Bachelor of Science

(B.Sc.)

WAVE MECHANICS AND NUCLEAR PHYSICS

Semester-III

Author- Dr. Priyanka Roy

SURESH GYAN VIHAR UNIVERSITY
Centre for Distance and Online Education
Mahal, Jagatpura, Jaipur-302017

EDITORIAL BOARD (CDOE, SGVU)

Dr (Prof.) T.K. Jain
Director, CDOE, SGVU

Dr. Dev Brat Gupta
Associate Professor (SILS) & Academic Head, CDOE, SGVU

Ms. Hemlalata Dharendra
Assistant Professor, CDOE, SGVU

Ms. Kapila Bishnoi
Assistant Professor, CDOE, SGVU

Dr. Manish Dwivedi
Associate Professor & Dy, Director, CDOE, SGVU

Mr. Manvendra Narayan Mishra
Assistant Professor (Deptt. of Mathematics) SGVU

Mr. Ashphaq Ahmad
Assistant Professor, CDOE, SGVU

Published by:
S. B. Prakashan Pvt. Ltd.
WZ-6, Lajwanti Garden, New Delhi: 110046
Tel.: (011) 28520627 | Ph.: 9205476295
Email: info@sbprakashan.com | Web.: www.sbprakashan.com

© SGVU

All rights reserved.

No part of this book may be reproduced or copied in any form or by any means (graphic, electronic or mechanical, including photocopying, recording, taping, or information retrieval system) or reproduced on any disc, tape, perforated media or other information storage device, etc., without the written permission of the publishers.

Every effort has been made to avoid errors or omissions in the publication. In spite of this, some errors might have crept in. Any mistake, error or discrepancy noted may be brought to our notice and it shall be taken care of in the next edition. It is notified that neither the publishers nor the author or seller will be responsible for any damage or loss of any kind, in any manner, therefrom.

For binding mistakes, misprints or for missing pages, etc., the publishers' liability is limited to replacement within one month of purchase by similar edition. All expenses in this connection are to be borne by the purchaser.

Designed & Graphic by : S. B. Prakashan Pvt. Ltd.

Printed at :

Introduction to Physics

Physics is the most fundamental of the sciences. Its goal is to learn how the Universe works at the most fundamental level—and to discover the basic laws by which it operates. *Theoretical physics* concentrates on developing the theory and mathematics of these laws, while *applied physics* focuses attention on the application of the principles of physics to practical problems. *Experimental physics* lies at the intersection of physics and engineering; experimental physicists have the theoretical knowledge of theoretical physicists, *and* they know how to build and work with scientific equipment.

Physics is divided into a number of sub-fields, and physicists are trained to have some expertise in all of them. This variety is what makes physics one of the most interesting of the sciences—and it makes people with physics training very versatile in their ability to do work in many different technical fields.

The major fields of physics are:

- *Classical mechanics* is the study the motion of bodies according to Newton's laws of motion, and is the subject of this course.
- *Electricity and magnetism* are two closely related phenomena that are together considered a single field of physics.
- *Quantum mechanics* describes the peculiar motion of very small bodies (atomic sizes and smaller).
- *Optics* is the study of light.
- *Acoustics* is the study of sound.
- *Thermodynamics* and *statistical mechanics* are closely related fields that study the nature of heat.
- *Solid-state physics* is the study of solids—most often crystalline metals.
- *Plasma physics* is the study of plasmas (ionized gases).
- *Nanoscience and Nanotechnology is the new emerging field of physics*
- *Atomic, nuclear, and particle physics* study of the atom, the atomic nucleus, and the particles that make up the atom.

- *Relativity* includes Albert Einstein's theories of special and general relativity. *Special relativity* describes the motion of bodies moving at very high speeds (near the speed of light), while *general relativity* is Einstein's theory of gravity.
- The fields of *cross-disciplinary physics* combine physics with other sciences. These include *astrophysics* (physics of astronomy), *geophysics* (physics of geology), *biophysics* (physics of biology), *chemical physics* (physics of chemistry), and *mathematical physics* (mathematical theories related to physics).

Besides acquiring knowledge of physics for its own sake, the study of physics will give you a broad technical background and set of problem-solving skills that you can apply to wide variety of other fields. Some students of physics go on to study more advanced physics, while others find ways to apply their knowledge of physics to such diverse subjects as mathematics, engineering, biology, medicine, and finance.

In this Book, there are five Blocks. Block I and II deals with Wave mechanics and Fundamental of Quantum mechanics, Block III and IV deals with Fundamentals of Nuclear physics and Radioactivity and, Block V deals with Nuclear reactions.

