\therefore \quad e^{-\lambda_2 t} = \frac{1}{10} \text{ or } \quad t = \frac{\log_e 10}{\lambda_2} = \tau_2 \frac{\log_e 10}{\log_e 2}$$

Hence, $\frac{t}{\tau_2} = \frac{\log_e 10}{\log_e 2} = 3.32$

UNIT 13 THE ATOMIC NUCLEUS

Structure

13.1 Introduction
Objectives
13.2 The Alpha-Particle Experiment
13.3 Binding Energy of Nuclei
13.4 Nuclear Fission
Liquid Drop Model of Fission
Critical Energy for Fission: Spontaneous Fission
13.5 Nuclear Models
The Shell Model
The Collective Model
13.6 Summary
13.7 Terminal Questions
13.8 Solutions and Answers

13.1 INTRODUCTION

You were introduced to the atom in your plus-two Physics course. You now know that atoms are too small to be seen even with the most powerful optical microscope. They are too small to be weighed even with the most sensitive balance. Their existence was initially postulated by Dalton for explaining the laws of chemical composition. However, the investigations of Thomson, which conclusively proved the existence of electrons as fundamental constituents of all matter, gave rise to the question: How are electrons distributed in an atom? If atoms contain electrons, they should also contain positive charges since electrons are negatively charged and atoms should be electrically neutral. That is, the existence of electrons implied that there were positive charges to be accounted for within an atom. On the basis of these observations, Thomson proposed his plum-pudding model of atom. He suggested that the atom consisted of a uniform positive sphere (radius about 10^{-10}m) of electrification with electrons embedded in such a way as to give it a most stable configuration. (According to this model, electrons are embedded in a spherical cloud of positive charges like seeds in a watermelon.) This model suggested that the mass of an atom is spread uniformly through out its volume. This hypothesis seemed reasonable, yet it could not withstand the experimental test of Geiger and Marsden who bombarded thin metallic foils by alpha-particles spontaneously emitted by certain naturally occurring radioactive elements. The picture of the atom that emerged from these experiments revealed that an atom consists of a centrally located positively charged core, called the **nucleus**, with negatively charged electrons some distance away. We will discuss this in detail in Sec. 13.2 Subsequent researches revealed that the nucleus is composed of protons and neutrons. The proton carries a positive charge of nearly 1.6×10^{-19}C, which is equal in magnitude but opposite in sign to that of an electron. Neutrons are uncharged particles slightly heavier than protons. And almost the entire mass of an atom is concentrated in its nucleus. Moreover, in its normal (electrically neutral) state, an atom has as many electrons as the number of protons in the nucleus.

You may now like to know: How do nucleons (protons and neutrons) cling together? The gravitational attraction between them is far too weak to hold the nucleus together and the Coulomb repulsion between protons should blow them apart. The forces which hold the nucleons together are called **nuclear forces** and are discussed in Sec. 13.3. You will learn that **stability of a nucleus is governed by the binding energy per nucleon.** An important consequence of this is that very light as well as very heavy elements are less stable and have tendency to fuse or fission under suitable conditions. The liquid drop model has been applied to predict the instability of heavier nuclei with reference to fission. This is the subject matter of Sec. 13.4. In Sec. 13.5 we have discussed various nuclear models. We are aware that you are familiar with some of these concepts. The repetition is for recapitulation.

Objectives

After going through this unit, you will be able to

- describe the importance of Rutherford's scattering experiment
- calculate the binding energy per nucleon based on semi-empirical mass formula

- explain how nuclear forces give nucleons
- describe the liquid drop and shell models of nuclei and
- discuss the stability of nuclei.

Ernest Rutherford

13.2 THE ALPHA-PARTICLE EXPERIMENT

To 'see' what is inside an atom, Rutherford suggested that alpha-particles from naturally occurring radioactive elements like polonium, radium etc. can be used as convenient tools to probe the structure of atom. A series of classic experiments were carried out under his guidance by Geiger and Marsden. The schematics of the apparatus used by them is shown in

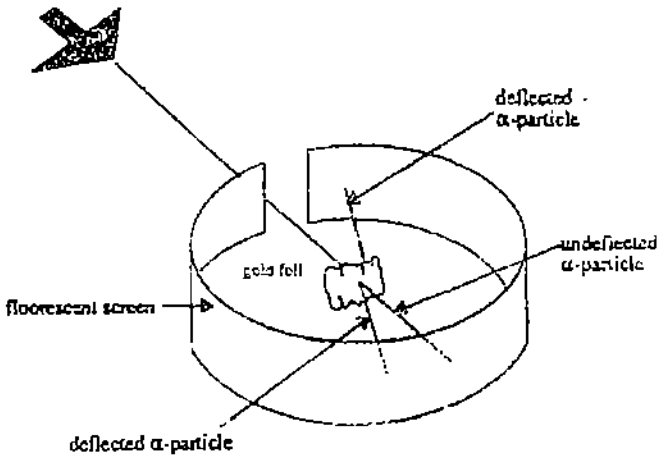


Fig. 13.1: Schematics of α-particle experiment

Fig. 13.1. In one of their experiments, 5.5 MeV alpha-particles ($v \cong 1.63 \times 10^7 ms^{-1}$) from ^{214}Bi source were made to enter an evacuated chamber and directed at a thin gold foil. The foil was a few micro-meters ($2.1 \times 10^{-7}m$) thick. From the previous unit you will recall that air absorbs α-particles. It is for this reason that the experiment had to be performed in vacuum. The alpha-particles were expected to strike the flouresceni ZnS screen producing a visible flash of light. On the basis of plum-pudding model it was anticipated that most of the alpha-particles would go straight through the foil and only a few would at best suffer slight deflection. What actually happened and was observed was :

- nearly all the alpha-particles emerged without much deviation ($\theta < 1°$). This suggested that atoms consist largely of empty space.

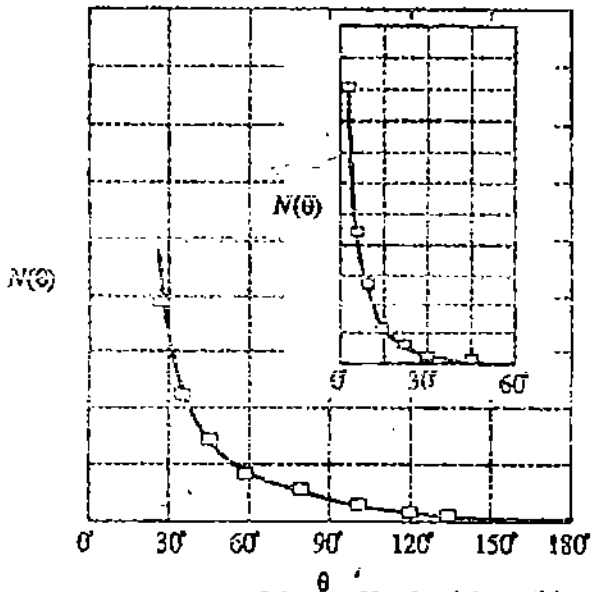


Fig.13.2: Plot of $N(\theta)$ versus θ for Geiger and Marsden alpha-particle experiment

- some (1 in 9000) alpha-particles were deflected through very large angles ($\theta \approx 90°$)
- a few (1 in 20000) alpha-particles were even deflected back towards the source ($\theta \rightarrow 180°$)

These results are shown in Fig. 13.2, which depicts the plot of the number $N(\theta)$ of scattered α-particles versus the angle of scattering, θ.

Large angle deflection of alpha particles, particularly the backward scattering, was rather surprising, in fact completely unexpected. In the words of Rutherford,

"It was the most incredible event that has ever happened to me in my life. It was almost as if you have fired a 15" shell at a piece of tissue paper and it came back and hit you."

It only suggested that the alpha-particles moving back towards the source had rebounded from the head-on collision from something immovable. It could happen only when these alpha-particles were acted on by very powerful force. Can you suggest what type of force this is? Since alpha-particles carry positive charge, it seems that the force between the atoms and the alpha- particles was electrostatic. Since an alpha-particle carries kinetic energy, it is gradually converted to electrostatic potential energy and at the point of the closest approach, the energy is completely potential so that an alpha-particle should come to rest momentarily. If b is the distance of closest approach, then from the principle of conservation of energy, we can write

$$E = \frac{1}{2} mv^2 = \frac{1}{4\pi\epsilon_0} \frac{2Ze^2}{b}$$

or

$$b = \frac{1}{4\pi\epsilon_0} \frac{4Ze^2}{mv^2} \qquad (13.1)$$

where v is speed of α-particles of mass m, Z is the atomic number of the foil atoms and e is electronic charge.

This relation suggests that the object carries the positive charge. To give you an idea about the numerical value of the radius of this object, let us calculate how close incident α-particles of energy 5.5MeV, which are scattered through 180°, get to it. Here

$Z = 79$, $e = 1.6 \times 10^{-19}$ C, $\frac{1}{4\pi\epsilon_0} = 9.0 \times 10^9$ Nm^2C^{-2} and

$E = 5.5\text{MeV} = (5.5 \times 10^6 \text{eV})(1.6 \times 10^{-19} \text{J eV}^{-1}) = 8.8 \times 10^{-13}\text{J}$. Therefore,

$$b = \frac{(9 \times 10^9 \text{ N m}^2\text{c}^{-2}) \times 2 \times 79 \ (1.60 \times 10^{-19} \text{C})^2}{8.8 \times 10^{-13}\text{J}}$$

$$= 4.14 \times 10^{-14} \text{ m}$$

This shows that the distance of closest approach in gold is 4.14×10^{-14}m and the positive charge in the targeted object is distributed over space confined to a radius less than this.

The radius of the atom is 10^4 times as big! So if the nucleus of an atom were of the size of a tennis or golf ball, the atom would fit into a large football stadium.

On the basis of a series of experiments, Rutherford proposed the nuclear model of atom, which is shown in Fig. 13.3. The genius in Rutherford argued that the rebound could occur

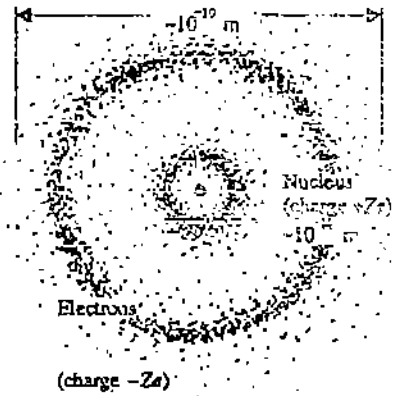


Fig. 13.3: Nuclear model of atom

because of the repulsive force between the positive charge of the alpha-particles and some positive charge in the atom. For the force to be sufficiently strong, he conjectured that all the positive charge in the atom must be concentrated in a small space. This positively charged centre must carry most of the mass of the atom. You can visualise this picture by analogy with the game of marbles (kancha). As a child, you must have played marbles. You will know from experience that a light marble bounces off a heavier one without much effect on it. (Gold atoms here are like a heavy marble and alpha-particles are analogous to light marbles.)

Proceeding along this line of thought Rutherford proposed that

- all the mass and the positive charge in an atom is concentrated in a tiny nucleus,
- the radius of the nuclei is of the order of a few fermi (1 fm = 10^{-15}m), and
- electrons reside outside the nucleus.

Rutherford's model of the atom resembles in many ways the model of our solar system. Just as a planet revolves around the Sun, the electrons revolve around the nucleus. Most of our solar system is empty space and so also is an atom.

Using this nuclear model of the atom, it is easy to picturise how most of the particles go right through the thin foil. Fig. 13.4 depicts the nuclei of atoms as small centres.

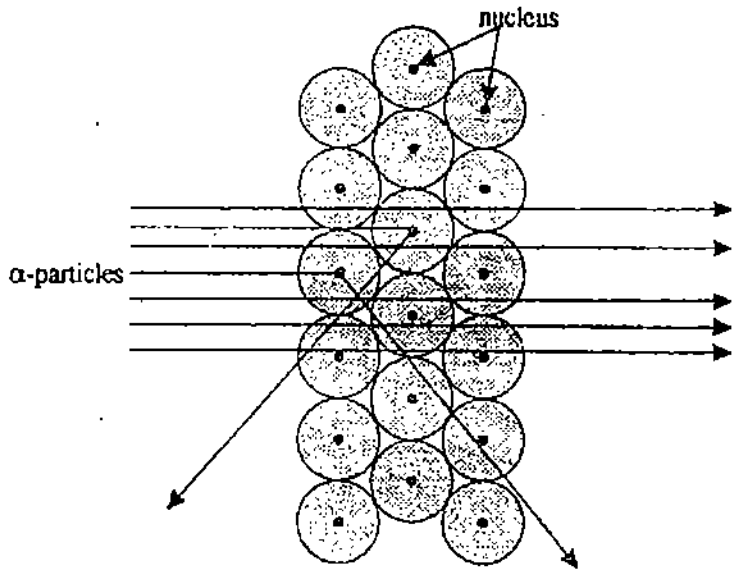


Fig.13.4: Scattering of alpha-particles by atomic nuclei on the basis of Rutherford's nuclear model

Rutherford's model is inherently simple and satisfactorily explains the scattering of alpha-particles. But, these investigations derive their real worth from the fact that they opened up very fertile and new avenues of research. A lot of good new physics of the atom emerged out of these studies.

We all know that Coulomb interactions involve two charges (an alpha-particle and the gold nucleus in the above experiment) and the distance between them. In 1920 Chadwick, a student of Rutherford, carried out a series of experiments to determine the charge on the nuclei of several elements. He found that the charges on the nuclei were exact integral multiples of the electronic charge but of opposite sign. In the gold foil used by him, there were millions of nuclei; each with the same positive charge of 79 units. The question then arose: How was it that the foil had no net electric charge? The only possible explanation was: the positive charge on each nucleus is exactly equal to the negative charge carried by the electrons. This means that each atom is electrically neutral and has 79 electrons. This suggests that atoms of each element can be uniquely characterised by a single number known as the atomic number, denoted by Z. You would now like to know: Do all nuclei consist only of protons? What is the structure of the nucleus? Let us consider the simplest element—hydrogen. Its nucleus has one proton and one electron. Does this mean that nuclei of all elements are composed only of protons? Definitely not. Even two other known forms of hydrogen-deuterium and tritium - conspicously differed in their masses. It puzzled all physicists.

The heavy isotope, deuterium, of hydrogen was discovered by Urey and his coworkers in 1931 by spectrographic methods. This however was predicted by Rutherford in 1920.

To circumvent the problem it was proposed initially that electrons also reside inside the nucleus and balanced the excessive positive charge of protons. However, this hypothesis was rejected as it was not consistent with Heisenberg's uncertainty principle, which required that for this to happen electrons given out by elements should be of energies of about 50MeV. Experimental evidences did not support it. This impasse, however, pointed to the possibility

Heisenberg's uncertainty principle states that one cannot simultaneously determine and location of a particle with absolute certainty. Mathematically, we write

$$\Delta p \times \Delta x = \hbar$$

where $\hbar = \frac{h}{2\pi}$

of existence of such inhabitants in the nucleus that contributed to its mass but not to its charge.

Rutherford conjectured the existence of a particle devoid of all charge but slightly heavier than the proton in the nucleus. In 1920, in a lecture to the Royal Society, Rutherford observed,

"It seems very likely that one electron can also bind two hydrogen nuclei and possibly also one hydrogen nucleus. In the one case, this entails the possible existence of an atom of mass nearly two, carrying unit charge, which is to be regarded as an isotope of hydrogen. In the other case, it involves the idea of the possible existence of an atom of mass one which has zero charge. Such an atomic structure seems by no means impossible..... Such an atom would have novel properties. Its external field would be practically zero, except very close to the nucleus, and in consequence it would be able to move freely through matter. Its presence would be difficult to detect by spectroscope and it may be impossible to contain it in a sealed vessel. On the other hand, it should enter readily the structure of atoms and may either unite with nuclei or be disintegrated by its intense field resulting in the escape of a charged hydrogen atom or an electron or both ...".

This 'atom' was named neutron by Rutherford.

The experimental evidence for the existence of neutrons came in 1932. The discovery of the neutron crystallised our ideas about nuclear structure. It became clear that all nuclei are composed of protons and neutrons (with the exception of a hydrogen nucleus). The number of neutrons and protons together define the mass number, denoted by A. The neutrons and protons are collectively referred to as nucleons.

We often come across elements whose atoms have the same number of electrons but their nuclear masses differ. They are called isotopes. For example, deuterium nucleus has one proton and one neutron and tritium nucleus consists of one proton and two neutrons. Since they have only one electron like hydrogen, they are different varieties of hydrogen and are called isotopes of hydrogen. These are denoted by the symbols of ^1H, ^2H, ^3H. Similarly, lithium has two stable isotopes - ^6Li, ^7Li and uranium has three isotopes ^{233}U, ^{235}U and ^{238}U and so on.

Nuclear Density

The alpha-particle experiment provided the first evidence that nuclei are of extremely small size (~ 10^{-15}m). Since then, a variety of experiments have been performed to determine nuclear dimensions using high energy electrons and neutrons as probes. These experiments have revealed that

- nuclei do not have sharp boundaries
- the density of nuclear matter is maximum at the centre of the nucleus and decreases gradually to zero as the distance increases.

You would now like to know the order of magnitude of the density of nuclear matter. Let us consider the lightest nucleus of hydrogen whose mass is 1.673×10^{-27} kg and the radius is 1.2×10^{-15} m. If we take it to be spherical, the density of nuclear matter can be computed as follows:

$$d_H = \frac{M_H}{\frac{4\pi}{3} R_H^3} = \frac{1.673 \times 10^{-27}\,\text{kg}}{\frac{4\pi}{3}(1.2 \times 10^{-15}\text{m})^3} = 2.3 \times 10^{17}\,\text{kg m}^{-3}.$$

This value is extremely high. Recall the density of water (= 10^3kg m^{-3}) or of mercury (13.6×10^3 kg m^{-3}). It shows that nuclear matter is extremely intensely packed. The mass of our earth (= 6×10^{24} kg) can be packed together into such a high density sphere of radius 184m. You should convince yourself by doing this calculation. Also, compute the radius of the nuclear sphere whose mass will be equal to the mass of our sun. Your answer should be nearly 10km!

Next, let us calculate the density of nuclear matter from the data for oxygen. It is reliably known that $R_O = 3 \times 10^{-15}$ m and $M_O = 2.7 \times 10^{-26}$ kg. Therefore,

$$d_O = \frac{2.68 \times 10^{-26}\,\text{kg}}{\frac{4\pi}{3}(3 \times 10^{-15}\text{m})^3} = 2.39 \times 10^{17}\,\text{kg m}^{-3}$$

What do you observe? The density of hydrogen and oxygen nuclei is nearly the same. Is it a mere coincidence? To discover the answer, solve the following SAQ.

In 1930, Bothe and Becker bombarded beryllium and some other light elements (Li, B, Mg, Al) with α-particles from a polonium source and observed, using Geiger-Müller counters, penetrating radiations with energies roughly equal to those of natural γ-rays. Using an ionization chamber sensitive to charged particles as well as γ-rays and a more intense polonium source, Irène Curie and her husband Frédéric Joliot continued the investigation of these penetrating radiations emitted by beryllium and confirmed the results of Bothe and Becker. In addition they found a penetrating radiation which could liberate protons of roughly 5 MeV energy from hydrogeneous materials such as paraffin wax. They proposed that protons were ejected by Compton scattering of high energy (50 MeV) γ-rays. However, on theoretical grounds it was difficult to account for such high energy γ-rays arising in this reaction of alpha particles with beryllium. This discrepancy induced Chadwick to repeat the observations. From his experiments, performed at the Cavendish Laboratory in 1932, Chadwick proved conclusively the existence of a neutral particle with rest mass nearly the same as that of the proton. This was the neutron, the particle predicted about 12 years earlier by Rutherford. As we now know, when an α-particle hits a beryllium nucleus, the reaction is

$^9Be + \alpha \rightarrow [^{12}C]^* + n$

which is often written as

$^9Be(\alpha,n)^{12}C$. This notation indicates that the carbon nucleus produced in the reaction is in an excited state and returns to the ground state by emitting γ-rays of a few MeV. These were the γ-rays that were originally observed by Bothe and Becker.

SAQ 1

Calculate the density of carbon and lead nuclei using the following data:

Spend 5 min

$M_C = 19.92 \times 10^{-27}$ kg, $R_C = 2.7 \times 10^{-15}$ m

$M_{Pb} = 3.4 \times 10^{-25}$ kg and $R_{Pb} = 7 \times 10^{-15}$ m

---

On solving this SAQ you will come to the conclusion that nuclei of all elements have the same density. This suggests that

- nucleus is analogous to a drop of liquid and
- there is an empirical relationship between radius of a nucleus, R, and its mass number A. Indeed, experimental evidences suggest that R and A are connected through the relation $R = 1.2 \times 10^{-15} A^{1/3}$ m.

To get a quick idea about this, let us calculate the radii of carbon and lead nuclei. To this end, we note that

$$M_C = M_H A_C = \frac{4\pi}{3} R_C^3 d_C$$

$$R_C = \left[\frac{3}{4\pi}\left(\frac{M_H}{d_C}\right)\right]^{1/3} A_C^{1/3}$$

Similarly, $$R_{Pb} = \left[\frac{3}{4\pi}\left(\frac{M_H}{d_{Pb}}\right)\right]^{1/3} A_{Pb}^{1/3}$$

On substituting the values given and/or calculated in SAQ 1, we find that

$$R_C = \left[\frac{3}{4 \times 3.1417}\left(\frac{1.673 \times 10^{-27}\ \text{kg}}{2.42 \times 10^{17}\ \text{kg m}^{-3}}\right)\right]^{1/3} A_C^{1/3}$$

$$= \left(\frac{16.73 \times 0.2387}{2.42}\right)^{1/3} \times 10^{-15} A_C^{1/3}\ \text{m}$$

$$= \left(\frac{3.993}{2.42}\right)^{1/3} \times 10^{-15} A_C^{1/3}\ \text{m}$$

$$= (1.6502)^{1/3} \times 10^{-15} A_C^{1/3}\ \text{m}$$

$$= 1.1181 \times 10^{-15} A_C^{1/3}\ \text{m}$$

$$= 1.12 \times 10^{-15} A_C^{1/3}\ \text{m}$$

Similarly for lead,

$$R_{Pb} = \left[\frac{3}{4 \times 3.1416}\left(\frac{1.673 \times 10^{-27}\ \text{kg}}{2.37 \times 10^{17}\ \text{kg m}^{-3}}\right)\right]^{1/3} A_{Pb}^{1/3}$$

$$= \left(\frac{16.73 \times 0.2387}{2.37}\right)^{1/3} \times 10^{-15} A_{Pb}^{1/3}\ \text{m}$$

$$= \left(\frac{3.993}{2.37}\right)^{1/3} \times 10^{-15} A_{Pb}^{1/3}\ \text{m}$$

$$= (1.685)^{1/3} \times 10^{-15} A_{Pb}^{1/3}\ \text{m}$$

$$= 1.1190 \times 10^{-15} A_{Pb}^{1/3}\ \text{m}$$

$$= 1.12 \times 10^{-15} A_{Pb}^{1/3}\ \text{m}$$

13.3 BINDING ENERGY OF NUCLEI

We now know that the nucleus of deuterium contains one proton and one neutron. The measured (rest) masses of the proton and the neutron are 1.6723×10^{-27}kg and 1.6747×10^{-27} kg, respectively. This means that the total rest mass of a neutron plus a proton is 3.34709×10^{-27} kg. But the rest mass of a deuterium nucleus is 3.34313×10^{-27} kg. This means that the measured mass of deuterium nucleus is 3.96242×10^{-30} kg less than the measured masses of a neutron and a proton. In fact, it is quite well known now that mass of any nucleus is always less than the sum of the rest masses of its constituent nucleons. This difference is termed the mass defect. Let us denote it by Δm. Mathematically we can write

$$\Delta m = (Z m_p + N m_n) - (M - Z m_e)$$

$$= Z m_H + N m_n - M \qquad (13.2)$$

where M is the actual mass of the neutral atom containing Z protons and N neutrons. $m_H = (m_p + m_e)$, m_p, m_n and m_e are the masses of hydrogen atom, the proton, the neutron and the electron, respectively. It is often convenient to express the mass defect by its equivalent energy through Einstein's mass-energy equivalence relation:

$$BE = \Delta m\, c^2$$

For deuterium

$$BE = (3.96242 \times 10^{-30}\text{kg}) \times (2.998 \times 10^{8}\text{ms}^{-1})^2$$

$$= 35.614 \times 10^{-14}\text{kg m}^2\text{s}^{-2}$$

$$= 3.5614 \times 10^{-13}\text{J}$$

$$= 2.223 \times 10^{6}\text{eV}$$

since $1\text{ eV} = 1.602 \times 10^{-19}$J.

This means that we will have to supply atleast 2.223 MeV energy to free the constituent nucleons – protons and neutrons – of deuterium nucleus. We can generalise this result to say that mass defect appears as the energy which binds the nucleons together. This is essentially used up in doing work against the forces which bind the nucleons together.

If we supply more energy than 2.223MeV, the extra energy goes to provide kinetic energy to freed nucleons. This result is confirmed by observations of the photo-disintegration of a deuteron. When deuterium is bombarded by gamma ray photons, it breaks up into a proton and a neutron on absorbing a photon of energy atleast equal to the binding energy:

$$E_\gamma = (m_p + m_n - m_d)c^2$$

This is shown in Fig. 13.5. Repeating the above argument we can say that when a neutron and a proton combine to form a deuteron, a small mass is found missing. Does this not suggest

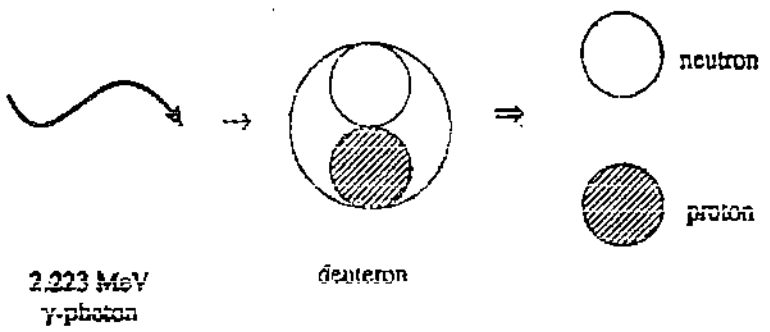


Fig. 13.5 : When deuterium is bombarded by a 2.223 MeV gamma ray photon, it breaks up into a proton and a neutron

that the binding energy can be looked upon as a direct measure of nuclear stability. To discover answer to this question, we would like you to solve SAQ 2.

Now a days, atomic masses are expressed in terms of the actual mass of ^{12}C isotope of carbon. The unit of atomic mass, abbreviated as u, is (1/12)th of the actual mass of the ^{12}C. This is equal to 1.66×10^{-27}kg. The energy equivalent of 1u is

$1\text{u} = (1.66 \times 10^{-27}\text{kg}) \times (2.998 \times 10^{8}\text{ms}^{-1})^2$
$= 14.92 \times 10^{-11}$ J
$= 931.3 \times 10^{6}$ eV
$= 931.3$ MeV

Spend 10 min.

SAQ 2

Calculate the binding energy of $^{4}_{2}He$, $^{35}_{17}Cl$, $^{56}_{26}Fe$ and $^{235}_{92}U$. Given m_n = 1.008665u, $M(^1H)$ = 1.007825u, $M(^4He)$ = 4.002604u, $M(^{35}Cl)$ = 34.968853u, $M(^{56}Fe)$ = 55.934939u and $M(^{235}U)$ = 235.043933u.

On solving this SAQ you will find that the binding energy of a nucleus is an increasing function of the mass number (28.3MeV for ^{4}He, 298MeV for ^{35}Cl, 492MeV for ^{56}Fe and 1784MeV for ^{235}U. Let us divide these binding energies by the mass numbers of particular nuclei. The binding energy per nucleon are found to be 7.1MeV, 8.5MeV, 8.8MeV and

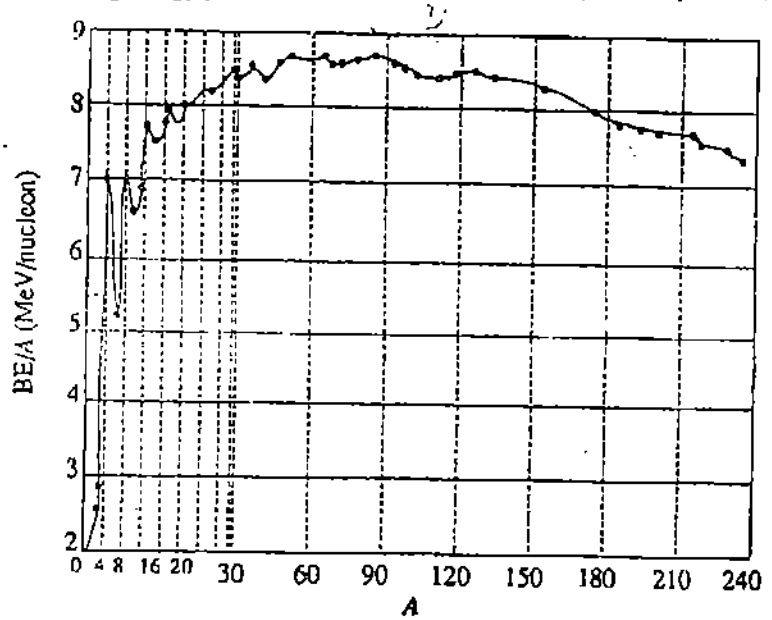


Fig.13.6: Binding energy per nucleon as a function of mass number

7.6MeV. Fig. 13.6 shows an explicit plot of binding energy per nucleon as a function of mass number. You will note that

- binding energy curve shows sharp peaks, particularly for ^{4}He, ^{9}Be, ^{12}C, ^{16}O and ^{20}Ne
- with the exception of light nuclei ($A \leq 20$), the values fall on a smooth curve
- binding energy per nucleon increases monotonically and humps in the vicinity of A = 56 (corresponding to the iron nucleus) with a maximum value of 8.8MeV
- beyond A = 56, BE/A decreases steadily droping to 7.6MeV at A = 238.

This means that nuclei at either extreme of the periodic table are less stable as compared to the nuclei in the middle. That is, BE/A, rather than BE, is a measure of stability of a nucleus.

The variation of the binding energy per nucleon with mass number hints at the possibility of tapping the energy of the nucleus. For instance, when two light nuclei fuse to produce a more stable nucleus, energy would be released. Such reactions are called fusion reactions and are responsible for the release of energy in stars. Efforts are now on to use controlled fusion reaction which, once achieved, holds the promise to meet our all future energy needs. Similarly, whenever a very heavy nucleus breaks into two parts, the binding energy per nucleon increases leading to liberation of energy. This process is called nuclear fission. The amount of energy released in fission is equal to the number of nucleons times the difference in binding energy per nucleon of the reactants and the products. For example, the binding energy per nucleon in ^{235}U is nearly 7.6MeV, whereas it is about 8.5MeV for nuclei with mass number around 120. Thus, if a ^{235}U nucleus splits into two nearly equal fragments, there would be a gain in binding energy of the system of 0.9MeV per nucleon. The total energy released in one fission event would therefore be nearly equal to $235 \times 0.9 \cong 212$MeV. The heat of combustion of a carbon atom is only about 4eV. Hence the energy released when 1kg of nucleus is fissioned is nearly the same as when about 2700 metric tons of coal are burnt.

A semi-empirical formula for the binding energy of nuclei was given by Weizsäcker by considering the similarity that exists between forces which make nucleons cling together in a nucleus and the forces that bind molecules in a liquid drop. For any nucleus containing Z protons and A nucleons, the binding energy can be expressed as

$$BE(\text{MeV}) = \alpha A - \beta A^{2/3} - \gamma \frac{(A - 2Z)^2}{A} - \delta \frac{Z(Z-1)}{A^{1/3}} \pm \frac{\epsilon}{A^{3/4}} \qquad (13.3)$$

where $\alpha, \beta, \gamma, \delta$ and ϵ are numerical constants having values $\alpha = 15.8$, $\beta = 17.8$, $\gamma = 23.7$, $\delta = 0.71$ and

$$\epsilon = \begin{cases} 34, \text{ for even - even or odd - odd nuclei} \\ 0, \text{ otherwise} \end{cases}$$

The first term represents the attractive energy of nucleons (volume energy); the second corrects for the over-estimation due to the weaker binding of nucleons near the surface and is proportional to the surface (surface effect), the third term is a negative correction due to excess neutrons (asymmetry effect) and fourth takes account of the electrostatic energy of protons (Coulomb effect). Because each charged particle in the nucleus repels all the other charged particles, this energy is negative and proportional to the number of pairs of protons,

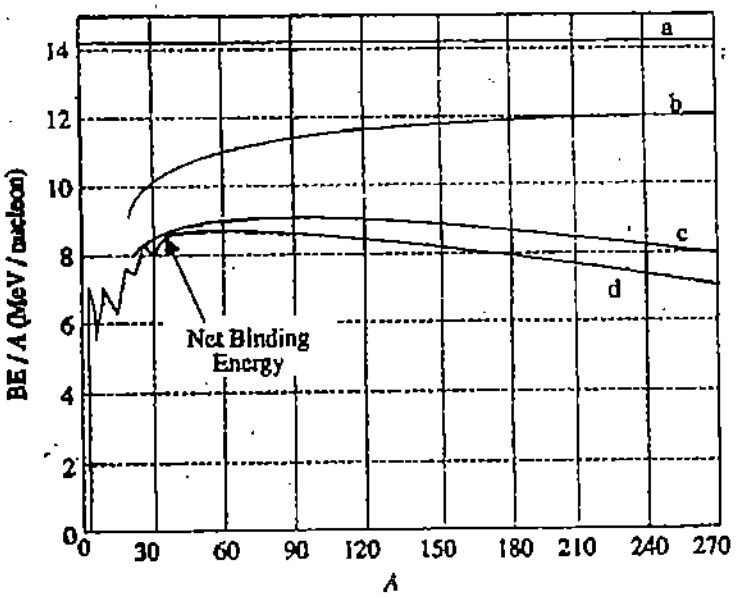


Fig. 13.7 : Plot of binding energy per nucleon as a function of mass number. Curve a represents the volume energy. Curve b represents the combined effect of volume and surface energies. Curve c represents the cumulative effect of the first, second and fourth terms in Weizsäcker's formula. When the contribution of asymmetric and spin terms is added, we get curve d for net binding energy per nucleon.

which is equal to $Z(Z-1)/2$. The last term is the spin term. This term is positive for even- even nuclei (i.e. both Z and N even), negative for odd- odd nuclei and zero for odd-even or even-odd nuclei. The relative contributions of the various terms in Weizsäcker's semi-empirical formula to binding energy per nucleon as a function of mass number are plotted in Fig. 13.7. Now we will illustrate how Weizsäcker's formula can be used to calculate BE/A.

Example 1

Using Weizsäcker's formula, calculate the binding energy per nucleon for ^{235}U. When it is made to undergo fission, suppose that $^{149}_{60}Nd$ and $^{85}_{32}Ge$ are the two fission products. Calculate BE/A for these nuclei as well.

Solution

In the case of ^{235}U, $A = 235$, $Z = 92$ and $N = 143$. Also since it is an even-odd nucleus, the contribution of spin term would be zero. Hence, using Eq. (13.3) we get

$$BE(\text{MeV}) = 15.8 \times 235 - 17.8 \times (235)^{2/3} - \frac{23.7 \times (51)^2}{235} - \frac{0.7 \times 92 \times 91}{235^{1/3}}$$

$$= 3713.00 - 677.85 - 263.31 - 963.25 = 1808.6$$

You will note that the Coulomb term (fourth) dominates the surface term (second). This is due to the large number of protons in the nucleus of ^{235}U.

The binding energy per nucleon is $= \frac{1808.6}{235} \cong 7.7\text{MeV}$.

For $^{149}_{60}Nd$, we have $A = 149$, $Z = 60$ and $N = 89$. As before, the contribution of the spin term is zero. Therefore,

$$\text{BE(MeV)} = 15.8 \times 149 - 17.8 \times (149)^{2/3} - \frac{23.7\,(29)^2}{149} - \frac{0.71 \times 60 \times 59}{(149)^{1/3}}$$

$$= 2354.20 - 500.28 - 133.77 - 474.18$$

$$= 1246.0$$

What do you observe? The Coulomb term and the surface term are nearly equal. In this case BE per nucleon comes out to be 8.4MeV showing that fission product nuclei ^{149}Nd is more stable than ^{235}U.

For the case of $^{85}_{32}Ge$, $A = 85$, $Z = 32$ and $N = 53$. The contribution of spin term is zero even here. Hence,

$$\text{BE(MeV)} = 15.8 \times 85 - 17.8 \times (85)^{2/3} - \frac{23.7 \times (21)^2}{85} - \frac{0.71 \times 32 \times 71}{(85)^{1/3}}$$

$$= 1343.00 - 344.11 - 122.96 - 160.19 = 715.7$$

Therefore, binding energy per nucleon is 8.4MeV.

You will note that in this case surface energy is more than the Coulomb energy and, of the two product nuclei, $^{85}_{32}Ge$ is more stable.

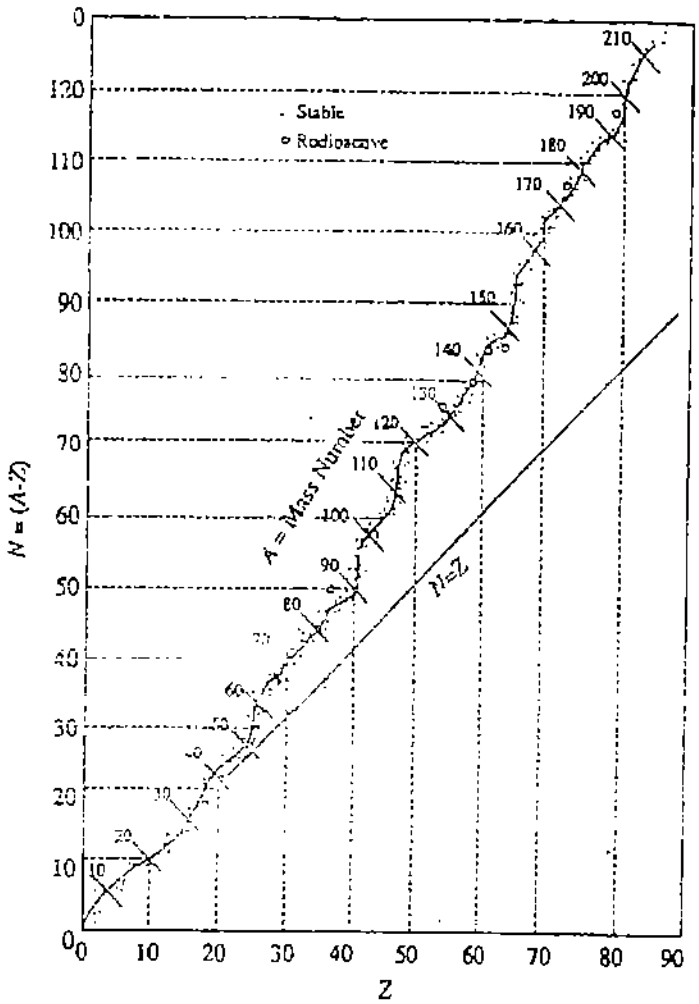


Fig.13.8: Plot of the number of neutrons versus the number of protons for naturally occurring nuclei

In the above example you have seen that stability of nuclei decreases as the difference between the number of neutrons and protons increases. It is now well known that in light stable nuclei the number of neutrons is nearly equal to the number of protons. As we move towards higher A, the neutron number increases relative to the proton number and the excess gradually increases with increasing A. And only certain combinations of protons and neutrons form stable nuclei. This is brought out in Fig. 13.8, where we have plotted the number of neutrons (ordinate) against the number of protons (abscissa) for stable nuclei (shown by solid circles). We note that for nuclei with $Z \leq 20$, the stability curve is a straight line with $Z = N$. For $Z > 20$, the stability curve bends towards $N > Z$. You can understand this on the basis of semi-empirical mass formula by answering the following SAQ.

Sp.. 12 min

SAQ 3

Show that for light nuclei, the fact $Z \approx N$ is explained by the semi-empirical mass formula.

It is an observed fact that nuclei like $^{58}_{28}$Ni, $^{50}_{22}$Ti, $^{40}_{20}$Ca, etc in which either N or Z, or both, are equal to 2,8,20,28,50,82 and 126 have some very special features (or properties) which are markedly different from those of other nuclei:

- they are more abundant in nature
- they are more stable than others

These numbers are called magic numbers. They have proved very helpful in speculating the structure of atomic nuclei.

How do nucleons cling together: Nuclear Force

Once physicists accepted the neutron-proton hypothesis of nucleus, an important question arose: How do nucleons cling together? In other words: What is the nature of force that is responsible for the binding of nucleons in a nucleus? Since gravitation and electromagnetic interactions explain most of the observed facts, you may be tempted to identify one of these forces as the likely force. The extremely small size of the nucleus, where all the protons and neutrons are closely packed, immediately suggests the existence of strong short range attractive forces to hold them together. These attractive forces cannot have electrostatic origin. Why? This is because electrostatic forces between protons are repulsive and if only these were operative, the nucleons would have been blown apart. Instead, the forces between nucleons are responsible for the large binding energy per nucleon (nearly 8 MeV) in a nucleus. Let us consider the other alternative. The force between nucleons may be gravitational since it is a force of attraction between every pair of nucleons. However, it is far too weak to account for the powerful attractive forces between nucleons. If the nucleon - nucleon force is taken to be unity, the gravitational force would be of the order of 10^{-39}. We may, therefore, conclude that the purely attractive forces between nucleons are of a new type with no analogy whatsoever with other known forces in the realm of classical physics. This new attractive force is called nuclear force.

The gravitational as well as electrostatic forces obey the inverse square law. The situation in the case of the nucleus is entirely different. All the nucleons are closely packed in the tiny nucleus like a set of marbles in a box. The force that holds the nucleons together must exist between the individual neighbouring nucleons in the nucleus. The nuclear force between nucleons is, therefore, a short range force operating over very short distances ($\sim 10^{-15}$m). The nuclear force is negligible at large distances. It suggests that each nucleon interacts only with its nearest neighbours.

Recent experimental evidences suggest that nuclear forces have a charge dependent part. However, it is quite small (<1%).

These nuclear forces must account for the attractive force between:

- a proton and a neutron
- two protons, and
- two neutrons.

Since BE/A is the same, irrespective of the mix of neutrons and protons in the nucleus, we are quite justified in considering the force between them equivalent. That is, nuclear force is charge independent.

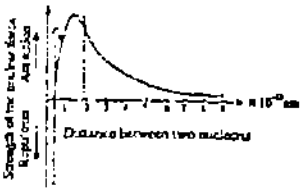
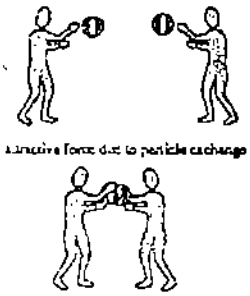


Fig. 13.9: Typical variation of nuclear forces with distance.

If nuclear forces have only attractive character, nucleons should coalesce under their influence. But we all know that the average separation between nucleons is constant, resulting in a nuclear volume proportional to the total number of nucleons. The possible explanation is that nuclear forces exhibit attractive character only so long as nucleons are seperated through a certain critical distance. For distances less than this critical value, the character of nuclear forces must change abruptly; attraction should change to repulsion. (You should not confuse this repulsion with electrostatic repulsion.) These qualitative aspects of nuclear forces are shown in Fig. 13.9.

Let us pause for a while and ask: How do nuclear forces operate between nucleons? In 1932 Heisenberg, suggested that electrons and positrons shift back and forth between nucleons. A neutron, for instance, might emit an electron and become a proton, while a proton on absorbing the electron would become a neutron. However, theoretical considerations showed that the forces resulting from electron and positron exchange by nucleons would be too small (by a factor of 10^{14}) to be significant in nuclear structure. In 1935 Japanese physicist Hideki Yukawa proposed that particles of mass in-between the masses of electrons and nucleons are responsible for nuclear forces. Now these particles are called pions. Pions may be charged (π^+, π^-) or neutral (π^0); the word pion is contraction of the original name pi-meson.



Attractive force due to particle exchange

repulsive force due to particle exchange

Fig. 13.10 : Attractive and repulsive forces can both arise from particle exchange.

According to Yukawa's theory, every nucleon continually emits and reabsorbs pions. An emitted pion can also be absorbed by another nucleon. The associated transfer of momentum is equivalent to the action of a force. One of the strengths of Yukawa's theory of nuclear forces is that it can account for their attractive as well as repulsive characters. There is no simple way to demonstrate this aspect formally. But as a rough analogy, let us imagine two boys exchanging volleyballs (Fig. 13.10).

You may now ask: If nucleons constantly emit and absorb pions, why are they not found with other than their usual masses? The answer lies in Heisenberg's uncertainty principle. We all know that the laws of physics refer to measurable quantities only and the accuracy with which certain combinations of measurements can be made is limited by the uncertainty principle. The emission of a pion by a nucleon which does not change in mass – a clear violation of the law of conservation of energy–is possible if the nucleon absorbs the same or another pion so soon afterward that even in principle it is not possible to measure any mass change. The uncertainty principle does not bar an event in which energy is not conserved for the time less than $\hbar/(2\,\Delta E)$. This condition enables us to estimate the pion mass. This is illustrated in the following example.

Example 2

Assume that a pion travels between nucleons at a speed of $v \sim c$. The emission of a pion of mass m_π, represents a temporary energy discrepancy of $E \sim m\,c^2$. Calculate m_π.

Solution

Nuclear forces have a maximum range of about 1.7 fm and the time t needed for the pion to travel this far is given by

$$\Delta t = \frac{r}{v} \sim \frac{r}{c}$$

Hence $$(m_\pi c^2) = \frac{r}{c} \sim \hbar$$

so that $$m_\pi \sim \frac{\hbar}{rc}$$

On substituting the known values we find that

$$m_\pi = \frac{1.05 \times 10^{-34}\ \text{J.s}}{(1.7 \times 10^{-15}\ \text{m}) \times (3 \times 10^8\ \text{ms})} \cong 2.1 \times 10^{-28}\ \text{kg}$$

This is about 230 times the rest mass of the electron. Yukawa's mesons were actually discovered in 1946 in cosmic rays by Powell. The rest mass of charged pions is 273 m_e and that of neutral pions is 264m_e.

13.4 NUCLEAR FISSION

You now know that intermediate mass nuclei such as barium, krypton and iron have somewhat higher binding energy per nucleon compared to heavy nuclei such as uranium, plutonium and thorium. This is essentially due to the increasing role of electrostatic repulsion. This accounts for the breaking up of a less tightly bound nucleus into more tightly bound nuclei. This process is called nuclear fission – a term coined by Lise Meitner and Otto Frisch from the analogy with biological cell division. In 1938 Otto Hahn and Fritz Strassmann established that barium was one of the products when uranium was bombarded with slow neutrons. This result defied all known tenets of nuclear physics at that time. However, these findings were reported in Nature in December 1938. Later a number of other fission fragments – from selenium to lenthanum – were identified chemically.

It was conjectured by a number of physicists that some of the fission fragments may decay by emitting neutrons. And the experimental evidence for production of such neutrons came in March 1939. As we now know, on an average, two to three neutrons are emitted per fission event. Moreover, a large amount of energy is released in this process. These immediately suggested the possibility of a neutron chain reaction – an observation with exciting practical utility for production of electricity.

Refer to Fig. 13.11. It shows the schematics of nuclear fission of ^{235}U by slow neutrons. You will observe that the (primary) fission products consist of two (middle weight) nuclei of unequal mass, 2 or 3 neutrons, a few γ-rays and nearly 200 MeV energy. It can be represented as

$$ {}^{1}_{0}n + {}^{235}_{92}U \rightarrow {}^{141}_{56}Ba + {}^{92}_{36}Kr + 3\ {}^{1}_{0}n + 200\ MeV $$

Investigations have revealed that fission event occurs within 10^{-17} s of the neutron capture and fission neutrons are emitted within about 10^{-14} s of the event.

Lise Meitner and Otto Frisch explained the results of Hahn and Strassmann on the basis of the liquid drop model of the nucleus. However, the detailed theory of fission was developed

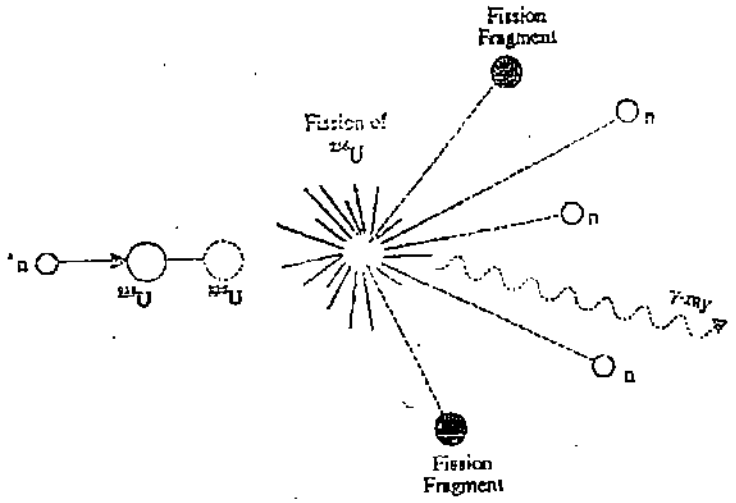


Fig. 13.11 : Schematics of nuclear fission of ^{235}U by thermal neutrons

by Bohr and Wheeler using the analogy between nuclear forces and the forces which bind molecules together a liquid. Using this model, they explained the existence of spontaneous fission and predicted the ability of ^{235}U to undergo fission more readily than ^{238}U. They also showed that if the energy of neutrons produced in fission could be reduced to about 0.025 eV, their effectiveness in causing fission of ^{235}U increases hundred-fold. Before we discuss the liquid drop model of fission, it is important to point out that the fission fragments are usually of unequal mass; one being much heavier than the other. Such a fission is said to be asymmetric. Studies have shown that ^{235}U can fission in more than 40 different ways. This means that about 80 different nuclei are directly produced in fission. The heavier fragments normally lie in the mass range 125-150 with a well defined maximum around 140, whereas the lighter fragments lie in the mass range 80 – 110 with a maximum around 95. As

examples, we may mention $^{147}_{57}La$, $^{140}_{54}Xe$, $^{135}_{52}Te$, $^{149}_{60}Nd$, $^{87}_{35}Br$, $^{94}_{38}Sr$, $^{99}_{40}Zr$, $^{85}_{32}Ge$, etc. Beside ^{235}U, ^{233}U and ^{239}Pu can also be fissioned with thermal neutrons.