Syllabus

BPHY-32: WAVE MECHANICS AND NUCLEAR PHYSICS

BLOCK-I

Dual nature of light –De' Broglie's concept of matter waves – De' Broglie wavelength – Wave and group velocity – Relation between wave and group velocity – Davisson and Germer experiment – G.P. Thompson experiments – Heisenberg's Uncertainity Principle.

BLOCK-II

Basic Postulates of wave mechanics – Quantum operators, Linear operator, Hermitian operator, Parity operators – Properties of wave Function – Orthogonal and normalized wave functions – Eigen Values and Eigen Functions – Schrodinger's Equations – Time Independent – Time Dependent – Application – Particle in a box-Rigid rotator.

BLOCK-III

Classification of Nuclei – General Properties of Nucleus – Size, Mass, Density Charge, Angular momentum and Dipole moments – Binding Energy – Packing fraction – Nuclear stability – Semi Empirical Mass formula – Liquid Drop Model – Shell Model.

BLOCK-IV

Radioactivity – Properties of Alpha, Beta and Gamma Rays – Geiger-Nuttal Law – Nuclear Isomerism – Soddy Fajan's displacement law – Radioactive disintegration Law – Half Life, Mean Life periods – Law of Successive disintegration-Linear Accelerator - Cyclotron-Betatron.

BLOCK-V

Types of Nuclear Reaction – Energy balance – Q value – Transmutation by Alpha, Proton, Deutrons and Neutrons – Artificial Radioactivity – Radio Isotopes – Applications – Nuclear Fission – Chain reaction – Nuclear Reactor –Fast Breeder reactor- Nuclear Fusion – Thermo Nuclear Reactions – Carbon-Nitrogen Cycle – Proton-Proton Cycle.

Contents

Block I

Wave Mechanics

STRUCTURE

OVERVIEW

We will begin this Unit by describing Dual nature of matter, and in this context, we will derive the expression for de Broglie wavelength. We will then touch wave and group velocity and its experimental verification. In addition, study about Heisenberg's Uncertainty principle

LEARNING OBJECTIVES

After completing this Unit, you should be able to:

- differentiate between particle and wave nature of matter;
- Derive the expression for de broglie's wavelength
- Derive the relation between wave and group velocity

1.1 Introduction

Nucleons in a nucleus do not behave like classical particles (colliding like billiard balls); instead, the wave behavior of the nucleon determines the properties of the nucleus. Therefore, we need quantum mechanics, which is a mathematical technique that enables us to calculate the wave behavior of a material particle.

The quantum behavior of light and the photoelectric effect that had been analyzed, showed that the light or electromagnetic waves should also be considered as if its energy were delivered not smoothly and continuously as wave but instead in concentrated bundles or quantum effect (particle of light).

The analogy between matter and light (wave) was made in 1924 by de Broglie; he postulated that associated with a particle moving with momentum p is a wave of wavelength $\lambda = h/p$ (Eq.1.22). Experimental confirmation soon followed through the diffraction of electrons (particles) like waves with de Broglie wavelength. The de Broglie theory was successful in those instances, but it is incomplete and unsatisfying in describing the particle by classical physics, since the classical particle has a definite position in space and unique momentum p.

The solution of this problem comes from quantum physics, the size of a quantum particle varies with the experiment performed. Thus, an electron or particle may have a certain size in one experiment and a very different size in another. Only through this coupling of the observing system and the observed object can we define observation in quantum physics. The particle then is localized within some region of space of dimension Δx. It is likely to be found in that region and unlikely to be found elsewhere. We improve our knowledge of Δx at the expense of our knowledge of momentum px, the very act of confining the particle to Δx destroys the precision of our knowledge of px. It is not our goal to take up the study of quantum mechanics as a topic by itself. On the other hand, we have no reasons to avoid using quantum mechanics whenever it is the proper way to understand nuclear concepts and radiation interactions.

In the following sections, the concepts and terminologies in quantum mechanics (Schrödinger equation) will be given, since they are such integral parts of nuclear concepts, nuclear structure and the interaction of radiation with matter that some knowledge of quantum mechanics is essential to having full command of the language of nuclear physics.

1.2 Dual nature of light

The true nature of light is difficult to assess. Experiments showed that light exhibited wavelike properties of diffraction and interference. On the other hand, photoelectric effect indicates that light has the aspects of a particle photon, with both energy and momentum. Thus light exhibits a wave-particle duality.