13.4.1 The Liquid Drop Model of Fission

In a large nucleus containing many nucleons two main forces are operative: Coulomb replusion between protons and nuclear forces between the nucleons. Nuclear forces have short range and are charge independent. Therefore, a nucleon deep inside the nucleus is surrounded by other nucleons on all sides and the average force acting on it will be zero. However, a nucleon on the surface will be attracted only by those nucleons which are inside the nucleus and within the range of nuclear forces. Therefore, there will be a net unbalanced force directed inwards. (This is similar to the force of surface tension in a liquid.) Hence, when we treat the nucleus as a charged liquid drop, it will be spherically symmetric in its lowest state, as shown in Fig. (13.12a).

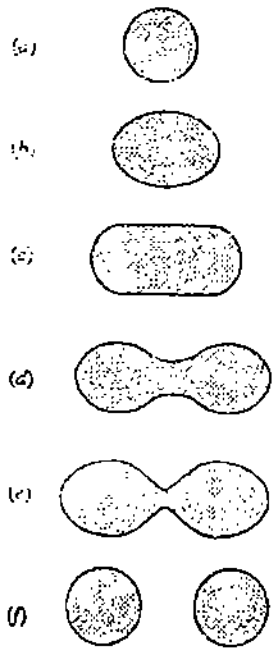


Fig. 13.12 : Fission of a nucleus according to liquid drop model

When a nucleus captures a neutron, it induces re-pairing of nucleons inside the nucleus. The compound nucleus, formed in an excited state, has an energy equal to the kinetic energy of the incident neutron and the energy liberated in re-pairing of nucleons. As a result, it undergoes strong oscillations which tend to distort its spherical shape i.e. the energetic compound nucleus is slightly elongated (Fig. 13.12b). The force of surface tension tends to restore the original shape while the Coulomb force tends to distort it further. When the energy of excitation is small, the distortions produced in the nucleus are small and the nucleus returns to the ground state by emitting γ-rays. However, if the energy gained by the nucleus is large, the elongation of the nucleus gradually develops into two bulges joined by a waist (Fig. 13.12c,d). That is, the whole system could be pushed into a bumb-bell shape. Once the separation between the two charge centres exceeds some critical value, nuclear surface tension forces lose out to electrostatic repulsion and the two halves are pushed apart. The nucleus is then irrevocably said to fission. The state e, where two charge centres are just in contact, is referred to as the scission state. The energy of the Coulomb field at e transforms into the kinetic energy of fission products, which fly away at great speed.

13.4.2 Critical Energy for Fission: Spontaneous Fission

The difference in energy between the system in the state e and the nucleus in ground state is called the critical energy of fission, denoted by E_c. To calculate E_c, we assume that

- the nucleus in the ground state is spherical
- in going from one state to another, the total volume of the system does not change, and
- the oscillations of the compound nucleus in the excited state deform the surface only.

When spontaneous fission occurs, more often that not, it is asymmetric. However, for simplicity, we consider the case of symmetric fission in which the two fragments are exactly alike in charge and mass. Thus, if the charge and mass of the original nucleus are Z and A, respectively, then both fragments will be of charge $Z/2$ and mass $A/2$ (we assume both Z and A to be even). Then, the surface energy of the initial spherical drop will be

$$E_{si} = \sigma S = 4\pi R^2 \sigma \tag{13.4}$$

where σ, S and R are the surface tension, the surface area and the nuclear radius, respectively.

The Coulomb energy of the original undeformed nucleus is

$$E_{ci} = \frac{3}{5}\frac{(Ze)^2}{R} \tag{13.5}$$

If we use the relation $R = r_0 A^{1/3}$ ($r_0 = 1.2 \times 10^{-15}$m), the energy of the nucleus in its ground state, obtained by adding the above two contributions, will be

$$E_i = E_{si} + E_{ci} = 4\pi\sigma r_0^2 A^{2/3} + \frac{3}{5}\frac{(Ze)^2}{r_0 A^{1/3}} \tag{13.6}$$

The surface energy of the two fragments in state e is

$$E_{se} = 2 \times 4\pi \left(\frac{R}{2^{1/3}}\right)^2 \sigma = 8\pi r_0^2 \left(\frac{A}{2}\right)^{2/3} \sigma \tag{13.7}$$

where $(R/2^{1/3})$ signifies the radius of each symmetric fragment. (This is a direct consequence of the second assumption.)

The Coulomb energy of these fragments at e will be a sum of two terms: one representing the (Coulomb) energy of the two undistorted fragments and the other arising from the repulsion between them:

$$E_{ee} = 2 \times \frac{3}{5} \frac{(Ze/2)^2}{(R/2^{1/3})} + \frac{(Ze/2)^2}{2 \times (R/2^{1/3})}$$

$$= \frac{3}{5} \frac{Z^2 e^2}{2r_0 (A/2)^{1/3}} + \frac{Z^2 e^2}{8r_0 (A/2)^{1/3}} \qquad (13.8)$$

The energy of fission fragments is therefore given by

$$E_f = E_{se} + E_{ee} = 8\pi\sigma r_0^2 (A/2)^{2/3} + \frac{3}{5} \frac{Z^2 e^2}{2r_0 (A/2)^{1/3}} + \frac{Z^2 e^2}{8r_0 (A/2)^{1/3}} \qquad (13.9)$$

The difference in the energy of the original nucleus and the energy of two equal fragments in contact is, by definition, the critical energy for fission:

$$E_c = E_f - E_i = 4\pi r_0^2 A^{2/3} \left[(2^{1/3} - 1) - \frac{3}{40\pi} \frac{Z^2 e^2}{r_0^3 A\sigma} \left(2 - 2^{1/3} - \frac{5}{12} \times 2^{1/3} \right) \right]$$

$$= \pi\sigma r_0^2 A^{2/3} \left[1 - \frac{3}{40\pi} \frac{Z^2 e^2}{r_0^3 A\sigma} \right] \qquad (13.10)$$

since $(2^{1/3} - 1) = 0.260$ and $\left(2 - 2^{1/3} - \frac{5}{12} \times 2^{1/3} \right) (2^{1/3} - 1)^{-1} \cong 1$.

Let us now discuss the conditions for the stability of a nucleus against spontaneous fission.

For spontaneous fission to occur, E_c must be zero. Then Eq. (13.10) reduces to a compact form:

$$1 = \frac{3}{40\pi} \frac{Z^2 e^2}{r_0^3 A\sigma}$$

or

$$\left(\frac{Z^2}{A} \right)_{SF} = \frac{40\pi r_0^3 \sigma}{3e^2} = 2 \times \frac{4\pi r_0^2 \sigma}{(3/5)(e^2/r_0)} \qquad (13.11)$$

This equation defines the limiting value of Z^2/A for stability of nuclei against spontaneous fission. From Weizsäcker's semi-empirical relation we recall that $4\pi r_0^2\sigma = \beta = 17.8$ and $(3/5)e^2/r_0 = \delta = 0.71$. Using these values in Eq. (13.11), we get

$$\left(\frac{Z^2}{A} \right)_{SF} = 50.1 \qquad (13.12)$$

A more generally accepted value of $(Z^2/A)_{SF}$ is 47.8. For ^{238}U, $Z^2/A = 35.56$ and for ^{239}Pu it is 36.97. From this you may conclude that within the framework of this theory, even in the case of the heaviest naturally occurring element, spontaneous fission is a rare occurrence (one in million).

In the above discussion of the liquid drop model of fission, we have made some simplifying assumptions. All of them are not always justifiable. Using this theory one cannot, therefore, account for all the observed properties of fission. This has led to several modifications of this theory and we now understand the mechanism of fission far better.

13.5 NUCLEAR MODELS

To have an idea about the structure of atomic nuclei, physicists had to resort to various conceptions in which the picture of the nucleus varied from a gas to a drop of liquid, and as a solid entity. Even as a solid, they visualised it in many forms and shapes. For example, they pictured it as a loosely bound structure, as a strongly bound structure, and as a structure where nucleons form groups (of a few) at a time. They also pictured it to be spherical in

shape, ellipsoidal in shape, as a pear in shape, and so on. These different 'imaginations' led to models like the Fermi Gas Model, the Collective Model, the Liquid Drop Model, the Shell Model and the Cluster Model, among others.

Different models had varying degrees of success in explaining/predicting the properties of atomic nuclei. The liquid drop model has had only a limited success; it has been able to explain only the binding energies of nuclei and the phenomenon of nuclear fission (where heavier nuclei break-up into two or more approximately equal parts). Because of its quite limited success, we are not discussing it beyond what is already done. However, we discuss two other models.

13.5.1 The Shell Model

The existence of magic numbers in nuclei led to the development of the shell model. According to this model, the neutrons and protons are distributed inside the nucleus in certain specified orbits (or paths), like the electrons in an atom. The maximum number of nucleons in any path is restricted by Pauli's Exclusion Principle (just as in the case of an atom, Unit 5).

Any given orbit can have equal number of neutrons and protons, as Pauli's exclusion principle applies to both these separately (Fig. 13.13). The orbits are designated by their nlj values. Here n refers to the principal quantum number of the orbit, l to the orbital angular momentum, and j to the total angular momentum of a nucleon. The number of allowed neutrons or protons in an orbit is equal to $2j+1$.

A group of orbits, quite farther apart from another such group, is usually spoken of as a shell. A number of shells form a shell type of a structure inside the nucleus. Whenever a shell is full to its maximum allowed capacity, it "sort of" becomes inert (not contributing to nuclear properties) and gives rise to a magic number.

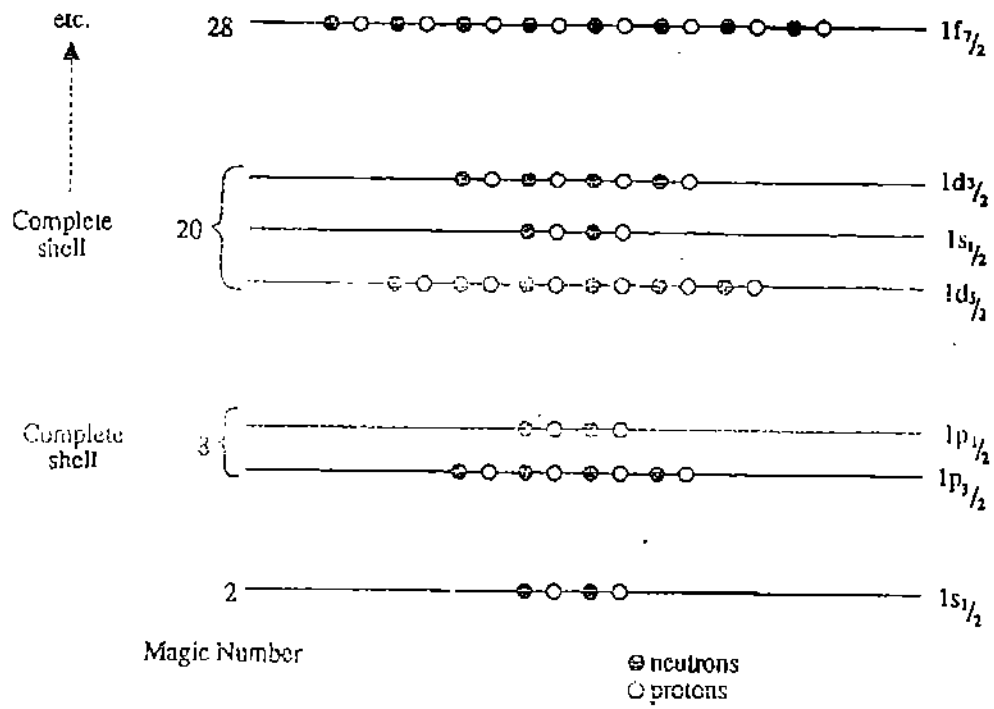


Fig.13.13: Representation of shell structure of nuclei. The straight lines signify orbits which may be circular or nearly circular in shape.

Fig. 13.14: Two types of motion of a nucleon. Circular motion of a nucleon of mass m moving with a velocity V in a orbit of radius r is ($V = r\omega$). A nucleon in a nucleus has two angular momenta: orbital (l) and spin (s). The total angular momentum j of a nucleon is their vector sum.

The neutrons and protons in orbits are assumed to execute two kinds of motion. A rotatory motion around some common centre of force and a spin motion around themselves. These are illustrated in Fig. 13.14. These two motions are similar to the two different types of motion of the earth: one around the sun and the other about its own axis.

The structure of a nucleus, on the whole, is very similar to that of an atom. The only difference is that in an atom the electrons revolve around a common centre of force – the nucleus, while in a nucleus no such common centre of force exists.

It is observed that nuclei with N and/or Z equal to magic numbers are spherical in shape. For nuclei which are spherical or nearly spherical in shape, the shell model is extremely successful in explaining and predicting their properties.

13.5.2 The Collective Model

Away from the magic numbers, the nuclei become non- spherical or ellipsoidal in shape. In such nuclei, the nucleons in the last unfilled (or incomplete) shell form clusters and deform the nuclear shape. This deformation changes its orientation in space as a function of time, leading to rotation of the nucleus as a whole (Fig. 13.15). Rotations of nuclei are indicative of the collective behaviour of nucleons.

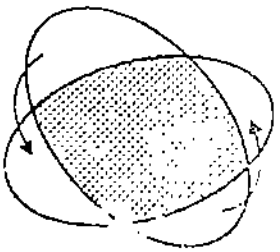


Fig.13.15: Rotation of an ellipsoidal nucleus. Nuclear matter contained in the shaded region is not supposed to rotate. Arrow heads indicate how the matter moves giving rise to a rotational behaviour.

Not so far away from the magic numbers, the nuclei exhibit surface vibrations about their equilibrium shapes (Fig. 13.16). These vibrations have a fixed frequency and exhibit simple harmonic character. The vibrations arise because of the continuous formation and dissolution of clusters. Vibrations are also indicative of the collective behaviour of nucleons.

The collective model has been extremely successful in predicting and explaining the properties of nuclei which are away from the magic numbers. This model has been particularly successful in explaining their quadrupole moments.

The shell model assumes that the nucleons within a nucleus act more or less independently. The collective model, on the other hand, assumes that the nucleons act collectively. Neither of these two models is able to explain the structure of nuclei throughout the periodic table. The shell model is successful for nuclei with N and/or Z values near the magic numbers. Collective model is successful for nuclei away from the magic numbers.

Since no model has been able to explain the structure of nuclei throughout the periodic table, an attempt is being continuously made to unify (combine) the ideas of these two models into one model. Though there has been some success, no concrete model has emerged so far which explains the structure of all nuclei and satisfactorily explains/predicts their properties.

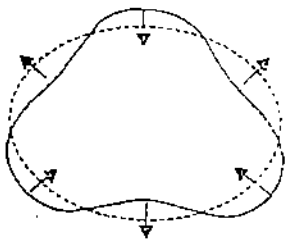


Fig.13.16: Vibrations of a not so deformed nucleus. The dashed figure is the equilibrium shape. Arrow heads indicate how the crests and troughs move with time giving rise to a vibrational behaviour.

13.6 SUMMARY

- Alpha particle experiment of Geiger and Marsden revealed that (i) nearly all alpha particles emerged without deviation suggesting that atoms consist largely of empty space (ii) large angle alpha particle scattering implied that they undergo a head-on collision with something immovable - the nucleus.
- The distance of closest approach of an alpha particle to the nucleus of an atom can be calculated using the relation

$$b = \frac{1}{4\pi\varepsilon_0}\,\frac{4Ze^2}{mv^2}$$

- Rutherford analysed the observations of Geiger and Marsden. He proposed that all the mass and the positive charge of an atom is concentrated in a tiny nucleus and electrons reside outside the nucleus.
- Nuclei of all elements have the same density, of the order of 10^{17}kg m^{-3}.
- The difference between the observed mass of any nucleus and its constituent nucleons is termed as the mass defect:

$$\Delta m = Zm_{\mathrm{H}} + Nm_{\mathrm{n}} - M$$

The energy equivalent of mass defect is called binding energy:

$$\mathrm{BE} = \Delta mc^2$$

- The binding energy (BE) of a nucleus is an increasing function of mass number. Binding energy per nucleon (BE/A) versus A graph shows (i) sharp changes for ^{4}He, ^{9}Be, ^{12}C, ^{16}O, ^{20}Ne etc (ii) monotonic increase beyond A = 20 and BE/A attains a maximum value of 8.8MeV for the iron nucleus (iii) gradual decrease for nuclei heavier than those of iron and drops to 7.6MeV for uranium. BE/A rather than BE is a measure of stability of nuclei.
- Nucleons cling together via nuclear forces, which show saturation, are charge

independent and attractive. However, they become repulsive when nucleon-nucleon distances are less than a critical value.

- According to Yukawa, pi-mesons, now called pions, are continuously emitted and absorbed by every nucleon. In this process, some momentum transfer takes place, which is responsible for the action of nuclear forces.
- When a heavy nucleus is bombarded by slow neutrons, its shape undergoes a series of changes before splitting into nuclei of unequal mass. Every fission event is accompanied by release of 2 or 3 neutrons and nearly 200MeV energy, which heats up the surrounding medium.
- The shell model of a nucleus is analogous to the electronic shell model of atom. It tells us that within the nucleus, nucleons are distributed in certain specified orbits. The maximum number of nucleons in any path is restricted by the Pauli's Exclusion Principle. The nuclei with filled nuclear shells correspond to magic nuclei and are more stable.
- In nuclei away from the magic numbers, the nucleons in the last unfilled shell form clusters and change its orientation in space and time leading to its rotation as a whole.

13.7 TERMINAL QUESTIONS

1. Calculate the distance of closest approach of an α- particle of energy 5.3MeV fired directly towards a nucleus of gold ($Z = 79$). Given, mass of a gold nucleus $(M) = 6.7 \times 10^{-27}$kg, charge of an electron $= 1.6 \times 10^{-19}$C, $\frac{1}{4\pi\epsilon_0} = 9 \times 10^9 \text{N m}^2 \text{C}^{-2}$.
2. Find the energy released when two ${}^2_1\text{H}$ nuclei are fused together to form a ${}^4_2\text{He}$ nucleus. The binding energy per nucleon of ${}^2\text{H}$ is 1.1MeV and that of ${}^4\text{He}$ is 7.0MeV.
3. Calculate the binding energy of a ${}^5\text{He}$ nucleus on the basis of semi-empirical mass formula and compare it with the value obtained on the basis of mass defect (SAQ 2).
4. From the semi-empirical mass formula as given by Eq. (13.3), calculate the value of the atomic number (Z) for the most stable nucleus at a given mass number. Calculate Z_0 for $A = 56$.

13.8 SOLUTIONS AND ANSWERS

SAQs

1. $$d_C = \frac{M_C}{\frac{4\pi}{3} R_C^3} = \frac{19.92 \times 10^{-27} \text{ kg}}{\frac{4}{3} \times 3.1416 \times (2.7 \times 10^{-15} \text{ m})^3}$$

$$= \frac{19.92}{82.45} \times 10^{18} \text{ kg m}^{-3}$$

$$= 2.42 \times 10^{17} \text{ kg m}^{-3}$$

$$d_{Pb} = \frac{M_{Pb}}{\frac{4\pi}{3} R_{Pb}^3} = \frac{3.4 \times 10^{-25} \text{ kg}}{\frac{4\pi}{3} (7 \times 10^{-15} \text{ m})^3}$$

$$= \frac{3.4 \times 10^{30} \text{ kg m}^{-3}}{\frac{4\pi}{3} \times (7)^3}$$

$$= 2.37 \times 10^{17} \text{kg m}^{-3}$$

2. From Eq. (13.2) we know that

$$\text{BE} = \Delta m\, c^2$$

where $\Delta m = Zm_H + Nm_n - m$.

For $^4_2$He, we have $Z = 2$ and $N = 2$. On substituting the given values, we get

$$\Delta m(^4\text{He}) = 2 \times (1.007825\text{u}) + 2 \times (1.008665\text{u}) - 4.002604\text{u}$$
$$= 4.03298\text{u} - 4.002604\text{u} = 0.030376\text{u}$$

and

$$\text{BE} = 0.030376 \times (1.66 \times 10^{-27}\text{kg}) \times (2.998 \times 10^8\text{ms}^{-1})^2$$
$$= 0.030376 \times 14.92 \times 10^{-11}\text{J}$$
$$= 0.030376 \times (14.92 \times 10^{-11}\text{J})/(1.602 \times 10^{-19}\text{JeV}^{-1})$$
$$= 0.030376 \times 9.313 \times 10^8\text{eV}$$
$$= 2.829 \times 10^7\text{eV}$$
$$= 28.3\text{MeV}$$

For $^{35}_{17}$Cl, we have $Z = 17$ and $N = 18$. Therefore,

$$\Delta m(^{35}\text{Cl}) = 17 \times (1.007825\text{u}) + 18 \times (1.008665\text{u}) - 34.96885\text{u}$$
$$= 17.133025\text{u} + 18.15597\text{u} - 34.96885\text{u}$$
$$= 0.320145\text{u}$$

Since 1u = 931.3MeV, we find that

$$\text{BE}\,(^{35}\text{Cl}) = 298.2\text{MeV}$$

For $^{56}_{26}$Fe, $Z = 26$ and $N = 30$. Therefore

$$\Delta m\,(^{56}\text{Fe}) = 26 \times (1.007825\text{u}) + 30 \times (1.008665\text{u}) - 55.934932\text{u}$$
$$= 26.20345\text{u} + 30.25995\text{u} - 55.934932\text{u}$$
$$= 0.528468\text{u}$$

and

$$\text{BE}(^{56}\text{Fe}) = 492.2\text{MeV}$$

For ^{235}U, $Z = 92$, $N = 143$ so that

$$\Delta m = 92 \times (1.007525\text{u}) + 143 \times (1.008665\text{u}) - 235.043933\text{u}$$
$$= 92.7199\text{u} + 144.239095\text{u} - 235.043933\text{u}$$
$$= 1.915062\text{u}$$

so that

$$\text{BE}(^{235}\text{U}) = 1783.5\text{MeV}$$

3. For a given A, the ratio $\frac{N}{Z}$ will be such that the total energy E tends to a minimum. Since $m_p \cong m_n$, the only terms which we have to consider in discussing this minimum for a particular value of A are $\frac{(N-Z)^2}{A}$ and $\frac{Z^2}{A^{1/3}}$. The first term demands $N = Z$ while the second term demands Z to be as small as possible. This is consistent with the fact that light nuclei have $Z \cong N$.

Qs

1. The distance of closest approach is given by

$$b = \frac{1}{4\pi\epsilon_0}\ \frac{4Ze^2}{mv^2} = \frac{1}{4\pi\epsilon_0}\ \frac{2Ze^2}{E}$$

where E is the energy of the α-particle.

Here $\frac{1}{4\pi\epsilon_0} = 9 \times 10^9\ \text{Nm}^2\ \text{C}^{-2}$, $Z = 79$, $e = 1.6 \times 10^{-19}$ C and

$E = 5.3\text{MeV} = 5.3 \times 1.6 \times 10^{-13}$J.

Substituting these values in the above expression, we get

$$b = \frac{(9 \times 10^9\ \text{Nm}^2\ \text{C}^{-2}) \times 2 \times 79 \times (1.6 \times 10^{19}\ \text{C})^2}{5.3 \times 1.6 \times 10^{-13}\ \text{J}}$$

$$= 4.3 \times 10^{-14}\text{m}$$

2. B.E. of a $_2^4$He nucleus $= 7.0 \times 4 = 28\text{MeV}$

B.E. of a $_1^2$H nucleus $= 1.1 \times 2 = 2.2\text{MeV}$

$\therefore$ Mass of $_2^4$He nucleus $= 2m_p + m_n - 28.0\text{MeV}$

Mass of $_1^2$He nucleus $= m_p + m_n - 2.2\text{MeV}$

Energy released in the fusion reaction

$E = 2 \times$ Mass of $_1^2$H - mass of $_2^4$He

$$= 2(m_p + m_n - 2.2) - (2m_p + 2m_n - 28.0) = 23.6\text{MeV}$$

3. For $_2^4$He, $A = 4$ and $Z = 2$ and $\varepsilon = 34$

Hence from Eq. (13.3) we get

$$\text{BE (MeV)} = 15.8 \times A - 17.8 \times 4^{2/3} - 0.71 \times \frac{2}{4^{1/3}} + \frac{34}{4^{3/4}}$$

$$= 63.2 - 44.9 - 0.895 + 12.02 = 29.43$$

This is slightly greater than the value calculated on the basis of mass difference of helium nucleus.