The wave-particle duality was extended to particles as matter waves by Louis de Broglie. His theoretical study on the nature of particles and waves led to the invention of a new mechanics of particles called quantum mechanics.

Matter waves

The radiant energy has dual aspects of particle and wave, hence a natural question arises, if radiation has a dual nature, why not the matter. In 1924, a French Physicist Louis de Broglie put forward the bold hypothesis that moving particles should possess wave like properties under suitable conditions. He reasoned this idea, from the fact, that nature is symmetrical and hence the basic physical entities –

matter and energy should have symmetrical characters. If radiation shows dual aspects, so should matter.

1.3 De *Broglie wavelength*

Consider a photon of energy 'hν' moving with the velocity of light 'c'. According to the Plancks theory of radiation, the energy of a photon is given by,

$$E = h\nu = \frac{hc}{\lambda}$$

(wave aspect) $\left[\nu = \frac{c}{\lambda}\right]$ ---------- (1)

where, h → Plank's constant

ν → frequency of radiation

If the photon is considered as a particle of mass 'm' then according to Einstein's mass – energy relation, the energy of the photon is given by ,

$E = mc^2$ (particle aspect) --------------------- (2)

The energy of the photon in the two cases is the same, hence from equations (1) and (2) we get,

$$\frac{hc}{\lambda} = mc^2$$

$$\lambda = \frac{hc}{mc^2}$$

$$\lambda = \frac{h}{mc}$$

If a particle of mass ‘m’ travels with a velocity ‘v’ then its momentum is ‘mv’. The wavelength corresponding to this particle is given by,

$$\boxed{\lambda = \frac{h}{mv} = \frac{h}{p}} \text{ -------------- (3)}$$

where, λ→ wavelength of matter waves.

Equation (3) is called as deBroglie wave equation.

deBroglie wavelength in terms of kinetic energy

The kinetic energy of a moving particle is,

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

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

$$m^2v^2 = 2mE$$

$$p = mv$$

$$\therefore\ p^2 = 2mE$$

But $$p = \sqrt{2mE}$$

deBroglie wavelength $$\boxed{\lambda = \frac{h}{p} = \frac{h}{\sqrt{2mE}}}$$

The above equation gives the deBroglie wavelength in terms of kinetic energy.

deBroglie wavelength associated with electrons

Consider an electron of mass 'm' charge 'e', which is accelerated by the potential 'V' from rest.

$\therefore$ Energy gained by the electron = eV ----------------- (1)

This energy gained is converted into kinetic energy of the electron.

$\therefore$ Kinetic energy of the electron = $\frac{1}{2}mv^2$ ------------------------- (2)

Equating (1) and (2)

$$eV = \frac{1}{2}mv^2$$

$$v^2 = \frac{2eV}{m}$$

$$v = \sqrt{\frac{2eV}{m}}$$

We know deBroglie wavelength $\lambda = \frac{h}{mv}$

Substituting the value of 'v' from the above equation

$$\lambda = \frac{h}{m\sqrt{\frac{2eV}{m}}}$$

$$\lambda = \frac{h}{\sqrt{2Vem}}$$

Substituting the values,

$$h = 6.6256 \times 10^{-34}\, Js$$

$$e = 1.6 \times 10^{-19}\, C$$

$$m = 9.1 \times 10^{-31}\, kg$$

$$\lambda = \frac{6.6256 \times 10^{-34}}{\sqrt{2 \times 1.6 \times 10^{-19} \times 9.1 \times 10^{-31} \times V}} = \frac{12.24}{\sqrt{V}}\, A^\circ$$

$$\boxed{\lambda = \frac{12.24}{\sqrt{V}} A\circ}$$

The above equation gives the relation of deBroglie wavelength associated with an electron.

1.4 Wave and group velocity

Phase and group velocity are two important and related concepts in wave mechanics. They arise in quantum mechanics in the time development of the state function for the continuous case, i.e. wave packets.

Harmonic Waves and Phase Velocity

A one-dimensional harmonic wave (Figure 1) is described by the equation,

$$u(x,t) = A_0 \cdot \sin(\omega t - k x + \varphi)$$

where A_0 is the wave amplitude, w is the circular frequency; k is the wave number; and j is an initial, constant phase. The argument for the sine function, $q\ (x,\ t) = wt - kx + j$ is called the phase. Sometimes the wave number is referred to as the spatial frequency or propagation constant.