4. For the most stable nucleus at a given mass number A, we have

$$\left(\frac{\text{d BE}}{\text{d}Z}\right)_A = 0 \text{ for } Z = Z_0$$

Hence we have from Eq. (13.3)

$$-2\gamma\ \frac{(A - 2Z_0)}{A}\ (-2) - \frac{\delta}{A^{1/3}}\ (2Z_0) = 0$$

$$\Rightarrow\ 4\gamma\ (A - 2Z_0) - \delta A^{2/3}\ (2Z_0) = 0$$

or $$4\gamma A - 8\gamma Z_0 - 2\delta A^{2/3} Z_0 + \delta A^{2/3} = 0$$

or $$Z_0\ (8\gamma + 2\delta A^{2/3}) = 4\gamma A + \delta A^{2/3}$$

$$\Rightarrow\ Z_0 = \frac{4\gamma A + \delta A^{2/3}}{8\gamma + 2\delta A^{2/3}}$$

$$= \frac{23.7 \times 4A + \delta A^{2/3}}{8 \times 23.7 + 2 \times 0.71\ A^{2/3}}$$

$$= \frac{94.8A + A^{2/3}}{1.42A^{2/3} + 189.6}$$

For $A = 56$

$$Z_0 = \frac{94.8 \times 56 + 0.71 \times 56^{2/3}}{1.42 \times 56^{2/3} + 189.6} = \frac{5319.2}{210.4}$$

$$= 25.3 \Rightarrow 26$$

UNIT 14 APPLIED NUCLEAR SCIENCE

Structure

14.1 Introduction
Objectives
14.2 Self-Sustained Chain Reaction
14.3 Nuclear Reactors
Classification of Reactors
General Features of a Reactor
14.4 Nuclear Radiations in Our Environment
14.5 Radioisotopes in Everyday Life
14.6 Summary
14.7 Terminal Questions
14.8 Solutions and Answers

14.1 INTRODUCTION

In Unit 13 you have learnt how Rutherford's investigations on transmutation in 1919 led to a new and fertile branch of physics - Nuclear Physics. Over the next fifty years its frontiers expanded rather rapidly. The enormous research was rewarded in the form of numerous applications of nuclear radiations and radioisotopes in various facets of human activity. However, we shall confine ourselves to typical examples in nuclear energy generation, medicine, agriculture, industry and research.

No power is as expensive as no power. These words of the architect of Indian Nuclear Programme - Homi Bhabha - at the 3rd UN Conference on peaceful uses of atomic energy held in 1964 vividly bring out the importance of energy for the people of under developed nations.

Energy is fundamental for our existence. We need energy to cook food, run machines or go from one place to another. We also need energy for agriculture, industry, communication, better living standards and economic growth. Moreover, in future, the ever growing population, our unending quest for comfort and material affluence and issues of security will raise energy demand to unprecedented levels. We all know that energy cannot be created out of nothing. We have to burn wood, coal, gas or oil to produce energy. And at the present rate of consumption our known limited reserves of such fossil fuels are bound to dwindle very soon. Moreover, increased reliance on them could even threaten to disturb ecological balance by adding to greenhouse effect, air pollution, ozone depletion and acid rain etc. The inevitable trend in energy supply, therefore, is to move away from fossil fuels. In this diversification, efforts are on to tap hydel power. The solar dream is yet to be weaved for commercial purposes. However, nuclear energy, presents itself as a practical alternative, which can cater to our immediate as well as long term energy needs.

Presently most of the nuclear energy has come from fission reactors in which fissile isotopes ^{235}U, ^{233}U and ^{239}Pu have been used. But public acceptability of these reactors has been rather low. It has now been realised that the growing electrical power requirements of the world in the 21st century can be met only by fusion energy. However, the choice between nuclear fusion and fission as the likely energy source will probably hinge upon the public perception as to which system is safer and eco-friendly.

When two light nuclei colliding in plasma coalesce and form heavier nuclei, a large amount of energy is released. This process is referred to as nuclear fusion. The device in which such thermonuclear reactions occur in a controlled manner is called the controlled thermonuclear reactor (CTR) or a fusion reactor. It is worthwhile to note that while nuclear fuel used in a fission reactor is limited, the likely ingredients of a fusion reactor — deuterium, lithium, helium etc. — occur in nature in enormous amounts. Therefore, successful operation of a CTR would signify that the world has acquired a practically inexhaustible source of energy. Further nuclear fusion seems attractive as it bye-passes many of the safety problems associated with nuclear fission. Also, it is 'clean' in the sense that it does not leave behind harmful radioactive wastes. Moreover, the energy released per unit mass of the reacting material is much greater than the corresponding quantity in nuclear fission involving heavy nuclei. For instance, energy released per nucleon in D-T fusion is nearly five times larger

than the energy released per nucleon in a fission event (though the energy released per fusion 17.6 MeV is much less than the available energy of about 200MeV in a fission reaction).

In the previous unit we discussed the fission process in some detail. You learnt that the average number of neutrons emitted in each fission event is greater than two and these neutrons are capable of causing further fissions. This immediately presented the exciting possibility of maintaining a fission chain reaction in which each fission event removes one neutron and replaces that by more than two. When the rate of production of neutrons equals the rate of loss of neutrons, the reaction is said to be **self-supporting** or **self-sustained**. The device designed to maintain a **self-sustained and controlled chain reaction** is called a **nuclear reactor**. In Sec. 14.2 you will learn how a self-sustained chain reaction can be achieved and what are the principal components that go to make a nuclear reactor.

Nuclear reactors can be classified according to the purpose for which a reactor is used or the mean energy of neutrons causing fission. A reactor can be used for power generation, basic research and conversion of one material into another. To give you a feeling of power reactors, we have briefly discussed Indian power programme in Sec. 14.3. Research reactors can be used for producing radioisotopes, which find so many and so varied applications in everyday life. You will learn some of their important applications in industrial as well as basic processes, agriculture, medical diagnosis and treatment.

Objectives

After going through this unit, you will be able to

- explain the term self-sustained chain reaction
- identify similarities between fossil fuel plants and nuclear power plants
- explain the function of various components of a nuclear reactor
- state essential differences between a research reactor and a power reactor, and
- describe the uses of radioisotopes in industry, agriculture and medicine.

14.2 SELF-SUSTAINED CHAIN REACTION

When the question of establishing a fission chain reaction was examined for the first time by Fermi, natural uranium was taken as fuel. He immediately ran into difficulties. The abundance of ^{238}U in natural uranium is 99.7% whereas that of ^{235}U is 0.3%. Of these, ^{238}U absorbs neutrons and does not necessarily undergo fission (except by high energy neutrons). It was experienced that, on an average, more neutrons are removed from the system than are created in fission. Hence, even if a fission reaction is started by the use of an external neutron source, no sooner is the source removed than the number of neutrons in the system begin to decrease rapidly and the chain reaction stops. On the other hand, the probability of ^{235}U to undergo fission is very large for low energy neutrons. This led Fermi to suggest that much more effective use of this isotope could be made if the energy of fission neutrons is reduced through collisions in a moderating medium before they had a chance to interact with other uranium nuclei. Fermi and his coworkers suggested that

- graphite should be used as a moderator to slow down neutrons born in fission
- uranium rods should be distributed uniformly in a moderator medium.

According to Laura Fermi, the news of having attained a self sustained nuclear chain reaction to the world outside was conveyed by Prof. A. Compton through Mr. Conant of the Office of Scientific Research and Development at Harvard. To quote her: "The Italian Navigator has reached the New World", said Compton as soon as he got Conant on the line. "And how did he find the natives?" "Very friendly."

When these two innovations were incorporated, Fermi succeeded in achieving the first self-sustained chain reaction on December 2, 1942 at the University of Chicago, USA.

For a fission chain reaction to be self-sustained, the rate of production of neutrons must be equal to the rate of loss of neutrons. Neutrons can be lost from a fission chain by basically two processes. They are either absorbed by nuclei of the medium or they escape from the system. The rate of loss by leakage depends on the shape and size of the assembly. The most efficient shape is a sphere. For a spherical assembly, the leakage rate is proportional to the surface area and the rate of production is proportional to the volume (assuming that uranium is distributed uniformly throughout the assembly). Therefore, the proportion of neutrons that leak out will decrease as the size of the sphere increases. Thus, if we start with a small assembly, the rate of loss exceeds the rate of production and a self-sustained chain reaction will not be possible. As the size is gradually increased, we will reach a size when a

self-sustained chain reaction just becomes possible. This size of the sphere of fissile material constitutes the **critical size**. The corresponding mass of the fuel is termed as the **critical mass**. For ^{235}U the critical sphere has a diameter of 18cm and the corresponding critical mass is 53kg. For ^{239}Pu, the critical mass is about one-third of this.

The behaviour of a neutron chain reaction is described quantitatively in terms of a constant called the **multiplication factor**. It is denoted by k and is defined as the ratio of the number of neutrons in any one generation to the corresponding number in the immediately preceding generation, i.e.

$$\text{Multiplication factor,} \quad k = \frac{\text{Number of neutrons in } (n+1)\text{th generation}}{\text{Number of neutrons in } n\text{th generation}}$$

If $k > 1$, the number of neutrons increases indefinitely with time and the chain reaction will get out of control. If $k < 1$, the number of neutrons decreases with time and the chain reaction will stop. In the special case when $k = 1$, the number of neutrons in any two successive generations remains the same and the chain reaction proceeds at a constant rate, independent of time. Depending upon whether $k > 1$, $k = 1$ or $k < 1$, the reactor is said to be **super-critical, critical** or **sub-critical**, respectively. For this reason, k is also very frequently referred to as the **criticality factor**. A knowledge of the multiplication factor for a given reactor is essential for reactor control.

An aproximate expression for the time behaviour of a neutron chain reaction can be derived as follows: Let us suppose that at any instant of time there are N_1 neutrons. We will consider this as the first generation. Then, in the nth generation, the number of neutrons will be

$$N_n = N_1 k^{n-1} \tag{14.1}$$

Let l be the average time between the creation of a neutron and its final removal either by leakage or absorption. Then, the time elapsed between the start of the first generation (which is our zero of time) and the birth of nth generation neutrons will be $t = (n-1)l$. On substituting for $(n-1)$ from this expression in Eq. (14.1), we get

$$N(t) = N(0)\, k^{t/l} = N(0)\exp\left(\frac{t}{l} \ln k\right) \tag{14.2}$$

where we have replaced N_n by $N(t)$ and N_1 by $N(0)$. Assuming k to be close to unity and noting that $\ln k \cong k - 1$, we have

$$N(t) \cong N(0) \exp\left[(k-1)\frac{t}{l}\right] \tag{14.3}$$

This tells us that neutron number increases exponentially with time for $k > 1$ and decreases exponentially with time for $k < 1$. For $k = 1$ we have a steady state.

For prompt neutrons the mean life time is nearly 10^{-3}s. For $k = 1.01$, the increase in the number of neutrons per second is $N(t)/N(0) = \exp(0.01/0.001) = e^{10} \cong 22000$. Such a reactor would be impossible to control by any mechanical means. Thus, if the above considerations were the only ones operative, even an excess of 0.01 (1%) over unity in the value of k would make a reactor to explode. Fortunately the presence of delayed neutrons appreciably increases the neutron life time (to about 0.1s) and nuclear reactors become controllable.

Neutrons emitted within 10^{-14}s of the fission event are said to be **prompt neutrons**. However, some neutrons are emitted by fission product nuclei considerably after the fission event. These are known as **delayed neutrons**.

14.3 NUCLEAR REACTORS

Ever since the world's first nuclear reactor was constructed by Fermi and his co-workers at the University of Chicago, USA, in 1942, a variety of nuclear reactors have been built, primarily to meet our increasing demands of energy. Nuclear reactors are highly complex installations and great care has to be exercised in designing them, in spite of the fact that they work on a very simple principle. The heat generated in fission is removed by circulating a fluid, called coolant, around the fuel and this heat is subsequently used to generate high pressure and high temperature steam. This is fed to a turbine-generator system to produce electricity. In reactors built primarily for experimental purpose, the heat is discharged into a river or the sea.

14.3.1 Classification of Reactors

Reactors have been built for a wide range of uses, from power generation to testing new reactor components. And the design specifications vary widely depending upon the purpose. One can classify reactors in different ways but two most important criteria are:

- the mean energy of neutrons causing most of the fissions, and
- the purpose for which a reactor is meant.

According to the mean energy of neutrons causing most fissions, reactors can be classified broadly as: fast and thermal. In a fast reactor most of the fissions are caused by neutrons having energies of the order of a few hundred keV, whereas in a thermal reactor most fissions are induced by neutrons of energy around 0.0253eV. In India thermal reactors are the only ones that have been built so far for power generation. Fast reactors, which hold great promise for future, form third stage of Indian Nuclear Programme.

According to the purpose for which they are meant, reactors can be classified as research reactors and power reactors.

Reactors built for basic research in various branches of science, for testing new reactor designs or new reactor components, for producing radioisotopes and for medical purposes (neutron therapy) are referred to as **research reactors**. Cirus, Apsara, Purnima and Dhruva at Trombay are some of our thermal research reactors and FBTR at Indira Gandhi Centre for Atomic Energy Research, Kalpakkam is a breeder research reactor. It may be remarked here that usually a single research reactor simultaneously serves many of these purposes.

Reactors built to produce power are referred to as **power reactors**. Today all our (Indian) nuclear power (~ 2000MW) comes from thermal reactors based on the fission of ^{235}U.

Reactors designed to convert one isotope (^{232}Th or ^{238}U) into another more useful isotope (^{233}U or ^{239}Pu) are called **converters**. If the amount of newly produced fissible isotope is more than what is burnt in maintaining the chain reaction, they are called **breeders**.

From the above discussion you may be tempted to think as to what is common in different types of reactors. But this is not true! There are certain features common to all reactors and before we discuss typical reactors, it will be useful to summarise them.

14.3.2 General Features of a Reactor

All nuclear reactors consist of the following basic components: reactor core, reflector, reactor vessel, radiation shield, structural materials, coolant loops and heat exchangers, etc. These are schematically illustrated in Fig. 14.1. In a fast reactor, a blanket is placed between the core and the reflector. Let us discuss these in detail.

Core

The central region of a reactor where fission takes place, resulting in the release of energy, is known as the core. In fast reactors it contains a nuclear fuel, a coolant, control rods and structural materials. In thermal reactors a moderator is also present. An ideal fuel should have high thermal conductivity and high melting point, high resistance to radiation damage, and be chemically inert. Moreover, it should be easy to fabricate and corrosion resistant. Usually the fuel is in the form of a ceramic, i.e., either an oxide or a carbide or a nitride. (Oxy-carbides and nitrides of uranium have also been proposed as nuclear fuels.)

To prevent escape of fission fragments into the coolant or the moderator (if present), apart from providing protection to the fissionable material from corrosion, the fuel rods are cladded or canned. An ideal fuel clad material is highly resistive to corrosion, a poor neutron absorber, cheap and readily available. It should also have good mechanical strength and high melting point. Zirconium, steel, aluminium, magnesium, nickel and some other similar materials are used for this purpose. Of these all, zirconium is the best and one generally uses zirconium alloy, known as Zircalloy-2(Zr-2), in thermal power reactors. In fast reactors, stainless steel is used. Aluminium is used mainly in research reactors. A single cladded unit of fuel is known as the fuel element. (In a fast reactor the fuel element is thinner than that used in a thermal reactor and is called a fuel pin.) Several such fuel elements when put together constitute a fuel assembly. (It is this single unit which is inserted or taken out of the reactor as a whole.) A reactor core contains a large number of such fuel assemblies arranged in the form of a regular lattice. The lattice is usually square or hexagonal in form.

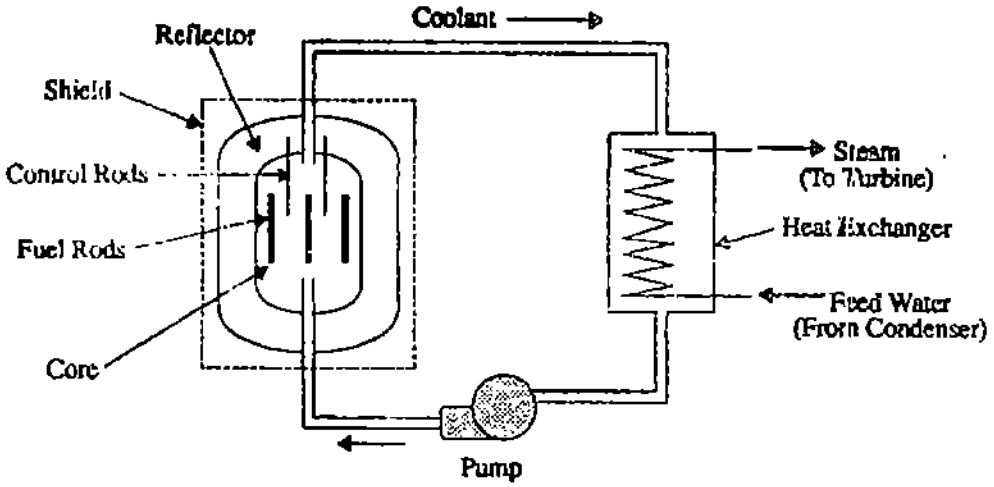


Fig.14.1: Schematic diagram of a nuclear reactor depicting its general features

To remove fission heat from the core (as also heat from any other part of the reactor), it is necessary to circulate a fluid - a liquid or a gas - through the reactor. This fluid is referred to as the **coolant**. A reactor coolant should have high thermal capacity, low cost, low neutron absorption cross section, good radiation and thermal stability, and compatibility with fuel and clad. In fast reactors, the coolant should further be a poor moderator. In these reactors liquid sodium or helium is used as coolant. In thermal reactors ordinary water, steam, heavy water, carbon dioxide, helium and some organic gases are frequently used as coolants.

To slow down neutrons born in fission, a moderator is also present in the core of a thermal reactor. It is a material of low mass number, large scattering cross section and small absorption cross section. (Other requirements for a moderator are almost the same as those for a coolant.) The best moderators are heavy water, light water and graphite. Of these, heavy water and graphite are so good that in their presence even natural uranium can sustain a chain reactor. To ensure safe operation of a reactor at a desired power level, to start up and to shut down a reactor (during routine operation or in an emergency) as well as to compensate for the gradual production of poisons and fuel burn up etc. provision has to be made in the reactor core to control the multiplication factor. This is achieved by having some control rods in the core. The control rods are made of some highly neutron absorbing material such as boron, cadmium, hafnium, gadolinium or their alloys. For example, boron is used in the form of boron carbide (B_4C) powder stuffed in stainless steel control rods. These move up and down between four fuel assemblies. Sometimes, empty tubes are built-in inside the core and the absorber in the form of a liquid, say gadolinium nitrate, can be pumped in and out. Similarly, in H_2O or D_2O moderated reactors, boric acid (H_3BO_3) is frequently added to the moderator. This is known as chemical shim. Many a time, a small amount of ^{10}B or ^{155}Gd is added to the fuel. A significant feature of these isotopes is that on neutron capture they change into isotopes (^{11}B and ^{156}Gd) which have negligible affinity for absorption of neutrons. Thus, as the fuel depletes, the poison also burns out itself resulting in the compensation to the loss of multiplication factor due to fuel burn up. Such a poison is termed as a **burnable poison**. Use of gadolinium oxide (Gd_2O_3) powder with UO_2 powder is a common example.

In fast reactors boron with increased fraction of ^{10}B is used. In experimental fast reactors periodic removal and installation of the reflector as well as the fuel (or core) have also been used to control the multiplication factor or to shut down the reactor.

Control rods may be subgrouped as shim rods, regulating rods or safety rods, depending upon their function. Shim control rods are used to bring the reactor to the desired power level when the system is started and to make occasional coarse adjustments in the power level. Before the reactor is started, these rods are completely inserted into the reactor core. They are then gradually raised until the reactor attains criticality; they are normally fully out of the core when the reactor is operating at its peak power. **Regulating rods** are used to make fine adjustments in the power level and to compensate for the effects of changes in temperature and pressure, fuel depletion and build up of poisons. **Safety rods** are provided to stop the chain reaction rapidly and are used for an emergency shut down of a reactor in

case of failure of the system. When reactor is critical, these rods are kept in a cocked position outside the core and to shut down the reactor they are moved rapidly into the core. They are distributed uniformly inside the core.

To support fuel elements, for making coolant channels and for various other purposes, it is necessary to use some structural materials inside the core. The desirable properties for such materials are the same as those of the clad and generally one uses the same material for both purposes. For example, stainless steel is used for spacer-grids - the structure which maintains spacing between fuel rods in a fuel assembly. Similarly, for coolant tubes and fuel channels, Zr-2 is used.

Blanket

In fast reactors, which are generally compact, a significant fraction of neutrons in the core leaks out of the system. To reduce this leakage and also to make proper use of the leaking neutrons, core in these reactors is surrounded by a region of fertile material (^{232}Th or ^{238}U). This region is referred to as the blanket. (This also serves as an additional neutron reflector as well as a shield.) Neutrons absorbed in the blanket eventually lead to the production of fissible nuclei, ^{233}U or ^{239}Pu.

Reflector

A region of non-moderating, non-absorbing material put next to the core (or the blanket, if it is present) to inhibit neutron leakage from the core (or blanket) is known as the reflector. In fast reactors high mass number material is used so that the mean energy of neutrons returned back from this region is not much different from that of neutrons entering it. Ni, Cu and Mo are frequently used as reflectors. In thermal reactors any good moderating material can be used as a reflector.

Reactor Vessel

The whole assembly is placed inside a vessel, called the 'pressure vessel'. Usually a few inches thick stainless steel is used to make the reactor vessel.

Shielding

To protect the scientists and other personnel working around the reactor as well as the equipment placed around it from radiations emanating from the reactor core, the reactor vessel is encased inside thick concrete walls. In some cases, alternate layers of heavy and light elements such as concrete and polyethylene or concrete and water are also used. This is referred to as the biological shield or the radiation shield.

The continuous absorption of leaking radiations by the walls of the reactor vessel as well as in the biological shield would result in the generation of heat, causing thermal stresses in them. To reduce this heating effect of nuclear radiation and hence prevent radiation damage of the pressure vessel, a thermal shield, usually made of stainless steel, is placed next to the reflector.

Reactor Building

The entire structure is placed inside a reactor building. It is air tight and is maintained at a pressure slightly lower than the atmospheric pressure so that no air leaks out of the building, except through the ventillation channels. In the event of an accident, the building also helps to contain the radioactive materials and prevents their dispersal into the surroundings.

Coolant Loops, Heat Exchangers and Electric Generators

The heat generated due to fission inside the reactor core is removed by circulating a coolant through it. Usually, the coolant circulates in a closed loop, called the primary (or reactor core) loop. The heated coolant carrying fission heat can become intensely radioactive when it comes out of the core. To prevent this radioactivity from spreading, it becomes necessary to introduce a secondary (or steam generating) loop, which may be closed or open. The primary fluid is made to give up its heat to the secondary fluid, usually water, in a heat exchanger. This results in the production of steam and for this reason, it is also called a steam generator.

In research reactors, secondary loop is usually open so that the heat produced inside the core is just discharged into a river or the sea. In power reactors, the secondary loop provides energy to an electric generator. However, in some reactors high pressure steam is produced in the core itself and it is directly fed to the generator so that no heat exchangers are needed

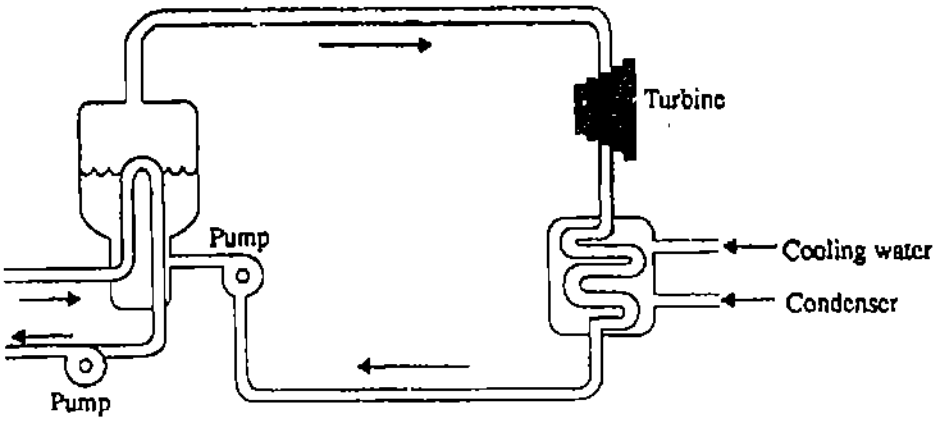


Fig. 14.2

in this case. Such reactors are called boiling water reactors (BWRs). A schematic diagram of a boiling water reactor is shown in Fig. 14.3. On the other hand, in fast reactors cooled by liquid sodium, an intermediate cooling loop (and hence heat exchanger) is also required. This is used to transfer heat from the primary sodium loop to the secondary loop.

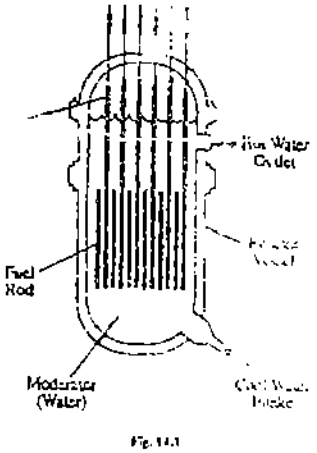


Fig. 14.3

The high temperature and pressure steam from the steam generator (or the core) expends in a turbine coupled to a large electric generator. The low pressure steam leaving the turbine is recondensed into liquid water in a steam condenser. The condensed water is then compressed and pumped back into the steam generator.

The efficiency of nuclear plants (33%) is lower than the efficiency of fossil plants (40%). This is because the temperature of the nuclear fuel (and hence steam) is kept lower than that of the fossil fuel to avoid melt down. However, it is hoped that soon it will also be raised to about 40%.