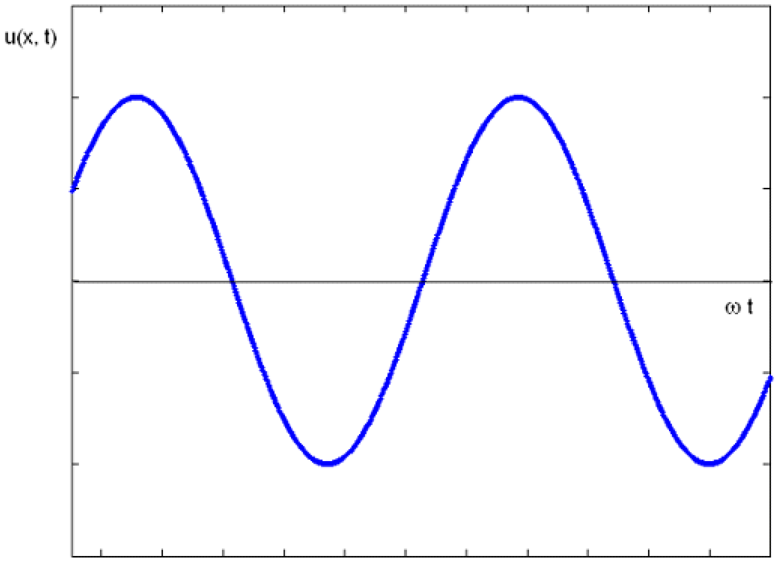


Figure 1: Harmonic Wave

This is a monochromatic wave (one frequency). There are no strictly monochromatic waves in nature. For example, the generating source of the wave may move slightly, introducing spurious frequencies.

In general, these waves propagate without warping. That is, the phase *q* (*x*, *t*) is a constant:

$$\frac{d\theta}{dt} = \omega - k\frac{dx}{dt} = 0$$

$$\frac{dx}{dt} = \frac{\omega}{k} = v_{phase}$$

v_{phase} is the phase velocity for a wave.

From the point of view of sending information, these waves are not useful. They are the same throughout time and space. Something must therefore be modulated, such as frequency or amplitude, in order to convey information. The resulting wave may be a perturbation that acts over a short distance, i.e. a wave packet. This wave packet can be considered to be a superposition of a number of harmonic waves, in other words a Fourier series or integral.

Group Velocity

In order to convey information, something more than a simple harmonic wave is needed. However, the superposition of many such waves of varying frequencies can result in an "envelope" wave and a carrier wave within the envelope. The envelope can transmit data. A simple example is the superposition of two harmonic waves with frequencies that are very close ($w_1 \sim w_2$) and of the same amplitude. The equations for the motion are,

$$\begin{aligned} u(x,t) &= A_0 \cdot \cos(\omega_1 t - k_1 x) + A_0 \cdot \cos(\omega_2 t - k_2 x) \\ &= 2A_0 \cdot \cos(\tfrac{\omega_2-\omega_1}{2} t - \tfrac{k_2-k_1}{2} x) \times \cos(\tfrac{\omega_2+\omega_1}{2} t - \tfrac{k_2+k_1}{2} x) \\ &= u_1 \times u_2 \end{aligned}$$

The plot of such a wave is shown in Figure 2.

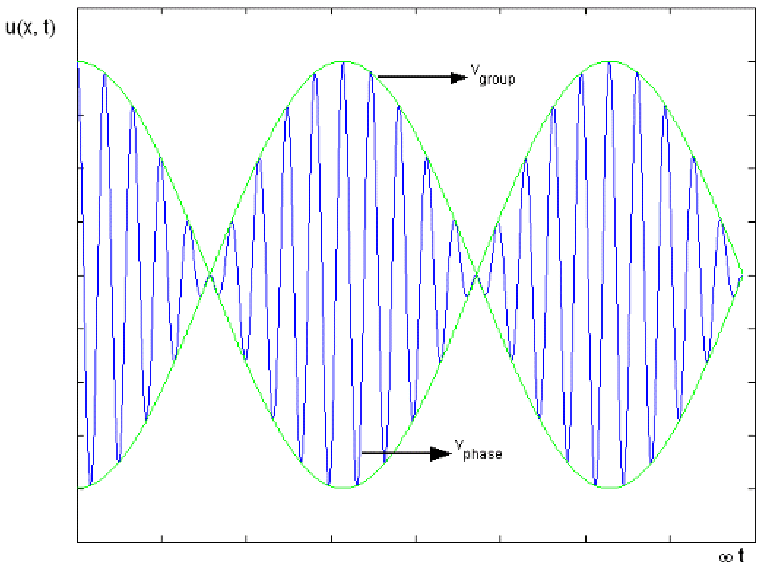


Figure 2: Group Velocity

The envelope (the green line) is given by u_1 and travels at the group velocity. The carrier wave (the blue line) travels at the phase velocity and is given by u_2. The wave packet moves at the group velocity. It is the envelope which carries information. Group velocity and phase velocity are not necessarily the same. Group velocity is given by,

$$v_{group} = \frac{\omega_2 - \omega_1}{k_2 - k_1} = \frac{\Delta\omega}{\Delta k} \Rightarrow \frac{d\omega}{dk}, \text{ for the continuous case}$$