14.3.3 Nuclear Energy: The Indian Saga

Pt. Jawaharlal Nehru dreamt of a technologically advanced post-independent India and set about laying its foundation immediately after taking over as Prime Minister. He was convinced that energy needs in the envisaged scenario would ever increase. On the question of energy options, he recognised inevitable reliance on nuclear 'apple' because of obvious limitations of hydel, coal and oil based sources of electricity generation. In this endeavour,

Fig. 14.4

he relied on the genius of Dr. Homi Bhabha, who responded to the challanges of nurturing nuclear programme of a 'bleeding' nation. The Atomic Energy Act of 1948 envisaged the development, control and use of atomic energy for peaceful and beneficial purposes, namely the generation of electricity and the development of nuclear applications in research, agriculture, industry, medicine and other areas. To achieve this objective, he initiated efforts to build up a versatile infrastructure of research facilities, trained scientific and technical manpower, raw material processing centres and capability to manufacture nuclear and electronic equipment to support the atomic energy programme and make India self-reliant. Today various nuclear installations are spread all over the country (Fig. 14.4).

The Indian nuclear programme is three tiered

- Construction of natural uranium-fuelled pressurised heavy water power reactors (PHWRs) producing plutonium as a bye- product (1st phase).
- Development of plutonium fuelled fast breeder power reactors producing plutonium and uranium-233 from thorium (2nd phase).
- Design and construction of reactors based on the thorium cycle producing more plutonium than uranium (3rd phase).

Though our policy has been to build a series of nuclear power stations burning natural uranium in the first phase of the programme, lack of infrastructure and scarcity of skilled manpower forced our scientists to resort to a modest change. India purchased two light water reactors of 200 MW each from USA in 1964 as a turn key project. These were installed at Tarapur (Maharastra) and burnt enriched uranium. On the basis of operational experience gained from these reactors, India embarked on the construction and installation of two 220 MW PHWRs based on Canadian design (proposed in the first stage) at Kota (Rajasthan). The bulk of equipment for the first reactor came from Canada. However, a major step towards indigenisation was taken in the construction of the second unit of the Rajasthan Station.

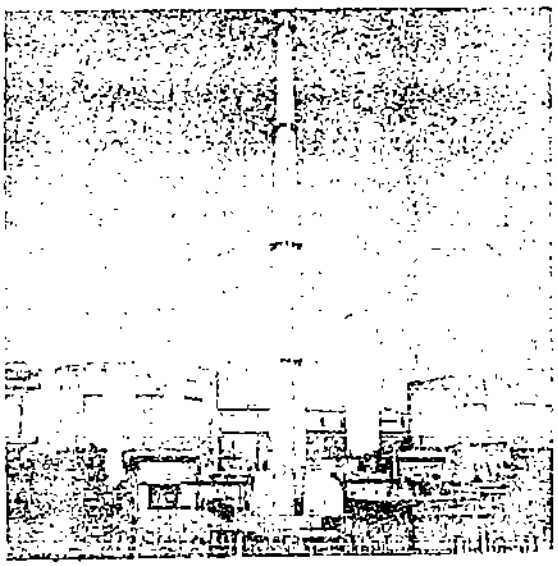


Fig. 14.5 : Kakrapara Atomic Power Station

The third atomic power station at Kalpakkam, 100 km from Madras, consisting of two units of 235 MW each, marked the coming of age of the Indian nuclear energy programme. Complete responsibility for the execution of the project, including design, construction, commissioning and operation rested with the Department of Atomic Energy. While the nuclear design of Rajasthan Reactors was retained, significant changes were made in nuclear containment. Major equipment and components were indigenous and the extent of foreign exchange spent was only of the order of 20%.

This success enthused Indian scientists. They started to build and put into operation nuclear power stations at Narora in U.P. This site lies in the Seismic Zone and the design of the station posed real challenge. Indian scientists and engineers rose to the occasion and the opportunity was used to redesign the reactor completely. Two units are in operation now. To coordinate the activities of a fast expanding nuclear energy programme, Nuclear Power Corporation was established in 1991. It started construction work at Kaiga (Karnataka) and

Kakrapara (Gujarat). Each of these projects consists of two units of 235 MW. Of these, both units at Kakrapara have attained criticality and are in commercial operation today (Fig. 14.5).

The recently formed Nuclear Energy Corporation proposes to build two units of 440 MW and two units of 220 MW. This may be followed by additional reactors depending on the availability of resources. In 70's DAE had projected a capacity of 10,000 MW by the year 2000 A.D. Now this seems to be too optimistic. As of the latest available information, India may generate only 3000- 4000 MW of electricity from its nuclear installations by the end of this century.

Nuclear Power Generation in India

Presently in commercial operation (2170 MWe):

Tarapur Atomic Power Station	–	2 × 160 MWe (TAPS)
Rajasthan Atomic Power Station	–	2 × 220 MWe(RAPS)
Kalpakkam Atomic Power Station	–	2 × 235 MWe
Narora Atomic Power Station	–	2 × 235 MWe (NAPS)
Kakrapara Atomic Power Station	–	2 × 235 MWe

Under Construction (470 MWe):

Kaiga Atomic Power Project	–	2 × 235 MWe

Proposed (1350 MWe):

Tarapur Atomic Power Project	–	2 × 440 MWe
Rajasthan Atomic Power Project	–	2 × 235 MWe

Second phase

The commissioning of the Fast Breeder Test Reactor (FBTR) in 1985 with a design capacity of 40 MW thermal and 13 MW electrical power at the Indira Gandhi Centre for Atomic Research (IGCAR), Kalpakkam (Tamil Nadu) marked the commencement of the second phase of India's nuclear programme.The next step after FBTR is to design and construct a Prototype Fast Breeder Reactor (PFBR) of 500 MW capacity.

Third phase

The third phase of the nuclear power programme envisages use of $^{233}U - ^{228}Th$ cycle to utilize the country's abundant reserves of thorium. India has the largest deposits of thorium in the world. It is proposed to use ^{239}U, derived from the first two stages to breed ^{233}U from thorium.

14.3.4 March towards Self-reliance in Nuclear Power Programme

Nuclear industry is very complex and to sustain it we need diverse paraphernalia: fuel, coolant, moderator etc., apart from heavy equipment for production of electricity. As you may have realised, a lot of radioactive waste is also generated. This necessitates additional infrastructure. Particular mention may be made of fuel processing/reprocessing complex, heavy water generation plants, radioactive waste disposal sites. We will now discuss these briefly within the framework of our nuclear energy programme.

a. Heavy Water Production

Heavy water is used as moderator as well as coolant in most of the Indian nuclear power reactors. Heavy water option was made primarily because a reactor can be made critical even with natural uranium. At present, several such plants are in operation: one each in Gujarat (Baroda), Punjab (Nangal), Rajasthan (Kota), Tamil Nadu (Tuticorin), Orissa (Talcher) and Maharashtra (Thal).

Reprocessing of Fuel

Based on the experience gained at Trombay, a 100 tonnes per year power reactor fuel reprocessing Plant has been set up at Tarapur for reprocessing zircalloy - clad uranium oxide fuel elements from the Tarapur Atomic Power Station. This plant will also have additional facility for reprocessing fuel from the Fast Breeder Reactor, India is the fifty country in the world to start reprocessing.

Chronology of the Indian Nuclear Programme		
Aug. 10, 1948	:	Setting up of the Atomic Energy Commission (AEC).
Aug. 3, 1954	:	Establishment of the Department of Atomic Energy (DAE).
Aug. 4, 1956	:	The first research reactor (in Asia) APSARA attains criticality at Trombay, Bombay.
Jan. 20, 1957	:	Pt. Jawahar Lal Nehru inaugurates Atomic Energy Establishment, Trombay.
Jan. 30, 1959	:	Nuclear grade uranium metal produced at AEE, Trombay.
July 10, 1960	:	40 MWt Research reactor CIRUS attains criticality.
Jan. 14, 1961	:	Research reactor ZERLINA becomes critical.
Jan. 22, 1965	:	Plutonium Plant at BARC commissioned and made operational.
Jan. 22, 1967	:	Atomic Energy Establishment, Trombay renamed as Bhabha Atomic Research Centre (BARC).
Dec. 31, 1968	:	Nuclear fuel complex set up at Hyderabad.
March 12, 1969	:	Setting up of Reactor Research Centre at Kalpakkam.
Oct. 2, 1969	:	Tarapur atomic power station starts commercial operation.
May 18, 1972	:	A plutonium fuelled fast reactor PURNIMA-I attains criticality.
Nov. 30, 1972	:	The first reactor at Kota, Rajasthan goes critical.
Nov. 1, 1980	:	Unit II at Kota goes commercial.
Nov. 19, 1982	:	Power Reactor Fuel Reprocessing plant at Tarapur is commissioned.
Nov. 15, 1983	:	Atomic Energy Regulatory Board is constituted.
Jan. 27, 1984	:	Madras atomic power station– Unit I at Kalpakkam starts commercial operation.
May 10, 1984	:	PURNIMA-I modified as PURNIMA-II to use ^{233}U as fuel.
Aug. 8, 1985	:	Research reactor DHRUVA (100 MWt) attains criticality.
Oct. 18, 1985	:	Fast breeder test reactor at Kalpakkam attains criticality.
March 21, 1986	:	Unit-II starts commercial operation at Kalpakkam.
March 12, 1989	:	Unit I of the Narora power project attains criticality.
Oct. 24, 1991	:	Unit II of the Narora power project attains criticality.
Sept. 3, 1992	:	Unit 1 of Kakrapar atomic power station attains criticality.
Jan. 8, 1995	:	Unit II of the Kakrapar atomic power station attains critically.

Radioactive Wastes Disposal

More than 99% of the total radioactivity in the entire nuclear fuel cycle is generated from the fuel reprocessing plant. In India a three-stage approach has been adopted to handle nuclear waste.

- waste will be incorporated in stable and inert solid matrices.
- conditioned waste to be placed subsequently in canisters and kept in a retrievable store under cooling and constant surveillance.
- store the canisters in suitable geological media.

A Waste Immobilisation Plant for incorporating high level radioactive wastes generated from the fuel reprocessing plant is set up alongwith the solid storage surveilance facility at Tarapur.

14.4 NUCLEAR RADIATIONS IN OUR ENVIRONMENT

Man has always existed in close contact with nature. Pre- historic man lived in caves and depended for all his needs on nature and its creations - flora, fauna and aquatic life. Passing through the stone, bronze and iron ages, man entered the industrial age. In the course of this progress, he succeeded in controlling several fatal diseases, which resulted in a steep fall in the mortality rate and exponential growth in world's population. In an effort to cater to the needs of vast population, man began misusing resources with adverse effects on our environment. Besides this, due to increasing carelessness in the management of industries and lack of adequate technical security, there have been so many fatal incidents. In the process, we have injected several chemical and nuclear pollutants in the environment. Bhopal gas tragedy, Chernobyl and Three Mile Island accidents, fires in Kuwait's oil wells are some of the familiar examples. Depleting forests, diminishing ozone layer, ever threatening greenhouse effect and global warming are some of the adverse effects of uncontrolled industrialisation with little concern for nature and its creations. We have to check release of pollutants. Otherwise, consequences would be fatal for all biological systems on this beautiful planet.

We all know that air pollution, particularly in cosmopolitan cities like Delhi, Calcutta, Bombay, Madras and Bangalore has assumed alarming proportions. Thermal power plants, industrial units, burning of fossil fuels and vehicular traffic release chemical pollutants, like carbon-monoxide, sulphur dioxide, hydrocarbons and oxides of nitrogen in our atmosphere. We are familiar with these pollutants to some extent. But since 1940 nuclear radiation has emerged as a more dangerous pollutant. It can neither be seen nor felt or smelt. Unfortunately, in our country many people know very little about it. Since there is a gradual shift from fossil fuels to other sources of energy and nuclear energy generation is high on our agenda, the level of nuclear radiations is likely to go higher in our environment.

What is Nuclear Radiation?

The term radiation is very broad and includes such things as light and radio-waves. But only those radiations which originate in the nucleus of atoms are termed as nuclear radiations. These include the radioactive emanations (alpha, beta and gamma rays), cosmic rays, neutrons and nuclei of light elements.

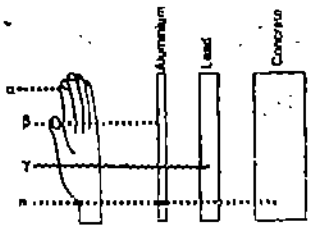


Fig. 14.6 : Type of Radiation and their penetrating powers

You all know that alpha particles are emitted by atoms of elements such as uranium and radium. The penetration power of alpha radiation is small and can be stopped completely by a sheet of paper; even thin surface layer of our skin. But if alpha emitting nuclei are taken inside the body (by breathing, eating or drinking), they can expose internal tissues directly and cause biological damage. Beta radiation can pass through 1 to 2 cm thick layer of water. But a few millimeter thick sheet of aluminium is capable of stopping these radiations. Gamma rays can pass through the human body but can be stopped by thick sheet of concrete or lead. These aspects are illustrated in Fig. 14.6

Cosmic rays are highly energetic particles which reach earth's surface from the outer space; the number of cosmic particles hitting us per minute is nearly 7000. Some of these particles carry energy as high as 10^9MeV.

Neutrons are uncharged particles and can penetrate large distances without any hindrance. They do not interact appreciably with electrons. As a result, they do not cause appreciable ionisation. Their detection is therefore difficult. They can be stopped only by thick walls of concrete.

Radioactivity is not new to man; it is a part of our earth — it has existed as long. In fact, the entire biological evolution on our planet has taken place with it. The natural sources of nuclear radiations are

- radio-active minerals present in earth's crust, the floors and walls of our houses, schools and offices and in the food we eat.
- radioactive air-borne nuclides which we breathe. Our own body — muscles, bones and tissues — contain naturally occurring radioactive elements.
- cosmic rays from the outer space.

Radioactive minerals like ^{40}K, ^{238}U and ^{232}Th are widely distributed in earth's crust and give rise to terresterial radioactivity. The concentration of these minerals in the soil determines

the intensity of nuclear radiations at a particular place. In India these minerals are abundantly available in Kerala and Bihar. Like terrestrial radioactivity, the strength of cosmic radiations reaching the earth depends on latitude, altitude, explosions, solar flares, etc. Their number is maximum at the poles and minimum at the equator. At the sea level, they are mostly μ-mesons and electron-positron pairs. At a height of about 15 km from the sea level, cosmic radiation consists mostly of protons and alpha particles. But for earth's magnetic field and the thick envelope of our atmosphere, the number of cosmic particles reaching us would have been much larger. People travelling by jet liners get an extra exposure to cosmic radiations and are that much more vulnerable to its ill- effects. Air-borne radioactivity is primarily due to ^{222}Rn and ^{14}C. The latter is continuously produced in the atmosphere by the action of cosmic rays on nitrogen. These nuclides affect us from within for their life time once they are inhaled.

It is true that the biological systems have evolved with natural radiation without much ill-effects but now man made (artificial) sources have begun to add large doses of nuclear radiation to the existing natural radioactivity. The radiation exposure of an average person in India and United Kingdom are given in Table 14.2. Major present day artificial sources are:

Table 14.2: Total population radiation exposure (%)

Source	Country	
	India	U.K.
Natural	79.7	87
Medical	16.2	11
Artificial	4.1	1.5

The world has seen more than 1000 nuclear detonations and two of them killed 160,000 people. Recent nuclear tests conducted by France in the South Pacific, in spite of the genuine criticism the world over, are a matter of great concern particularly for the people of the region.

radioisotopes, nuclear installations, radioactive waste products and fall-out from weapons tests. In a nuclear explosion or a reactor accident, a large quantity of long-lived radio-nuclides are released to atmosphere, which get distributed via winds all over the globe. The Chernobyl reactor accident is fresh in our memory. The radio-nuclides usually settle down with rain and mix with soil, water and vegetation. Once deposited in human body through food, they are sure to radiate internally for almost the whole life-span of an individual. The radiations which arise as a result of a nuclear explosion are generally insignificant compared to the destructive force of the blast and heat, but in the long run their effects are quite devastating.

You must have read about the use of an atomic device during the second world war. On August 6, 1945, the US dropped an atom bomb on Hiroshima, Japan. The heat released in the explosion devastated 10 sq. km of the city of 3,43,000 inhabitants. Nearly 66,000 people were killed instantly due to heat and 'nuclear burns' and another 69,000 were injured. The deadly consequences of nuclear radiations were again experienced in less than a week (on August 9, 1945) when the second atom bomb was dropped on Nagasaki, Japan.

Radioisotopes administered to patients and used in radiation therapy and scientific research are now proving to be an important source of nuclear exposure. The ability of radiations to kill diseased cells has made them an indispensable tool in the cure of some lethal diseases. However, their indiscriminate use, over-doses to patients and improper handling can be a cause of great concern.

Leakage of nuclear radiations from nuclear reactors and nuclear research laboratories may also increase with the increase in number of such facilities. Even with proper handling of radioactive materials and the use of shielding techniques, some radiations, such as radioactive emanations and neutrons, do leak out of research laboratories and reactor cores. The level of radiation from such sources is bound to increase in future, if more nuclear reactors are installed to meet increasing energy demands. What can happen in a nuclear accident like reactor core melt-down is no longer anybody's guess — the consequences of Chernobyl accident in the former USSR have been really frightening for all forms of life even beyond Europe and Asia.

Burnt-up fuels of nuclear reactors, also called radioactive wastes, contain vast quantities of long-lived radionuclides. Transportation and proper disposal or storage of these wastes is a major problem today. If not properly disposed of, these wastes can be an ever-rising source of nuclear radiations.

Because of the shortage of conventional building materials and the necessity to provide cheaper houses, industrial waste - products are often used in buildings. The flyash from power stations and slag from steel plants are the common examples. This increases our exposure to nuclear radiations. In a scientific study carried out in Norway it has been found that people living in concrete and brick houses receive about 30% more radiation than those living in wooden houses.

Biological Effects

Nuclear radiations produce no sensation unless absorbed in very large quantities. In many cases they have delayed effects, as late as 20 to 30 years. There are two types of biological effects associated with nuclear radiations; somatic and genetic. The somatic effects are limited only to the persons exposed, while the genetic ones may affect later generations too.

Nuclear radiations dissociate the complex molecules of living tissues through ionisation and kill the cells. They can induce malignant growth, cause sterility, severe skin burns and lower the body resistance against diseases. They disrupt the genetic process, maim the unborn child and show their effects even upto five generations.

Nuclear radiations can affect us even indirectly through the flora, fauna and the aquatic life around us. They kill and maim vegetation, fish and animals. As such, we have no control on natural radiation. But efforts are to be made at individual, collective, community, national and international fronts to keep the level of radiation from artificial sources at the minimum. Otherwise, cumulative, long term and indirect effects of nuclear radiations would outweight the benefits accruing from them; they can even prove to be a scourge for mankind.

The damage caused by nuclear radiation depends on the exposed part of body as well as on the energy, intensity and the nature of the radiation. Different parts of human body show different sensitivities to radiation. In general, hands and feet can receive much larger radiation without being affected than the other parts. Radiations are usually more harmful if they are more energetic and more intense. The alpha particles are, as a rule, quite harmful because of their high ionising power. The damaging effects of different radiations are generally compared in terms of their relative biological effectiveness, called the RBE factors. These factors for different particles/rays are given in Table 14.3.

Table 14.3: RBE factors of different radiations

Particles/rays	RBE factors
γ-rays, β-particles	1
Thermal neutrons	2 to 5
Fast neutrons	10
α-particles, high energy ions of O, N, etc	10 to 20 .

The amount of radiation which a human body absorbs is referred to as a dose. The unit used to measure this dose is the *rem*, which is the abbreviation of **Roentgen Equivalent Man**. 1 rem = RBE factor × rad, where rad is a unit of absorbed energy equivalent to 100 ergs of absorbed energy per g of absorbing material. For any individual, the safety limit of radiation exposure is usually put at 500 milirem per year. This may be compared to a dose of 130 millirem per year from natural background (Table 14.3).

It has been observed that higher forms of life are generally more prone to harmful effects from nuclear radiations than certain bacteria, which seem to have a safe dose limit of even upto 5×10^5 rem per year.

Cities located at sea level near the equator receive a cosmic ray dose of 35 millirem per year, while those located at latitude 50 north receive on an average, 50 millirem per year. At an altitude of 1850 meters above sea level, the cosmic ray dose is 90 millirem per year which increases to about 300 millirem per year at a height of 4600 meters. This suggests that people living in hilly regions are more susceptible to its ill effects.

Of the radioactive isotopes which enter our body through food, the one which releases highest radiations is 40k. It releases approximately 20 millirem per year.

14.5 RADIOISOTOPES IN EVERYDAY LIFE

Apart from generation of electricity, other vital contribution of a peaceful nuclear power programme is in the production of radioisotopes. As the name suggests, isotopes which are radioactive in nature are called radioisotopes. Their uses in everyday life are too many and too varied. These are based on the penetrating radiations they emit. You must be eager to know them. But before discussing them it is worthwhile to know how are radioisotopes produced.

14.5.1 Production of Radioisotopes

Naturally occuring radioisotopes have been available ever since the phenomenon of radioactivity was discovered in 1896 (Unit 12). However, they were available in minute quantities and were restricted to a few elements (Ra, Po, Th, U etc). However, with the discovery of artificial radioactivity (using charged particle nuclear reactions) in 1934, it became possible to produce radioisotopes of many elements. Even then, only limited quantities became available and that too at considerable cost. A wide range of abundant and relatively inexpensive supply of radioisotopes became possible with the commissioning of nuclear reactors in the fifties. Today, about 170 radioisotopes are in use for various purposes. Of these, about 120 are produced in research reactors while others are produced in accelerators (Unit 15).

In India, we now have an exclusive infrastructive for production of radioisotopes though a modest beginning was made towards the end of 1957 soon after the research reactor APSARA became critical. The experience gained with Apsara widened the scope of our programme on the commissioning of CIRUS in 1960. Comprehensive facilities for large scale processing of radioisotopes were installed at Trombay in the seventies. The large quantities of ^{60}Co began to be produced in the RAPS. During the eighties, Dhruva reactor was commissioned at Trombay for production of isotopes.

The applications of radioistopes are based on

- tracer studies
- effect of materials on emitted radiation
- energy of emitted radiations

Of these, the radioisotope tracers have the widest spectrum—ranging from studies with single living cells to movement of sediment on sea bed. Important industrial applications of radiotracers include leak detection in buried pipelines, seepage in dams and canals, measurement of flows etc. Applications of radioisotopes based on the effect of materials on radiation are very common in industry. These include non-destructive testing (using gamma radiograph), measurement of thickness of metal sheets, foils etc. The main applications of radiation energy include sterilisation of medical products, radiation therapy and preservation of food products. We will now discuss these briefly.

14.5.2 Isotopes as tracers

Before we discuss the applications of radioactive tracers, it is important to know: What is a tracer? A substance which enables us to locate something of interest deep inside a body is known as tracer. This is very similar to the tagging of birds while studying their migrating habits. A tracer must be identifiable and not present in the system being studied. Moreover, it should not interfere with the normal working of the system. Here we shall briefly discuss how radioisotopes are used as tracers to follow the flow of blood through the body (of animals or human beings) (physiology), photosynthesis (plant physiology), oil flow (industry), self-diffusion in metals (metallurgy), wear studies, leak detection etc. It is essentially because a radioisotope does not lose its identity during physical and chemical reactions and its path can be followed at all times using sensitive detectors.

(a) Study of Blood Flow

Suppose that we wish to follow the path of iron in an animal's body once it enters as a constituent of food. You can do this by labelling the iron contained in the food by the radioisotope ^{59}Fe. Most of the iron in the body resides in hemoglobin (which causes the colour of red blood corpuscles). You would expect the ^{59}Fe to enter the hemoglobin and show up through its β^- and γ activity. But this does not happen. The iron is found to go in

store in the body in the form of an iron-protein combination (called ferritin). Only when there is loss of blood, rapid growth or pregnancy, this stored iron is utilised, the amount of ferritin decreases and the body draws iron from food to replenish its iron stock.

(b) Photosynthesis

Green plants are known to manufacture carbohydrates from carbon-di-oxide and water in the presence of sunlight and evolve oxygen in the process. One reaction can be summarized as:

$$6\,CO_2 + 12\,H_2O + \text{solar energy} \xrightarrow{\text{chlorophyll}} C_6H_{12}O_6 + 6H_2O + 6\,O_2$$

You may now ask: What is the source of the evolved oxygen (CO_2 or water) or the carbon in the carbohydrates (CO_2 or plant's own body)? To find the answer, we have to label the oxygen by the stable isotope ^{18}O and the carbon in CO_2 by radiocarbon ^{14}C, as shown in Fig. 14.7.

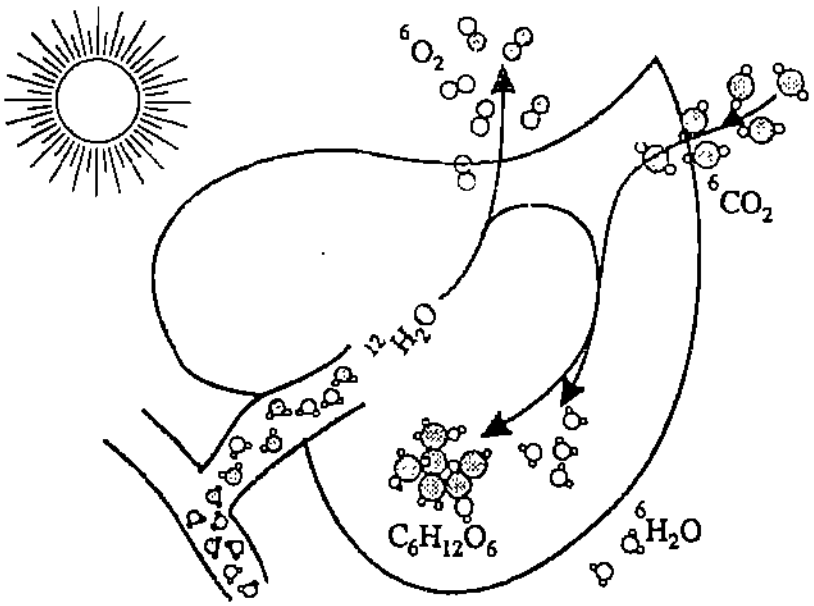


Fig. 14.7: Use of tracer in photosynthesis

Ordinary oxygen contains only 0.2% of ^{18}O. When ^{18}O enriched oxygen in water is absorbed by the plant, the evolved oxygen is also found to be correspondingly enriched in ^{18}O. It means that the plant produces the evolved oxygen by breaking the water molecules.

To study the path of carbon, carbon-dioxide was bubbled in one experiment through a suspension of algae (single-celled green plants) in a nutrient liquid medium. While photosynthesis was in progress, some bicarbonate solution containing ^{14}C as tracer was added to the medium. Within 30s of adding the radiocarbon, some 20-30 compounds were found to have been formed that contained the isotope suggesting that the carbon in the CO_2 is used by the plant in the synthesis and that the process is almost instantaneous.

(c) Oil Flow

When oils from different stocks are transported through a single long pipe-line one after another, we may like to keep track of the surface of separation, besides the degree of intermixing. The problem is solved with ease by injecting a γ- emitting tracer in between two oil stocks. A tracer commonly used is Antimony-124 (^{124}Sb) which decays with a half-life of 60 days into stable ^{124}Te, emitting β and γ rays of energy upto 1.69MeV. As oil flows, the injected tracer moves along, making the oil interface. The tracer spreads somewhat depending on the intermixing of the oils. The γ-rays easily penetrate the wall of the pipe-line and can be detected by a counter placed outside the pipeline at the receiving end. When the counter starts ticking, arrival of the new stock of oil is announced (Fig. 14.8).

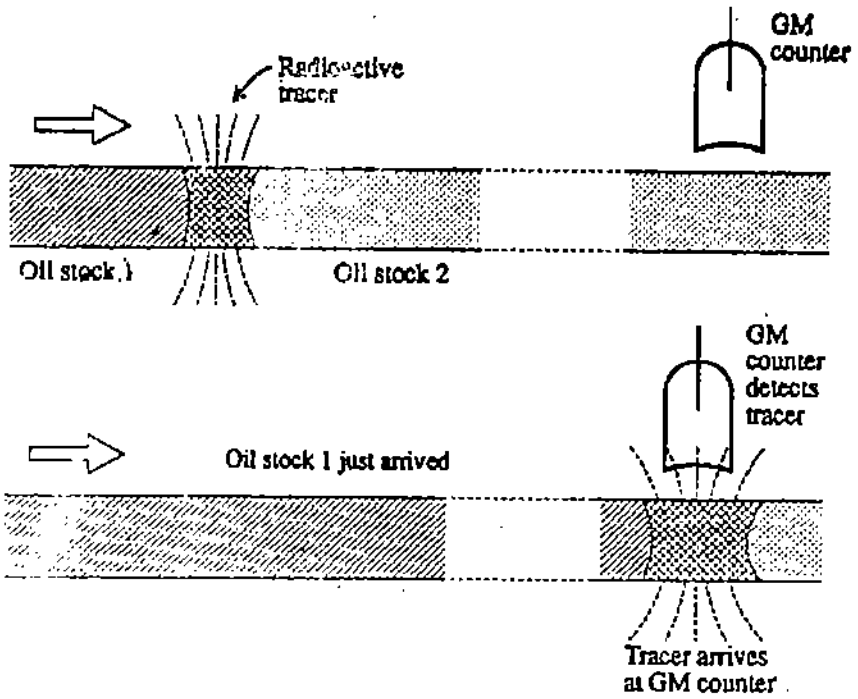


Fig. 14.8: Oil flow in pipeline

(d) Self-diffusion in Metals

With the tracer technique, we can study how atoms in a metal move about within a crystal lattice. A block of copper, for example, consists normally of ^{63}Cu and ^{65}Cu atoms. To investigate the self-diffusion of copper atoms, we deposit a layer of copper containing ^{67}Cu (half-life = 59 hours) on the block. The copper block may now be subjected to any treatment, the effect of which we want to study. We then scrape off layers of copper from the block and the activity of ^{67}Cu in these layers readily indicates the amount of self-diffusion that has occured.

(e) Wear Studies

The conventional methods for measurement of wear rate of an engine, machine parts as cutting tools under different operating conditions are very laborious and time consuming. The use of radiotracers not only makes it possible to measure the wear rate precisely and continuously but also rapidly and at low cost. The component, whose wear is to be measured, is made radioactive by irradiating it with neutrons in a nuclear reactor or by using charged particles such as protons and deuterons in a cyclotron. The irradiated component is placed in the test rig and radioactivity in the debris present in the lubricating oil is measured by a scintillation detector. The studies carried out at Pune, based on this technique, have enabled rapid evaluation of wear of different makes of piston rings and its correlation with the composition of the piston ring.

(f) Hydrology

Radioisotopes have proved to be indispensable tools for investigations in hydrology and water management. They have been extremely useful in the measurement of recharge to groundwater (by tagging moisture layer with tritium tracer), detection of seepage in canals and dams and for measurement of flows in rivers and canals. In all these applications, other techniques are either not available or do not provide data with required precesion or ease. In India, radioisotope tracers have been used for measurment of recharge to groundwaters in arid zones, detection of seepages in dams and reservoirs (Shrisailam dam, Aliyar dam) and for measurement of high discharges in turbulent rivers like Tapi and Beas. Environmental isotope techniques have also provided valuable information to water management programmes and they have been used to study the efficacy of percolation tanks in Maharashtra, intrusion of saline waters in coastal acquifers in Tamil Nadu etc.

(g) Leak detection

The detection of leaks in buried pipelines and in equipment of chemical process plants is one of the very common uses of radioisotopes as tracers. The detection of leaks in buried

pipelines normally involves removing the soil cover over the entire length of the pipeline (which is very time consuming and expensive) for visual inspection of the suspected leaky section. By using tracer technique, leaks can be detected without removing the soil covering the pipelines. In one technique used for this purpose, the pipeline is filled with the radiotracer solution which is pressurised so that a small quantity of radiotracer leaks into the soil at the point of leak. Subsequently, the pieline is flushed with water and the point of leak is identified by portable radiaiton detector which is moved along the length of the pipeline. This technique has been used for detection of leaks in the 140 km long Viramgam Koyali pipeline. Five minute leaks were determined in a period of 6 weeks and this enabled commissioning of the plepline on schedule. Conventional methods are estimated to have taken one year for completing the work and at almost ten times the cost incurred in carrying out tracer studies. Similarly study carried out on the pipeline between an oil refinery in Bombay and the petroleum storage and distribution centre near Pune enabled the detection of minute leaks in the pipeline prior to its commissioning.

The above examples have been chosen from various fields of study to show the usefulness and versatility of the tracer technique. Numerous other applications of the tracer technique have been made and it is not difficult to devise a new one whenever identification of atoms is the key requirement. In case you are interested in details, see Nuclear India, publised by BARC, Bombay.

14.5.3 Nuclear Radiations from Isotopes

You now know that over exposure to nuclear radiations can be fatal but we donot have to be afraid of nuclear radiations per se. Today they have been put to several uses in various facets of human activity. You are already familiar with nuclear fission induced by neutrons for power generation. Other important practical applications in everyday life are in the fields of medicine, agriculture, industry and research. We shall now discuss some of these.

(a) Radiography

You must be familiar with the use of X-rays for photographing bones and internal organs for diagnostic purposes. Even for non-destructive examination of welds and castings, use of X-rays in known for many years. But X-ray units cannot be conveniently carried from one place to another. Moreover, not every organ/location is accessible to such units. Radioactive γ-sources are more compact and are found to be more advantageous.

Suppose we want to test a metal casting for flaws like cracks and bubbles. We expose the sample to a γ-source, backed by a photographic plate. Any irregulatkes in the casting would show up in the pin-hole picture on the photographic plate. Radiography therefore plays an important role in industry quality control of castings and inspection of welded pipes and fabricated machine parts.

The choice of the γ-source depends on the penetrating power (or energy) of the γ-rays and their half-life. Some commonly used sources are ^{60}Co (1.33,1.17 MeV), ^{137}Cs (0.662 MeV), ^{192}Ir- (0:47 MeV), ^{143}Sm-(0.061 MeV) and ^{125}I-(0.035 MeV). The quantities in brackets signify the energy of γ-rays. In Indian industries, over 600 γ-radiography cameras using upto 100Ci of ^{192}Ir and 20Ci of ^{60}Co are is use to test high pressure vessels, ships, aircraft, nuclear and thermal power stations, fertilisers and petrochemicals. The radition sources, γ-radiograph equipment is provided by BARC.

(b) Electric Power Generation

Radio-isotopes can be used as small and self-contained power sources. When their radiations are absorbed in matter, their energy is ultimately converted into heat. We can use a thermocouple or a thermionic emission device to utilise this heat to generate electricity. and β-emitters are preferred since these particles are absorbed in relatively small thicknesses of matter and a large temperature rise is easier to achieve.

The sources commonly used ^{90}Sr, ^{137}Cs, ^{210}Po, ^{238}Pu etc. A typical power source, SNAP - 3 (SNAP = Systems for Nuclear Auxiliary Power) uses a ^{210}Po source surrounded by hot junctions of 54 thermocouples in series. With hot and cold junctions at 550C and 150C, it can generate 5 W of electric power.

Isotopic power sources find use in satellites, unattended weather stations and light-houses. Their working life depends on the half-lives of the isotopes used. In 1969, the first weather satellite fuelled with PuO_2 was launched.

The Apollo missions have included spectacular applications of radioisotope generators. For the first time, ^{238}Pu fuelled heater in the seismometer of Apollo-11 was used to keep the equipment warm. The Apollo-12 mission left a SNAP-27 generator on the moon. It produced 63 W of electricity. It was also fuelled by ^{238}Pu.

(c) Radiation Chemistry

We can use nuclear radiation as a catalyst in producing many chemical reactions, with the advantage that it leaves no by- products. A common example is the use of β and γ radiation in polymerization, i.e. chemical formation of long chain molecules called polymers by joining smaller molecules of monomers.

Many chemical compounds have been formed by exposing the constituents to fission fragments in a nuclear reactor. The process has actually been utilised in the industrial production of hydrogen peroxide from water, ozone from oxygen, hydrazine from ammonia and so on.

(d) Sterilization and Food Preservation

Radiation is capable of destroying bacteria and other micro- organisms. It is possible to kill all bacteria with 2-5 milli rad of radiation, while 10% of this dose will produce pasteurization i.e. destruction of enough bacteria so as to increase the shelf-life of food. Thus a convenient method of preservinig fish, meat, fruit or vegetables is irradiation at low temperature. Sterilization of surgical dressings, hypodermic syringes, raw wool etc. can be done by irradiating them to γ-rays from ^{60}Co.

(e) Industry

Radiaisotopes find varied and ingenious applications in industry. In plants manufacturing paper, plasting films, metal or rubber sheets etc., radiation can be used as a thickness gauge. You only have to place a source of radiations for metal sheets and β radiations for less absorbing materials, against a radiation detector with the sheet passing between the two. The fluctuations in the thickness of the sheet are detected from variations in the count rate due to attenuation of radiation. Radioisotope gauges are also used for checking the filling of LPG cylinders. More than 1500 nucleonic gauges are in use in Indian industry today.

You may have seen watch-dials with figures written in self- luminous paints. The paints are made by mixing some β-emitting isotopes with phosphorescent materials like zinc sulfide. The β-radiation excites the phosphorescence and the paint glows in darkness.

(f) Activation analysis:

It is often very difficult by chemical means to detect and estimate very small amounts of impurities of one substance in another. If you bombard the impure substance with a suitable radiation - neutrons or some charged particles, it converts the impurity nuclei into radioisotopes. The radiations emitted by these isotopes and their decay half-lives give away their identity and we can infer the impurity of parent isotope.

The activation method is capable of detecting and estimating 10^{-6} to 10^{-12}g of a substance. It can even detect variation in isotopic composition of an element in two samples, a job which chemical analysis just cannot do.

14.5.4 Radioisotopes In Agriculture

In recent years, radioisotopes have found vast applications in agricultural research. These investigations have led to the development of improved methods of agriculture and soil management. Radioisotope labelled fertilizers have been used to study the role of essential nutrients such as phosphorus, calcium, sulpher etc.

Phosphorus is applied to soil to improve plant growth. Various types of phosphate fertilizers are available and it is helpful to know which variety is utilized most effectively by the particular soil. In fact, we can optimise the use of fertilisers on different soils under different climatic conditions. By using a fertilizer in which radiophosphorus is employed it is possible to know to what extent the plant absorbs phosphorus from that particular fertilizer. This facilitates the selection of the fertilizer most suitable to the soil. Without the use of radio phosphorus it is not possible to distinguish between the phosphorus from the soil and that from the fertilizer.

are available and it is helpful to know which variety is utilized most effectively by the particular soil. In fact, we can optimise the use of fertilisers on different soils under different climatic conditions. By using a fertilizer in which radiophosphorus is employed it is possible to know to what extent the plant absorbs phosphorus from that particular fertilizer. This facilitates the selection of the fertilizer most suitable to the soil. Without the use of radio phosphorus it is not possible to distinguish between the phosphorus from the soil and that from the fertilizer.

(a) Crop Mutations:

Plant genetic studies have helped in obtaining crops and vegetables with high yield, develop resistance to disease, easier to grow and adaptability to new environment. Normally one has to wait for natural changes in the plants or crops that are passed on to next generations. Natural mutations and particularly desired ones are, however, rare. A more convenient alternative is to expose seeds or whole plants to radiation of varied types and intensities. Desirable mutations are induced by irradiation of seeds by γ rays or neutrons. High quality wheat, barely and rice mutants resistant to certain diseases and several jute mutants are some of the many successes of this method. In India, several improved qualities of seeds, spices and vegetables have been successfully developed.

Fig. 14.9: Improved quality of groundnuts developed at BARC

(b) Pest Control

Certain insect pests can be controlled by adopting the so-called sterile male technique. The procedure consists of laboratory breeding of large numbers of male insects, sterilizing them with a radiation and releasing them for mating in the infested area. They compete with, in fact out number, normal males and mate with female insects. As a result, most of females lay unfertilized eggs. This rapidly depletes the population of the insect. In parts of the United States, the screw worm flies, a pest which caused great damage to livestock, was eliminated by adopting this procedure.

Through their versatility, the radioisotopes have immensely benefitted mankind. This however in just a beginning; we are far from exhausting their full potential.

Let us now summarise what you have learnt in this unit.

14.6 SUMMARY

- A fission chain reaction is said to be self-sustained when the rate of production of neutrons is equal to the rate of loss of neutrons. The first self-sustained chain reaction was achieved by E. Fermi in 1942. He used graphite to slow down neutrons before they caused fission in uranium rods lumped uniformly.
- The time behaviour of neutrons in a chain reaction can be apaproximately know through the relation

$$N(t) = N(0) \exp [(k\text{-}1) t/l]$$

- Most of the nuclear energy generated so far has come from thermal reactors where fission is caused by 0.0253eV neutrons.
- The Indian nuclear power programme is three tiered. In the first phase, uranium fuelled heavy water moderated nuclear reactors of Canadian design were constructed. The second phase envisaged construction of plutonium fuelled fast breeder reactors which produce power and breed ^{233}U from ^{232}Th. In the third phase, we propose to build thorium fuelled power reactors.
- At present nuclear power generated by 10 reactors is around 2170 MWe and it is hoped that by the end of this century we shall generate 3000MWe-4000MWe.
- Nuclear radiations have been an integral part of evolution of life on this planet without adverse effects. But man made sources pose a potential threat from rise in the level of these radiations in our environment.
- Nuclear radiations have two types of biological effects: Somatic and genetic. Fast neutrons are most dangerous.
- Radioisotopes and radiations coming from them are being used as tracers, in radiography, power generation in unmanned weather stations and satellites. These have also yielded fruitful results in increasing as well as preserving food stocks,

14.7 TERMINAL QUESTIONS

1. A water pipe made of cast iron runs through wet soil and is suspected to have a leak. Suggest a way of detecting the leak by using a tracer.

2. A substance X decays by β-emission into a daughter Y. A rock contained some amount of X but no Y at the time of its creation. What would be the ratio of the amounts of Y to X when the age of the rock is $T/20$, T and $20\,T$ respectively; T being the half-life of X? Also explain from the result of this problem, why it is necessary in isotopic dating to use an isotope with a half-life comparable to the age being measured.

3. A 2.5 g sample of dead wood is found to give 36 β counts in a run of one hour. The self-absorption of βs by the wood is known to be 40%. Estimate the age of the sample.

14.8 SOLUTIONS AND ANSWERS

TQs

1. The leak can be detected by mixing a β-emitting solute with the water near the suspected leak and searching for the β activity along the length of the pipe. The pipe will stop all βs emitted inside it and the activity will be found only at the point of leak.

2. Amount of Y/Amount of $X = 0.0353$, 1 and 1.048×10^6 respectively. When the age to be measured is too small compared to the half-life, the amount of daughter present is too small to measure accurately. On the other hand, when the age is too long, practically all the parent isotope has decayed and the amount of parent cannot be measured accurately. Hence the necessity of an isotope of half-life comparable to the age being measured.

3. No. of true counts per min. per g. $= \dfrac{360}{0.6 \times (60\ \text{min}) \times (25\ \text{g})}$

$= 4.0$ per min per g.

Hence, $\text{Req. age} = \dfrac{1}{1.21 \times 10^{-4}} \ln\left(\dfrac{15.3}{4.0}\right) = 11100$ yr

UNIT 15 ELEMENTARY PARTICLES

Structure

15.1 Introduction
Objectives
15.2 Discovery of Positron: The First Antiparticle
15.3 Early Systematics of Particles
15.4 The Machines of Nuclear Physics
Particle Accelerators
Particle Detectors
15.5 Inventory of Particles
15.6 Conserved Quantities
15.7 The Quark Model
15.8 Summary

15.1 INTRODUCTION

You have now learnt about the structure of nucleus and the forces which hold the nucleons together. Experiments designed to study the properties of nuclei resulted in the discovery of new particles. You are familiar with electrons, protons, neutrons and photons. These particles, with the exception of free neutrons and possibly protons, are stable; left to themselves they would live forever. Many unstable particles are created due to conversion of energy into mass in high energy collisions.

The existence of one of these extraordinary particles - the positron - was predicted by Dirac in 1928 on the basis of his mathematical theory developed to describe the properties of the electron. The positron, then completely unknown, was supposed to be identical to an electron but with positive electric charge. The experimental confirmation of this consequence of Dirac's theory came from the work of Anderson on cosmic rays in 1932. It marks one of the most beautiful examples of how far human intellect can probe into the mysteries of nature by logic. You will learn these aspects in Sec. 15.2 without going into nicities of mathematics. Subsequently positrons were produced artificially by the bombardment of targets with beams of high energy charged particles – protons or electrons – from accelerators. The early systematics of elementary particles are discussed in Sec. 15.3. In Sec. 15.4 you will get a general overview of the principles of particle accelerators used to produce them and particle detectors used to 'see' them. You will appreciate that particle accelerators are very convenient tools of nuclear research; experiments can be carried out in a controlled and efficient manner. As a result, physicists have constructed more and more powerful accelerators. Today we have moved to giant installations. And this has made the inventory of elementary particles richer; today there are more than 300 such particles. Fortunately, this zoo of elementary particles in all its chaotic diversity can be broadly classified into a few categories. You will learn about it in Sec. 15.5.

You may now like to know: How elementary are elementary particles? Can they be regarded as fundamental building blocks of matter? OR what makes up matter? How do we measure their properties and analyse their dynamics? Other related questions are: Do nucleons have a sub-structure? If so, how do we probe such a constituent configuration? The sustained efforts directed to answer these questions have put physicists on the threshold of achieving an understanding of particle dynamics and structure comparable to our knowledge of atomic and nuclear structure. Though we do not intend to present theoretical details, a general discussion of the supposedly ultimate theory of matter is presented in the last section.

Objectives

After going through this unit, you will be able to

- explain Dirac's logic for the existence of positrons
- explain the working principles of particle accelerators and detectors used in studying particle dynamics and structure

- classify elementary particles into four broad categories and explain their decay processes, and
- describe quark model of matter.

15.2 DISCOVERY OF POSITRON: THE FIRST ANTIPARTICLE

From Unit, Block of this course you would recall how Goutsmit and Uhlenbeck, in attempting to account for anomalies in X-ray spectra, suggested that every electron has a permanent spin: That it rotates about its own axis with an angular momentum of $\frac{\hbar}{2}$, where $\hbar = \frac{h}{2\pi}$, h being Planck's constant. Though it disposed off the spectroscopic inconsistency, the new model of the atom was difficult to comprehend theoretically; nothing in the then prevailing views required the electron to spin. And need arose to transform spinning electron hypothesis from adhoc assumption to a satisfactory quantum theory.

In 1928, Dirac developed a wave equation for electrons in an electromagnetic field which incorporated special relativity into quantum mechanics. His procedure was mathematically highly elegant. Dirac's equation gave the electron a magnetic moment of one Bohr magneton $\left(= \frac{e\hbar}{2m}\right)$ and a spin angular momentum of $\hbar/2$. The experimental verification of these predictions followed soon. The most striking result of this theory, however, was that relativistic equation had solutions corresponding to negative energy states. That is, electrons can exist in states of positive energy as well as negative energy states:

$$E = \pm \sqrt{m_0^2 c^4 + p^2 c^2} \qquad (15.1)$$

The positive root permits the total energy of an electron to have any value from m_0c^2, the rest energy to ∞, as the momentum increases. The negative root permits values from $-\infty$ to $-m_0c^2$. Any electron in a positive energy state would make a transition to a negative energy state; the difference in energy would be radiated in the form of a photon of appropriate energy and momentum. Similarly, an electron in a negative energy state could fall to a still more negative state. Since every system has an inherent tendency to evolve in the direction of minimum energy, you may be tempted to think that all the electrons should have energy $E = -\infty$! But this seemed to defy all known tenets and Dirac's theory encountered a serious difficulty. In Dirac's own words:

"This seemed to be a stumbling block to begin with, but it turned out that one could get over that difficulty in a very neat way, at the expense of changing one's concept of the vacuum".

The vacuum was thought to be a region where there was nothing at all. Dirac interpreted the vacuum as a state of lowest energy. He argued that if the electrons can have negative energies too, one would want to have as many of these electrons as possible in order to get the lowest energy. However, Pauli's exclusion principle allows no more than one electron in a state. Therefore, Dirac proposed that vacuum is one in which all negative energy states but no positive energy state are occupied. Again, Pauli's exclusion principle will be operative to prevent any catastrophic transition of an electron with positive energy into a negative energy state in vacuum (as all states are already filled). That is, such electrons are not directly observable. However, a transition can take place to a negative energy state provided, to begin with, it is unoccupied. Absence of occupation of a negative energy state was interpreted as the presence of a hole in the sea of filled states. The hole behaves like a particle of positive charge, in much the same way as a hole in the valence band of a semiconductor. The hole state is an inevitable consequence of relativistic theory of spin ½ particles. When a positive energy electron jumps into the hole, both disappear. Physically, an electron annihilates a positron and their energy will be emitted as chargeless photons.

The existence of positrons was demonstrated by Anderson in 1932 when he discovered the tracks of electron -like particles of positive charge while observing cosmic ray particles in a cloud chamber. He had placed his cloud chamber in a strong magnetic field. This enabled him to distinguish particles of negative and positive charges by their directions of deflection. He observed some electron - like tracks in the direction corresponding to positive charge.

Subsequent measurements established that positrons are of the same mass, same spin and same magnetic moment as electrons. That is, but for the positive charge, positrons are identical with electrons in all respects.

15.3 EARLY SYSTEMATICS OF PARTICLES

The set of familiar particles out of which all species of nuclei, atoms and molecules can be made are protons ($m_p c^2 = 938.2$ MeV, $s = \frac{1}{2}$, $q = +1$unit), neutrons ($m_n c^2 = 939.6$ MeV; $s = \frac{1}{2}$, $q = 0$), electrons ($m_e c^2 = 0.511$ MeV, $s = \frac{1}{2}$, $q = 1$) and photons ($m_\gamma c^2 = 0$, $s = 1$, $q = 0$). To this set we may add the antiparticles $\bar{p}$, $\bar{n}$ and e^+.

You now know that Yukawa(in 1935) had postulated the existence of pions as quanta of strong nuclear force between nucleons. (This was much like a photon, which is responsible for electromagnetic interaction among charged particles.) From the short range nature of the nuclear force, he deduced that the mass of these particles would be a few hundred MeV (0.113u). Moreover, the coupling strength should be characteristically large to offset the Coulomb repulsion. With the advent of cyclotrons that could accelerate protons to several hundred MeV and the development of techniques to study cosmic rays, the experimental evidence for the existence of pions ($\pi^\pm$ mesons) came from the work of Powell and Occhialini more than a decade later. We now know that there is a triplet of them; π^+, π^0 and π^-, carrying +1, 0, –1 units of charge ($m_{\pi^\pm} c^2 = 140$ MeV, $m_{\pi^0} c^2 = 135$ MeV, $J = 0$).

While electrons are absolutely stable and protons are believed to be so, neutrons and pions are unstable. Their decay results in a number of new particles:

$$n \rightarrow p + e^- + \bar{\nu}_e$$

$$\pi^- \rightarrow \mu + \bar{\nu}_\mu$$
$$\downarrow$$
$$e^- + \bar{\nu}_e$$

$$\pi^+ \rightarrow \mu^+ + \nu_\mu$$
$$\downarrow$$
$$e^+ + \nu_e$$

A neutron decays into a proton, an electron and an antineutrino. A pion decayed into a muon, which in turn decayed into an electron. These decays are similar in nature to radioactivity; the probability of decay is independent of the age of the particle and is characterised by the decay half-life. We find $\tau_n = 896$s, $\tau_{\mu^\pm} = 2.6 \times 10^{-8}$s, $\tau_{\pi^\pm} = 0.7 \times 10^{-16}$s. New species of particles produced include neutrinos, which are (perhaps) massless, carry no charge and are spin ½ particles. They play no part in strong nuclear interaction and being neutral, experience no direct electromagnetic force. Since their interaction is very weak, typical of β-decay, they are elusive and difficult to detect. Their presence is very often inferred from missing momentum and energy in a reaction.

Muon (μ meson) is the principal decay product alongwith its species of neutrino (υ_μ) of pion. This was observed in the cosmic ray showers by Anderson in 1937 and was initially mistaken for the pion, the postulated mediatory particle in nuclear interaction. It used to be called μ meson for this reason. But it turns out to be just a fat electron. Like electron, it has no role in nuclear force but since it carries the same charge, the electromagnetic properties are very similar. It is about 205 times heavier than electron. Electrons, muons and neutrino's are collectively known as Leptons.

Ever since the discovery of nucleus in an atom, scientists have been interested in knowing the structure of nuclei and the forces which hold the nucleons together. To look into the nucleus one has to somehow disintegrate the nucleus. The first such disintegration of nitrogen nucleus with α-particles was achieved by Rutherford in 1919. The energies of these naturally occurring projectiles were far in excess of the energies available otherwise in the laboratory. To be able to perform disintegration experiments efficiently, need was felt for artificial sources of high energy particles. These sources, now called particle accelerators, are the machines of nuclear physics which produce high speed particles. Today, accelerators are the indispensable tool of nuclear research. You will learn about these now.

15.4 THE MACHINES OF NUCLEAR PHYSICS

The machines of nuclear physics can be broadly classified into two groups: Particle accelerators and Particle Detectors. While accelerators facilitate availability of high energy particles, detectors enable detection of particles created in high-energy collisions. We will now discuss these separately.

15.4.I Particle Accelerators

The simplest way to accelerate a charged particle is to 'drop' it through a high potential difference (pd). If a particle of charge q is made to move through p.d of V, it acquires a kinetic energy of qV. The largest p.d that can be maintained under accelerator conditios is 10^7V and ions acquire an energy of about 10MeV per unit charge. This is just the energy we need for many nuclear structure studies. The technology of electrostatic accelerators therefore consists of establishing and maintaining a high voltage terminal to accelerate the charged particles from the source.

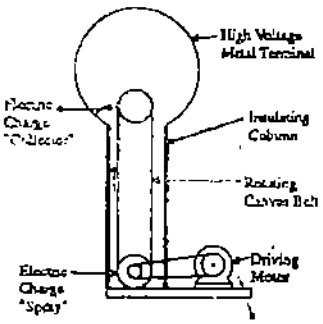


Fig.15.1: Schematic of the van de Graaff generator.

Though the development of particle accelerators began in late 1920s, the first particle accelerator was constructed by Cockcroft and Walton at the Cavendish Laboratory, Cambridge in 1932. It was an electrostatic accelerator in which charged particles were accelerated by means of a strong electrostatic field produced by a large amount of electrostatic charge. This machine was the first to "split" a nucleus of ^{7}Li with artificially accelerated protons of about 400 keV energy:

$$p + {}^7Li \rightarrow {}^4He + {}^4He$$

The most common type of electrostatic accelerator in use today in nuclear physics laboratories is based on the Van de Graaff generator, shown schematically in Fig. 15.1. When a charged inner conductor and a hollow outer metallic conducting shell are placed in electrical contact, all the charge from the inner conductor will flow to the outer one. The resulting potential on the outer conductor, given by $V = Q/C$, can be increased, in principle, without limit by adding more and more charge. The charge is transferred mechanically through a continuously moving belt and brush system. In going from one terminal to the other, voltages of about a few million volts enable charged particles to gain energies of a few MeV.

A basic disadvantage of all electrostatic accelerators arises from insulation breakdown at high voltages. Perhaps the weakest link in an accelerator facility is the ion source. Discharge filaments may burn out and require replacement and changing the type of accelerated ion often required changing the ion source, partly or fully. Placing the source inside the high voltage terminal creates annoying problems, which force accelerator shut down for sometime. Therefore, it was thought necessary to have machines which could accelerate particles in numerous small steps of low voltages. Such a machine was called the cyclotron, and formed the basis of a whole series of particle accelerators — from synchrotrons to super colliders. These machines have been developed the world over in a short span of time; their size growing bigger with the increasing need of more and more high energy particles. Their number in the world is now well over a few hundred. We have discussed the physics aspects of these machines in the next sub-section.

A. Cyclotron

A cyclotron is an accelerator in which charged particles are accelerated by a constant frequency alternating electric field that is synchronised with the movement of particles in spiral paths in a constant magnetic field normal to their path.

Lawrence was awarded the 1939 Nobel Prize in Physics for the invention of cyclotron.

The essential design of the first cyclotron was conceived in 1929 by Lawrence at the University of California, Berkeley (U.S.A.). In 1932, this machine was used by Lawrence and Livingston to disintegrate a nucleus with 1.2MeV protons, the highest energy then available in the laboratory. In the following years, they built a series of progressively larger cyclotrons and achieved energies of up to 27 MeV with proton projectiles.

Refer to Fig. 15.2 which schematically illustrates a cyclotron, which consists of two hollow semi-circular metallic boxes D and D', usually called the dees because of their shape. These dees are placed in an evacuated chamber C which is fixed within the pole faces of an electromagnet. The magnetic field, which acts perpendicular to the plane of the dees is uniform in the centre, but decreases slightly as one moves away from it. The dees are connected to a high frequency oscillator to provide them with a high frequency alternating

potential. The dees then act as electrodes. When one dee is positive, the other becomes negative, and vice versa.

Positive ions are produced in the ion source S placed near the mid-point of the gap between the dees. If at the time of formation of a particular ion the dee D is negative, the ion gets accelerated towards it. Once inside D, it moves in a circular path under the influence of the magnetic field quite unaffected by the electric field. If at the time of its exit from D the oscillator reverses the polarity of D' making it negative, the ion again gets accelerated in going from D to D'. As a result, the ion gains kinetic energy, which increases its speed. It then moves in a circle of greater radius in D'. On exit from D', it again gets accelerated if D by then becomes negative. This process continues till the ion reaches the periphery of the dees where it is pulled out by applying a strong negative potential to the deflecting plate P, and is made to bombard the target at T. In short, the magnetic field "guides" the ion to follow a circular path while the electric field gives it small "pushes" at regular intervals to increase its speed. This makes the ion path a spiral of increasing radius. When a particle of charge e and speed v moves in a field of strength B, the Lorentz force in the circular orbit, evB, provides the necessary centripetal acceleration to maintain the circular motion:

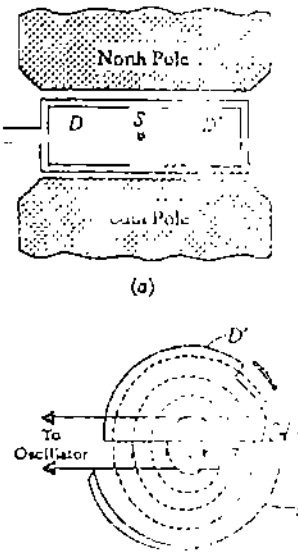


Fig.15.2: Principle of cyclotron: (a) Vertical view of the main part of the machine, (b) Path of the particle within the dees.

$$F = evB = \frac{mv^2}{r} \qquad (15.1)$$

where r is the instantaneous radius of the orbit.

We can rewrite this expression as

$$r = mv/eB \qquad (15.2)$$

so that the time taken to traverse a semicircular orbit is given by

$$t = \frac{\pi r}{v} = \frac{\pi m}{eB} \qquad (15.3)$$

The orbital frequency can be calculated from the expression

$$\nu = \frac{1}{2t} = \frac{eB}{2\pi m} \qquad (15.4)$$

The maximum kinetic energy of the particle when it reaches the outer most (largest) radius (R) of the cyclotron is given by

$$\text{K.E.} = \frac{1}{2} m v_{max}^2 = \frac{1}{2} \frac{e^2 B^2}{m} R^2 \qquad (15.5)$$

This relation shows that it is advantageneous to build cyclotrons with large fields and large radii. You may now like to answer an SAQ.

It is important to emphasize here that in a cyclotron, the angular velocity of the ions remains constant. That is, the slower ions move in circles of smaller radii while the faster ones in those of bigger radii. This forces all the ions to return to the gap (between the Ds) at the same time irrespective of their speeds. This means that the time a particle takes in traversing one semicircular path is independent of the radius of the path. The ions always gain energy in the gap provided the oscillator frequency is such that the time taken in reversing the polarities of the dees equals the time ions take in traversing the semi-circles. The final energy of the ion is equal to that gained in one passage through the gap multiplied by the number of such traversals.

The energy acquired by the particles in a cyclotron is independent of the dee voltage. When voltage is small, the particles make large number of turns before reaching the periphery.

Note that in a cyclotron, we can accelerate only charged particles (i.e. protons, deuterons, alphas, ions of light and heavy elements) since an electric field can exert force only on the charged particles.

The diameter of the magnetic pole faces is often used as a measure of the size of a cyclotron. The dees of the first Berkeley cyclotron had a 12.5cm radius. This cyclotron produced 1.2MeV protons in a field of 1.3T (13kG); the corresponding frequency being about 20 MHz.

The energies to which various particles can be accelerated in a cyclotron are limited by (a) the decreasing magnetic field near the periphery of the pole faces (Fig.15.3) necessary for focussing (keeping particles on their path in the horizontal plane), and (b) the relativistic increase in mass of the particles at high velocities (near the velocity of light). Because of these, the angular velocity of the ions becomes low near the periphery of the magnetic pole faces and they start getting out of step with the frequency of the alternating voltage. As a result, the ions are not further accelerated.

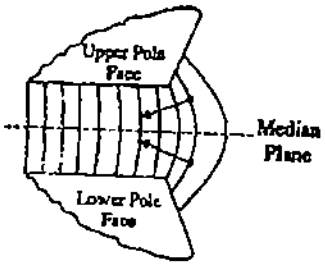


Fig.15.3: The magnetic field lines deviate from the vertical near the edge of pole faces

Ordinarily, a cyclotron can accelerate particles to velocities upto about $0.2c$, where c is the velocity of light. The relativistic increase in mass at this velocity is about 2%. This limits the size of the cyclotron and the energies to which various particles can be accelerated. The limit for accelerating protons is about 25 MeV. Deuterons and alpha particles can be accelerated to still higher energies.

For such particles, the expression for orbital frequency modifies to

$$\nu = \frac{eB}{2\pi m}\left(1 - \frac{v^2}{c^2}\right)^{1/2}$$

B. Synchrocyclotron

For many nuclear experiments, we need particles of a few hundred MeV of energy. It is possible to push up the particle energies in two different ways:

- By continuously varying the frequency of the accelerating voltage from some maximum to a minimum value, as the particles move away from the centre of the machine. The cyclotrons based on the this modification are called the **synchro-cyclotrons** or the **frequency modulated (FM) cyclotrons.** These machines do not produce particles in a continuous stream, but in small bursts. Energies of about 750 MeV for protons with field strengths of about 2.3 T corresponding to a frequency of about 20 MHz have been obtained using the Berkeley synchrocyclotron. (Other comparable machines have operated at Dubna (erstwhile USSR) and CERN in Geneva. The FM cyclotrons are usually used in high- energy (elementary particle physics) research.

- By continuously increasing the intensity of the magnetic field from the centre towards the periphery of the magnetic pole faces. (Here frequency of the accelerating voltage is kept fixed.) With the increasing magnetic field, the particles experience greater bending force and move in circles of smaller radii. The shortened paths compensate for the decreased angular velocity at high speeds. However, a field which increases as we move away from the centre of the machine can cause defocussing in the vertical direction. Because of the vertical defocussing, particles move up or down and may get lost in the walls of the dees. In 1938 Thomas suggested a way out but his suggestion was not put to test for over a decade probably due to the advantages gained through the development of synchrocyclotrons. The first two machines based on Thomas' idea, called **isochronous-cyclotrons, sector focussing or AVF (azimuthally varying field) cyclotrons,** were constructed at Berkeley in 1952. But these machines generated real interest only after 1955 when Lawrence reported about them at the "Atoms for peace" conference in Geneva.

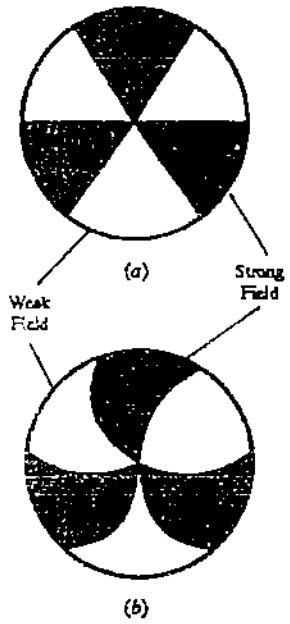


Fig.15.4: Pole faces of AVF cyclotron. Shaded areas represent ridges or "hills" while unshaded areas represent empty spaces or "valleys" (a) Radial ridges, (b) Spiral ridges

C. AVF Cyclotrons

In AVF Cyclotrons, the magnetic pole faces are fastened with iron slabs either in the radial direction or spiraling outward in gentle curves (Fig.15.4). These slabs, called sectors or ridges, have gaps (empty space) in-between and these are of about the same size. As the pole gap between the ridges is small, the magnetic field is stronger there compared to the one in the "valley" regions. This gives alternating regions of strong and weak magnetic fields. The magnetic field depends on the radius as well as the azimuthal angle. Such a field provides strong focussing of particles, particularly if spiraling sectors are used. The path of the particle in these machines is only nearly circular, with pronounced curvature in the "hill" regions than in the valleys (Fig. 15.5). The major advantage of AVF cyclotrons over synchrocyclotrons is the continuous beam and thus the larger possible beam currents (100A). Energies beyond 600MeV for protons have been obtained using these machines.

Most of the present day cyclotrons operate over wide range of oscillator frequency and can, therefore, accelerate particles to a desired range of energies (of course, upto a maximum possible from the machine in question). These are then also called the variable energy cyclotrons (VEC). In them, the strength of the magnetic field is also varied to make the particles return to the gap "in step" with the frequency of the oscillator. In India, we have a VEC at Calcutta.

The increase in magnetic field from the centre to the periphery of the magnet in an AVF cyclotron is often achieved by the use of some pairs of concentric circular coils, usually called the "trim" coils. These coils are placed on the upper and the lower pole faces of the electromagnet. By passing suitable current through these coils appropriate increase in magnetic field with radius may be obtained.

Particle beams from these machines are usually of good intensity but of large variation in energy. However, by magnetically analysing the beam before bombardment, this energy variation may be decreased even to about 5keV. Usually cyclotrons are best suited for intermediate energy (> 5MeV) experiments; for low-energy work, electrostatic accelerators are preferred. To decrease the output energy from cyclotrons, sometimes metallic foils are inserted in the path of the beam before the bombardment.

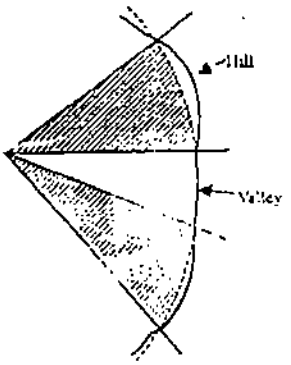


Fig.15.5: Trajectory of a particle in an isochronous cyclotron Circular path, for reference, is shown by a shaded line

D. Synchrotrons

You will now agree that extending the cyclotron or synchrocyclotron to higher energy means that we have to build machines of larger radii. This results in huge in-put costs. Therefore, to partly overcome this problem, it was thought necessary to vary the magnetic field strength as well as the resonant frequency. Such accelerators are termed **Synchrotrons**. Fig. 15.6 shows the simplest design of such a machine. The most critical feature of this device is that the orbital radius of the particles is (nearly) constant even at high energies.This is achieved by increasing magnetic field only within an evacuated circular beam pipe that serves as "racetrack" for the accelerated particles. Recall that in an ordinary cyclotron, the field is applied over the entire volume. Particles traversing a circular path are accelerated by the electric field while crossing the gap in each orbit. As the energy increases, the frequency of the voltage across the gap is increased to maintain the orbital frequency constant and the magnetic field strength is increased to keep the orbital radius constant. It is important to realise that the field is varied in time, not space as in the AVF cyclotron. Since the accelerator uses a varying frequency and magnetic field, it gives out bursts. All the high-energy proton accelerators are synchrotrons.

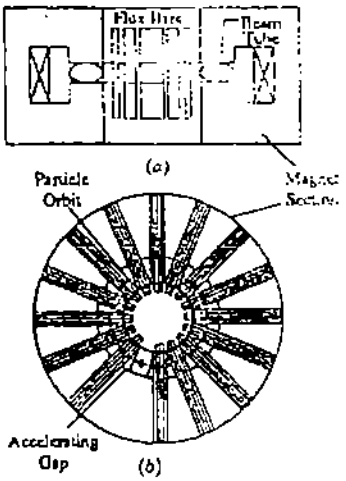


Fig. 15.6 : Vertical and horizontal cross-sectional views of electron synchrotrons. The magnets bend the beam into a circle and the electric field accelerates the particles once in each orbit

Table 15.1: List of Some Accelerators in Operation/ Under Development

Accelerator	Start of Operation	Particles	Energy (GeV)
Brookhaven Alternating Gradient Synchrotron (AGS), New York	1961	Protons	33
Standard Linear Accelerator (SLAC), California	1961	Electrons	22
Cornell Electron Synchrotron, New York	1967	Electron	12
Serpukhov Proton Synchrotron, USSR	1967	Protons	26
Fermilab Main Ring, Illinois	1972	Protons	500
Deutsches Elektronen Synchrotron (DESY), Germany	1974	Electrons	22
CERN Super Proton Synchrotron (SPS), Switzerland	1976	Protons	500
Fermilab Tevatron, Illinois	1985	Protons	1000
Japanese National Laboratory (KEK)	1986	Electrons	30
CERN Large Electron-Positron Storage Ring (LEP)	1989	Electrons	85
Superconducting Super Collider (SSC), USA	1995	Protons	20000

The first large synchrotron was the Bevatron at Berkeley, which produced protons of energy upto 6.4GeV. By bombarding a target with this proton beam, E.G. Segre and O. Chamberlain succeeded in creating antiprotons in 1955.

Table 15.1 lists some accelerators now in operation/under development. The two largest accelerators in operation are the proton synchrotrons at the Fermi National Accelerator Laboratory (Fermilab, Chicago) and at the Centre Europeen de la Recherche Nucleaire (CERN, on the Swiss-French border near Geneva). In the Fermilab, the accelerator is buried underground in a circular tunnel with a radius of 1km. The CERN accelerator is slightly larger, with a radius of about 1.2km. This accelerator system starts with a Cockroft-Walton generator (800keV) at the first stage. The energy of the beam is then pushed to about 200MeV in a linear accelerator and 26GeV in a proton synchrotron. Following this stage, the beam can be steered into Super Proton Synchrotron which raises the energy beyond 400GeV. The Fermilab accelerator produces beams of protons with an energy of 1000 GeV or 1 TeV, and a speed of 99.99995% of the speed of light.

The next largest accelerating machine is the Stanford Linear Accelerator (SLAC), 2 miles long. This accelerator produces 22 GeV electrons whose speed is within 8 cm s^{-1} of the speed of light. For such electrons, a circular accelerator is ruled out by pragmatic considerations; the centripetal acceleration would cause the electrons to lose a prohibitive amount of energy by electromagnetic radiation (synchrotron radiation). Once the projectiles have been given their maximum energy, they are guided out of the accelerator by steering magnets and directed against a target consisting of a block of metal or a tankful of liquid. The reactions that occur in the collisions between the projectiles and the protons and neutrons of the target create many new particles by conversion of energy into mass. However, only the centre of mass energy of the incident particles is available for reactions and in the quest for production of new particles, an accelerator designer intends to convert maximum possible energy into mass. This has led to the design of the colliding beam accelerators. We will discuss these now.

c. Colliding Beam Accelerators

When a proton of energy E_{lab} strikes a stationary proton of mass m_p, the centre of mass energy in the reaction is given by

$$E_{cm} = \sqrt{2m_p c^2 E_{lab}}$$

Thus when a 30 GeV proton hits a stationary target, available energy for the reaction in the centre of mass is about 7.6GeV only. However, if it is possible to store a pulse of the 30 GeV proton in a ring in say clockwise direction and another pulse in a similar ring in anticlockwise direction and allow them to intersect at a chosen location of a detector, all of the 60GeV energy is available for interaction. Of course, while a stationary target has huge number of protons available for collision, in the collider the collisions are much rarer in view of the limited number of protons in a typical beam. For a collider to be effective, we need the beam current to be as high as possible and also have the beam sharply focussed to increase the probability of interaction in a collision. Notwithstanding what one loses in statistics, one gains in increased energy scale for the interaction.

A 30GeV proton on collision with a stationary proton target has enough energy to produce an anti- proton. The $\bar{p}$ produced can be accumulated and injected into the supersynchrotron so that, as the magnetic field increases, $\bar{p}$ gets accelerated simultaneously with the protons circulating in the same machine. On allowing them to collide at predetermined spots, where a composite detector is placed, we can observe a $\bar{p}p$ collision in the centre of mass. Entire energy of this collision is available for the nuclear interaction. In one such experiment at CERN, with 270GeV protons and 270GeV antiprotons, it was possible to discover new species of particles W^+ and W^- of 81GeV mass.

Stanford (US), DESY (Hamburg), Frascati (Italy) and CERN (Geneva) have colliders in which e^+ and e^- are the accelerating particles. A charged particle, while moving uniformly in a circular orbit, is subject to centripetal acceleration. Charges, when accelerated radiate; the energy loss being proportional to the square of the acceleration. Since the acceleration is inversely proportional to the radius of curvature, larger orbits are necessary to minimise this energy loss in a circular machine. Lighter electrons radiate much more than heavier protons, the energy loss being proportional to the fourth power of their masses. Hence the electron machines have typically much larger diameter than proton machines of comparable energy; 50GeV × 50GeV e^+e^- collider has a circumference of 26.7 kms. In comparison the 400GeV SpS at CERN has a circumference of 6.9 km.

15.4.2 Particle Detectors

The devices used to detect elementary particles, for example those created in high-energy collision experiments at accelerator laboratories, are called detectors. The detection of a particle is made possible by the property that a charged particle loses energy by ionising or exciting the atoms of the medium. The ions left behind cause an electric pulse which can be sensed electronically or act as nuclei for condensation of supersaturated vapour or formation of bubbles in a superheated liquid. Accordingly, detectors can be classified into two broad groups: **counters**, such as Gieger-Muller counter, scintillation counter and Cherenkov counter, that register the passage of each charged particle; and track-recording devices, such as cloud chamber, bubble chamber, spark chamber, streamer chamber, and proportional chamber that provide pictures of the paths of the particles. We begin this discussion by describing counters.

A. Counters

The first ingenious device used to detect charged particles was devised by Geiger. It is based on the phenomenon of ionisation by collision. Later on, Muller introduced important modifications in this 'counter' and today we refer to it as the Geiger-Muller (G.M.) counter. You may get an opportunity to have a closer look at it when you visit your laboratory for PHE-12 (L) course. G.M. counter consists of a partially evacuated glass tube containing an open copper tube as cathode. A tungsten wire is stretched along the axis of the tube. When a charged high energy particle enters the tube, the gas in the glass tube is ionised giving a pulse of current, which operates a counting device. It is sensitive to the production of even a single ion pair and gives constant output, independent of initial ionisation.

Scintillation counter is a universal counter. It consists of a scintillating material (inorganic crystals, plastics or organic liquids) that gives off a brief and faint flash of light when struck by the (charged) ionising particles. This luminescence is recorded by a photomultiplier and counted by an electronic circuit (Fig. 15.7). These counters are robust, simple and efficient giving large sharp output pulses.

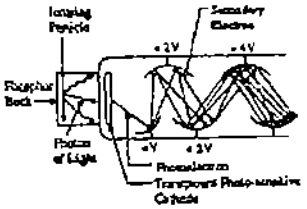


Fig. 15.7 : Schematic view of a scintillation counter

Cherenkov counter consists of a volume of dielectric (usually a tank filled with gas at high pressure) within which the speed of light is less than $3 \times 10^8 ms^{-1}$. When a high-energy charged particle traverses this dielectric, its speed will exceed the speed of light. Under these conditions, the particle emits an electromagnetic shock wave, analogous to the sonic boom emitted by a supersonic aeroplane. This Cherenkov radiation can be picked up by a photo multiplier tube. The usefulness of this counter lies in the fact that the measurements can be made at any given angle. This information enables an experimentalist to know the velocity of the ionising particle.

B. Chambers

The Wilson Cloud Chamber has the unique distinction of being the first chamber used to 'see' the tracks of the charged particles in their passage through matter. When air mixed with saturated vapour of water, alcohol or ether contained in a cylinder fitted with a piston is expanded rapidly, the vapour becomes supersaturated and condenses in a cloud of water drops. The production of such a cloud is facilitated by the presence of ions produced by charged particle induced ionisation. As super-saturation increases, the negative ions serve as centers of condensation. But with increasing volumes, both positive and negative ions serve as the nuclei of droplets. A close array of fine droplets gives rise to a cloud track, which is photographed to obtain a 'record'. The cloud chamber suffers from the limitation that it requires sometime to recover after the expansion and therefore can not give continuous record of events.

The bubble chamber consists of a tank filled with a liquified gas, such as liquid hydrogen, helium or freon at a high pressure (5-20atm) and a temperature slightly below its boiling point. The chamber is suddenly expanded adiabatically by means of a piston attached to the tank. This lowers the boiling point and the liquid becomes superheated. The liquid is then unstable and begins to boil as soon as some disturbance supplies energy for the formation of the first few bubbles. A high-energy charged particle passing through the chamber ionises molecules of the liquid along its track. The electrons so freed quickly deposit their energy in the liquid and trigger the formation of bubbles. Thus, a fine trail of bubbles marks the track of the charged particle. An array of high-speed cameras take photographs of these trails of bubbles, after about 10ms of their formation. Then, within a few hundredths of a second, the chamber is compressed again. This quenches the bubbles and the chamber is ready for the next cycle of operation.

The bubble chamber at CERN is surrounded by a massive electromagnet whose magnetic field curves the orbits of the particles and permits the determination of their momenta from the observed radii of curvature. Photographs of the tracks of particles passing through the bubble chamber are taken simultaneously with several cameras at different angles so as to obtain a stereoscopic view of the tracks. Bubble chambers are expensive to operate and maintain, but they attain a higher spatial resolution than any other track-recording device. Careful measurements of the photographs determine track positions to within 0.05mm, and the curvature of the track in the magnetic field determines the momentum of the particle to within about 0.1%.

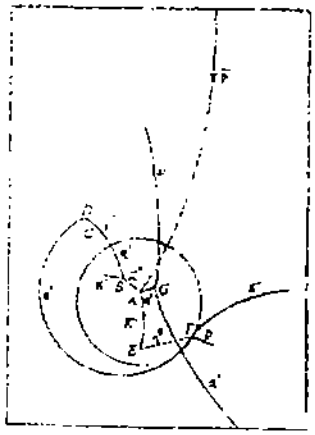


Fig. 15.8: Trace of the photograph of tracks of particles in a bubble chamber

Fig. 15.8 is an interesting trace of the photograph of tracks of particles made visible with a bubble chamber at CERN. It shows a sequence of events involving the creation of several particles in the collision of an antiproton and a proton, and the subsequent decays and collisions of these particles. The antiproton ($\bar{p}$) enters the field of view from above; its track is slightly curved to the left by the magnetic field. This antiproton is produced outside the bubble chamber by the impact of a beam of protons on a target of metal. At the point *A*, the antiproton collided with a proton at rest in the liquid of the bubble chamber. The antiproton and the proton are destroyed in this collision, which creats two kaons ($K^\circ.K^-$) and two pions (π^+, π°). By a remarkable coincidence, all of these particles, except for the π° cause further events within the field of view of the bubble chamber. The K° particle is electrically neutral and hence leaves no visible track in the bubble chamber; nevertheless, it is possible to reconstruct its trajectory because, after a short while, it spontaneously decays (at *B*) into two pions that leave tracks. One of these pions further decays into an antimuon and a neutrino (at *C*), and the antimuon in turn decays into an antielectron and two neutrinos (at *D*). Meanwhile, the K^- created in the original antiproton-proton collision suffers a collision with another proton at rest in the liquid of the bubble chamber (at *E*). This collision has led to the creation of a lambda particle (Λ°) and a pion (π°). The lambda particle is neutral and leaves no visible track. However, it decays into a pion (π^-) and a proton at (*F*). Furthermore, one of the pions created in the original antiproton-proton collision suffers an elastic collision with another proton at rest (at *G*), which causes this proton to recoil.

Spark chambers yield somewhat cruder pictures of particle tracks than bubble chambers, but they are much simpler. A spark chamber consists of many parallel screens or thin plates of metal, each separated from the next by a gap of about a centimeter. The space between the plates is filled with a gas, usually neon. Alternate plates are connected to the positive and the negative terminals of a high-voltage supply, which produces an electric field of about $10^6 Vm^{-1}$ between the plates. A high-energy charged particle passing through the chamber ionises the gas along its track. The electrons released by this ionization are accelerated in the strong electric field, strike gas molecules and release more electrons, so on. This produces an electric discharge between the plates, with a visible spark. Thus, a succession of sparks marks the passage of the particle through the chamber. Cameras record these sparks photographically; usually mirrors are placed around the spark chamber so that a single photograph can simultaneously record several views of the tracks of sparks seen from several angles.

The spark chambers do not have the high spatial resolution of bubble chambers but they have better time resolution. The high voltage is usually applied to the chamber for a short time, about 10^{-6}s; the chamber is therefore sensitive only during this brief interval. This permits the chamber to capture the track of one individual particle, provided it is triggered at a suitable moment. The trigger is provided by auxiliary scintillation counters placed around the chamber; these make preliminary identifications of the arriving particles and trigger the spark chamber whenever there is an interesting event.

Streamer chambers are similar to spark chambers but they use only one pair of widely separated plates instead of the many closely spaced pairs used in the spark chamber. A very brief pulse of high voltage, lasting only about 10^{-8}s, is applied to these plates. Under these conditions the electrons released by the ionization of the gas in the chamber do not have enough time to develop a full-sized spark from one plate to the other; instead they only give rise to small, faint protosparks, or streamers, which delineate the track of the particle. Streamer chambers are capable of good spatial resolution; when photographed from a direction perpendicular to the plates the streamers determine the positions of points on the track of the particle to within 0.1 mm.

Proportional chamber is one of the oldest devices used for recording ionisation. It consists of a gas-filled cylindrical metal or glass tube maintained at a negative potential and a fine central wire at a relative positive potential to serve as anode. A significant modification of this counter was achieved when single wire was replaced by a grid of thin parallel (anode) wires at positive potential sandwiched between two plates or screens at negative potential.

This is shown in Fig. 15.9. Electrons released by the ionisation of the gas in the chamber drift to the nearest positive wire and give rise to an electric discharge in the strong electric field in the immediate vicinity of the wire. This electric discharge registers as a current pulse on the wire (the magnitude of the current pulse is proportional to the amount of ionization, hence the name proportional chamber). Each wire is connected to its own electric current, and therefore each wire constitutes an independent detector, which signals the location of the discharge. In multiwire proportional chambers, good spatial resolution is attained by filling the entire chamber with a large number of wires in a dense array (typically one wire per millimeter). In drift proportional chambers, the spacings between the wires are much larger but excellent spatial resolution is attained by measuring the time delay between arrival of the high-energy particle in the chamber and detection of the current pulses on the wires. Since the electrons released by ionization drift toward the positive wires at a known speed, the measurement of the time delays permits computation of the precise distances that these electrons travel from their points of origin to the wires. Drift proportional chambers can determine track positions with an accuracy of upto 0.05mm.

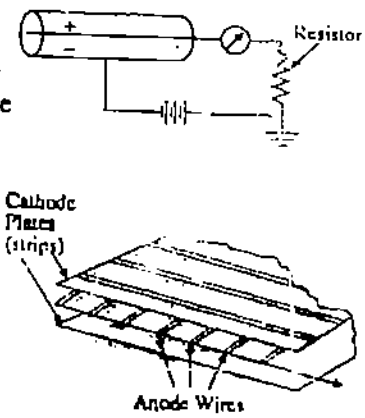


Fig. 15.9 : Arrangement of positive and negative electrodes in a proportional chamber

Proportional chambers do not provide photographs of the tracks of particles. Instead, they record the tracks as a sequence of electric signals from their detector wires. These signals are fed into a computer which constructs an image of the tracks. Because they have good spatial resolution and deliver data in a form that can be directly manipulated by computers, proportional chambers have become very popular in high energy physics.

Shower Detectors and Calorimeters Electrons and photons of high energy cause the occurrence of cascade showers. A parent electron will radiate photons, which convert to pairs and the number of particles increases exponentially with depth in the medium. These electromagnetic shower detectors are built from high Z material with small radiation length (which is defined as the length over which the energy of the particle drops by a factor 1/e). A hadron shower arises when an incident hadron undergoes inelastic collision and produces secondary hadrons and the same process goes on. Various types of electromagnetic and hadron calorimeters are now in use and they become an integral part of a large complex detector.

15.5 INVENTORY OF PARTICLES

The known particles fall into four broad groups: the leptons, the baryons, the mesons and field bosons. Mesons and baryons are collectively designated as hadrons. We will now discuss these in brief.

Leptons are particles that do not take part in strong interactions but are produced and absorbed in weak interactions. Electrons, muons (μ), tau particles (τ) and their associated neutrinos ν_e, ν_μ and ν_τ are members of this group. The anti-particles of these six leptons constitute the family of antileptons e^+, μ^+, τ^+, $\bar{\nu}_e$, $\bar{\nu}_\mu$ and $\bar{\nu}_\tau$. Physicists believe that six pairs of leptons and antileptons complete their family. Leptons have half-integral spins and are fermions. Table 15.2 lists various observable properties, such as their masses, spins, electric charges and lifetimes of leptons, apart from their decay modes.

Table 15.2: Observable Properties of Leptons

Particle	Mass (MeV/c^2)	Spin	Electric charge	Mean life	Principal Decay Modes	
					Mode	Fraction (%)
e	0.511003	1/2	−1	stable		
μ	105.659	1/2	−1	2.1973×10^{-6} s	$e\,\nu\bar{\nu}$	100
τ	1784	1/2	−1	3.4×10^{-13} s	$\mu\,\nu\bar{\nu}$ $e\,\nu\bar{\nu}$ hadrons, neutrals	17.6 17.4 51.6
ν_e	0*	1/2	0	stable		
ν_μ	0	1/2	0	stable		
ν_τ	0	1/2	0	stable		

* According to recent estimates, the mass of the neutrino is about 20 eV/c^2.

you have learnt about the discovery of the muon in Sec. 15.3. The tauon was discovered in 1975 by M. Perl and his co-workers in e^-e^+ collision experiments at the Stanford Linear Accelerator. (He is co-recepient of the Physics Nobel Prize for the year 1995.) The muon and the tauon are essentially heavy versions of electrons. The tauon decays preferably into hadrons.

The neutrinos (ν_e, ν_μ and ν_τ) are (perhaps) massless, uncharged particles. The electron neutrino, ν_e, is emitted in β^--decay. The muon and tauon neutrinos, ν_μ and ν_τ are emitted in the decays of the muon and the tauon:

$$\mu \rightarrow e^- + \nu_\mu + \bar{\nu}_e$$

$$\tau \rightarrow \mu^- + \nu_\tau + \bar{\nu}_\mu$$

$$\rightarrow e^- + \nu_\tau + \bar{\nu}_e$$

While ν_e and ν_μ have been detected directly in reactions which absorb these neutrinos, the existence of ν_τ has been inferred indirectly from energy and momentum conservation in the decay of the tauon.

Baryons are strongly interacting particles. They include protons, neutrons and heavier members like the omega-minus, whose final decay products include either a proton or a neutron (Table 15.3).

Table 15.3: Observable Properties of Baryons

Particle	Mass in (MeV/c^2)	Spin Parity	Isospin	Mean-life (s)	Decay modes
p	938.3	$\frac{1}{2}^+$	$\frac{1}{2}$	$> 10^{32}$ years	
n	939.6	$\frac{1}{2}^+$	$\frac{1}{2}$	898	$p e^- \bar{\nu}_e$
Λ	115.6	$\frac{1}{2}^+$	0	2.6×10^{-10}	$p\pi^-$, $n\pi^0$
Σ^+	1189.4	$\frac{1}{2}^+$	1	0.80×10^{-10}	$p\pi^0$, $n\pi^+$
Σ^0	1192.5	$\frac{1}{2}^+$	1	5.8×10^{-20}	$\Lambda\gamma$
Σ^-	1197.3	$\frac{1}{2}^+$	1	1.5×10^{-10}	$n\pi^-$
Ξ^0	1314.9	$\frac{1}{2}^+$	$\frac{1}{2}$	2.9×10^{-10}	$\Lambda\pi^0$
Ξ^-	1321.3	$\frac{1}{2}^+$	$\frac{1}{2}$	1.6×10^{-10}	$\Lambda\pi^-$
Ω^-	1672.5	$\frac{3}{2}^+$	$\frac{1}{2}$	0.82×10^{-10}	$\Lambda\kappa$, $\Xi^0\pi^-$, $\Xi^-\pi^0$
Λ_C^+	2282	$\frac{1}{2}^+$	0	unknown	$pK^-\pi^+$, pK^0, $\Lambda^+ \ldots e^+$, ...

These are the most numerous group of particles. For every baryon, there exists an anti-baryon. As in the case of leptons, these anti-particles have the opposite charge but the same mass and spin as the corresponding particles. Moreover baryons are fermions.

Mesons are particles that take part in strong interactions but are not baryons (Table 15.4). They can be sub-classified in various ways according to their spin and other quantum numbers. Their masses range from a mere 135 MeV, as in the case of pions (π^0), to more than 10,000 MeV as in the case of 'upsilons'. All mesons have integral spin – they are bosons whereas all baryons are spin half particles (fermions). For every meson, there exists an anti-meson. For instance, the anti-particle for π^+ is π^- and the antiparticle for π^0 is π^0 itself. This means that interaction between two π^0 leads to their annihilation.

Table 15.4: Observable Properties of Mesons

Particle	Mass (MeV/c^2)	J^P	I	Mean life (s)	Decay modes
$\pi^\pm$	139.6	0^-	1	2.6×10^{-8}	$\mu^\pm \nu_\mu$
π°	135.0	0^-	1	0.83×10^{-16}	$\gamma\gamma$, $e^+e^-\gamma$
η°	549	0^-	0	10^{-18}	$\gamma\gamma$, $3\pi^\circ$, $\pi^+\pi^-\pi^\circ$, $\pi^+\pi^-\gamma$
$K^\pm$	493.7	0^-	½	1.24×10^{-8}	$\mu^\pm\nu$, $\pi^\pm\pi^\circ$, $\pi^+\pi^-\pi^\pm$, $\pi^\circ\pi^\circ\pi^\pm$ $\pi^\circ\mu^\pm\nu$, $\pi^\circ e^\pm\nu$
$K^\circ, \bar{K}^\circ$	497.7	0	½	Decays as K_1° (50%, K_L° (50%)	
K_S°				0.89×10^{-10}	$\pi^+\pi^-$, $\pi^\circ\pi^\circ$
K_L°				5.2×10^{-8}	$3\pi^\circ$, $\pi^+\pi^-\pi^\circ$, $\pi^\pm\mu^\pm\nu$, $\pi^\pm e^\pm\nu$
$D^\pm$	1869	0^-	½	9.2×10^{-13}	electron or K meson + other particles
$D^\circ, \bar{D}^\circ$	1865	0^-	½	4.4×10^{-13}	
$F^\pm$	1971	0^-	0	1.9×10^{-13}	$\phi\pi$ (ϕ is a short-lived particle decaying into K^+K^-, $K^\circ\bar{K}^\circ$ etc.)
$B^\pm$	5271	0^-	½	1.4×10^{-12}	D meson or electron or μ + other particles
$B^\circ, \bar{B}^\circ$	5274	0^-	½		

Field bosons are particles carrying the electromagnetic and weak forces. The particles in this group are photons carrying the electromagnetic force and the $W^\pm$ and Z° bosons.

You now know that the only absolutely stable particles are the electron, the photon, the neutrinos and (perhaps!) the proton. This means that almost all of the known particles are unstable and decay into several other particles. The decay of a large number of such particles is caused by the force of strong interaction. Decays caused by the strong force are extremely fast. Such particles have lifetimes of only about 10^{-23}s. Such particles are regarded as unstable. If a particle lasts only 10^{-23}s, we cannot even detect it directly since the maximum distance travelled even by light over this time interval is $\sim10^{-15}$m ($10^{-23}s \times 3 \times 10^8 ms^{-1}$), which is roughly the same as the diameter of a proton. Obviously, such particles do not leave any visible track in any detector. However, their existence is usually deduced from circumstantial evidence. Short lived particles are often called resonances. Particles of this short-lived class, some of which are baryons and other mesons, are not listed in Tables 15.2 and 15.3. Some particles may survive for about 10^{-10}s from the instant of production. In high energy physics, a time interval of 10^{-14}s to 10^{-10}s is regarded as a rather large interval and particles with such lifetimes are regarded as stable. Electromagnetic decays that always involve photons are slower. Lastly, particles which cannot find any allowed strong or electromagnetic mode of decay, may find a weak mode to decay. Decays involving weak interactions are, in general, the slowest. The β-decay of a neutron with a mean life of about 15 min is a rather well known example of a lethargic decay mode.

15.6 CONSERVED QUANTITIES

The reactions among particles obey several conservation laws. Some of these conservation laws are already familiar from classical physics, and they are based on general theoretical principles. Some other conservation laws are new, purely empirical rules, without any obvious theoretical justification. Such empirical conservation laws were proposed by particle physicists to account for the (puzzling) absence of diverse hypothetical reactions. In general, we expect that a reaction, which is not forbidden by conservation laws, will occur perhaps at a low rate. The absence of certain reactions, otherwise not forbidden by any known conservation law, suggests the intervention of a new conservation law involving some new conserved quantity. However, when we formulate a new conservation law on such purely empirical ground, we cannot claim to have achieved an explanation of why the reactions in question do not occur; we may have merely codified the observed facts. Such a validation nevertheless has predictive value, since it empowers us to make predictions as to which reactions are possible. Furthermore, the empirical conservation laws provide us with valuable guidance in the construction of theories of interactions.

Some of the conservation laws are absolute, i.e. they are obeyed by all reactions under all circumstances. The absolutely conserved quantities are energy, momentum, angular momentum, electric charge, lepton number and baryon number. The first four of these quantities (energy, momentum, angular momentum, and electric charge) are familiar; their conservation is a consequence of the basic laws of classical or quantum mechanics and electromagnetism. The other quantities are less familiar; we will discuss them one by one.

Lepton Number. In nuclear reactions, the number of electrons is conserved. This notion is recognised by assigning a lepton number $L_e = 1$ for electron and $L_e = -1$ for positron. There are three separate conservation laws for lepton number, corresponding to the three separate varieties of leptons: e, ν, and τ. The conservation law for electron lepton number states that the net number of electron-type leptons remains constant in a reaction. Antileptons are reckoned as making a negative contribution to the net lepton number. In essence, this means that whenever a lepton is created or destroyed in a reaction, a corresponding antilepton must be created or destroyed. For instance, consider the decay of a free neutron

$$n \rightarrow p + e + \bar{\nu}_e$$

You will note that the net electron lepton number before and after the reaction is zero.

The conservation laws for muon-lepton number and tauon-lepton number are formulated similarly. The conservation of lepton number is supported by abundant empirical evidence.

Baryon Number. Analogous to conservation of leptonic number, nucleon number is also conserved in nuclear reaction. This notion is specified by assigning a baryon number to all particles that participate in strong interactions. If we assign B = +1 to baryons B = −1 to the anti-baryon and B = 0 to all non-baryons (mesons,leptons and field particles). The sum of the B's must be the same on both sides of the reaction. For example, production reaction of anti-protons in p-p collisions requires three protons in the final state:

$$p + p \rightarrow p + p + p + \bar{p}$$

so that the net value of B = +2 is conserved on both sides. The conservation law for baryon number states that the net number of baryons remains constant. Like the conservation of lepton number, the conservation of baryon number is an empirical law.

Baryon number conservation forbids decays such as

$$\overline{\Lambda^0} \rightarrow p + \pi^-$$

$$B = -1 \rightarrow +1 + 0$$

The observed stability of the proton can be regarded as a corollary of the conservation laws for energy and for baryon number. Since the proton is the lightest of all baryons, energy conservation requires that the hypothetical decay products be non-baryons, and the decay would violate baryon number conservation. The unified theory of electromagnetic and weak interactions, **the electroweak theory**, formulated by Weinberg, Salam and Glashow, suggested that proton is unstable; it should decay with a half-life of 10^{31} yr. To gather experimental evidence for this aspect of electroweak theory, experiments are being carried out throughout the world. In India, such experiments are on, in collaboration with Japan, in Kolar gold mines under the guidance of Prof. M.G.K. Menon. Since these experiments have failed to detect proton decay so far, the available empirical evidence seems to be consistent with baryon number conservation. (If proton decays are detected, the law of baryon number conservation will have to be regarded an approximate law.)

A complete mathematical definition of other conserved quantities - isospin, strangeness, and parity - falls beyond the scope of this course. We will therefore not discuss these here.

15.7 THE QUARK MODEL

You must now be convinced that the vast variety of new particles with diverse characteristics arose out of our quest for high energies. The question now arises: Can we understand the systematics of such large number of particles in terms of a basic theory or model? The answer to this question is in the affirmative. It is now believed that all hadrons are composite bodies made out of two or three quarks, which cling together through forces

mediated by gluons (which are massless and have spin 1). They are analogous to photons which are carriers of electromagnetic forces among charged particles. The baryons consist of three quarks while the mesons are made up of two quarks.

In the quark model, there are only two sets of elementary particles: leptons and quarks. All quarks have spin 1/2 and carry fractional electric charges, $\left(\frac{2}{3}e \text{ and } -\frac{e}{3}\right)$. The quarks are supposed to exist in six flavours (which have nothing to do with taste). Each quark, like any other particle, has an antiparticle. The charge on each antiparticle is opposite to that on the corresponding quark. These properties are summarised in Table 15.6.

Table 15.5: Properties of Quarks

Name	Symbol	Charge	Spin	Baryon Number	Antiquark
Up	u	$+\frac{2}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\bar{u}$
Down	d	$-\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\bar{d}$
Strange	s	$-\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\bar{s}$
Charmed	c	$\frac{2}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\bar{c}$
Top	t	$\frac{2}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\bar{t}$
Bottom	b	$-\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\bar{b}$

To ensure that Pauli's exclusion principle holds, the concept of colour field was added to flavours. Each flavour of quark occurs in three primary colours: red, green and blue. Each antiquark is assigned the corresponding complementary colour: cyan, magneta and yellow. Thus there are eighteen quarks and an equal number of antiquarks.

Colour is to the strong interaction between quarks as electric charge is to electromagnetic interaction between electrons.

An essential feature of the quark model with colour is that the combination of colours for hadrons always gives white. That is, all observed meson and baryon states are colourless. On mixing three primary colours we get a baryon and mixing a primary colour with its complementary colour yields a meson. (Note that the analogy with ordinary colours is very strong but the particles are not coloured.)

As of now, quarks can be regarded as most elementary particles since all known particles can be constructed out of quarks by joining them in different ways. For example, a proton is made of two u quarks and one d quark. A positive pion is made of one u quark and one d antiquark. You may now ask: Have we reached the ultimate? Perhaps No! Theoretical physicists are proposing new schemes to assemble quarks out of smaller units. Shall we ever reach the end of our quest for the ultimate constituents of matter? Questions never cease!

15.8 SUMMARY

- According to Dirac's relativistic theory, electrons can exist in positive as well as negative energy states.
- In an electrostatic accelerator, changed particles are accelerated by a strong electrostatic field.
- Lawrence designed a cyclotron which consisted of two hollow semi-spherical metallic 'dees'.
- In a cyclotron, changed particles are accelerated by a constant frequency alternating electric field.
- The time a particle takes in traversing one semicircular path is independent of the radius of the path.

- To increase the energy of charged particles in a synchrocyclotron, the frequency of the accelerating voltage is continuously varied as a particle goes away from the centre of the machine.
- In a synchrotron accelerator, the magnetic field strength as well as resonant frequency are varied continuously.
- The devices used to detect particles created in high energy collisions are called detectors.
- Detectors are classified into two groups: counters and chambers depending on whether the device registers the passage of a charged particle or tracks its path.
- GM counter is based on the phenomenon of ionisation by collision.
- In a Wilson Cloud Chamber, a cloud of water drops is facilitated by the presence of ions produced by charged particle induced ionisation.
- All known particles can be classified into four broad groups: the leptons, the baryons, the mesons and field bosons.
- Leptons (e, μ, τ, ν_e, ν_μ and ν_τ) are produced and absored in weak interactions. Baryons (n, p, Λ, Σ^+, Ω^-, Λ_c^+)are strongly interacting particles. Mesons take part in strong interactions but are not baryons. Field bosons are particles carrying the electromagnetic and weak forces.
- All hadrons are composite bodies made of two or three quarks, which cling together through forces mediated by gluons. They are analogous to photons, which are carriers of electromagnetic forces among charged particles.
- In quark model with colour, combination of colours for hadrons always gives white. All observed meson and baryon states are colourless.

Note

Note