Phase and group velocity are related through Rayleigh's formula,

$$v_{group} = v_{phase}\left(1 - \frac{\omega}{v_{phase}}\frac{dv_{phase}}{d\omega}\right)^{-1}$$

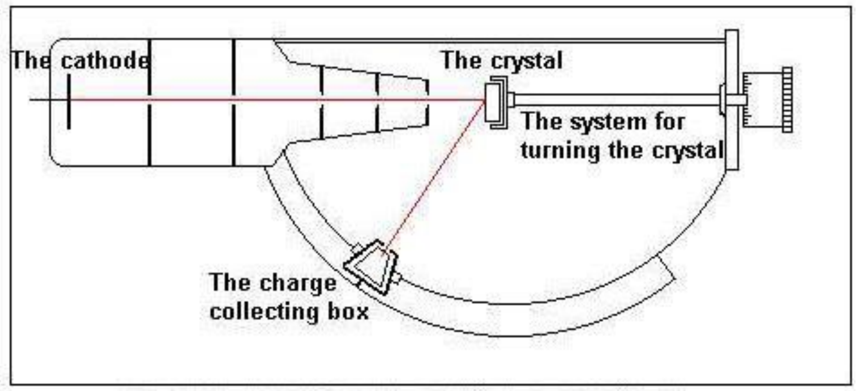
If the derivative term is zero, group velocity equals phase velocity. In this case, there is no dispersion. Dispersion is when the distinct phase velocities of the components of the envelope cause the wave packet to "spread out" over time. The components of the wave packet (or envelope) move apart to the degree where they no longer combine to complete the envelope.

EXPERIMENTAL STUDY OF MATTER WAVES:-

Several years after Debroglie's work, Davision and Germer and G.P Thomson independently demonstrated that streams of electrons are diffracted when they are scattered from crystals.

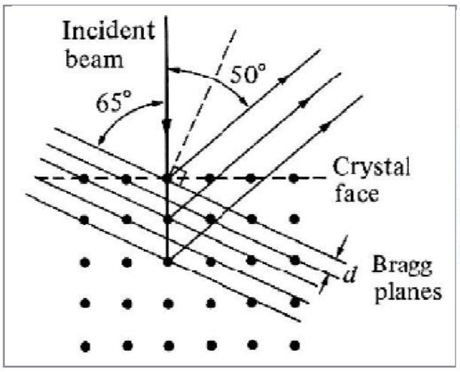
1.5 Davisson and Germer experiment

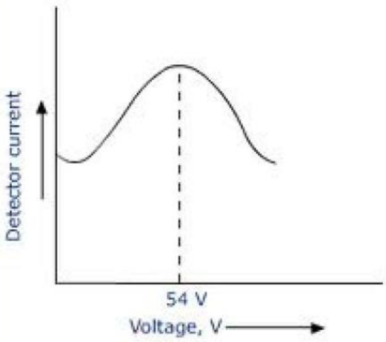
Principle:- Based on the concept of wave nature of matter fast moving electrons behave like waves. Hence accelerated electron beam can be used for diffraction studies in crystals.



The diagram for the Davisson-Germer experiment.

An electron gun which consists of a tungsten filament F heated by a low battery B1, produces electrons. These electrons are accelerated to a desired velocity by applying suitable potential from high tension source B2. The accelerated electrons are collimated into a fine beam by allowing them to pass through a system of pin holes provided in the cylinder ‘C’. The fast moving electrons is made to strike the target (Ni crystal) capable of rotating about an axis perpendicular to the plane of the diagram i.e. incident ray direction. The electrons are now scattered in all directions by the atomic planes of the crystal. The intensity of the electron beam scattered in direction can be measured by the electron collector which can be rotated about the same axis as the target. The collector is connected to a sensitive Galvanometer whose deflection is proportional to the intensity of the electron beam entering the collector. The instrument is kept in an evacuated chamber.





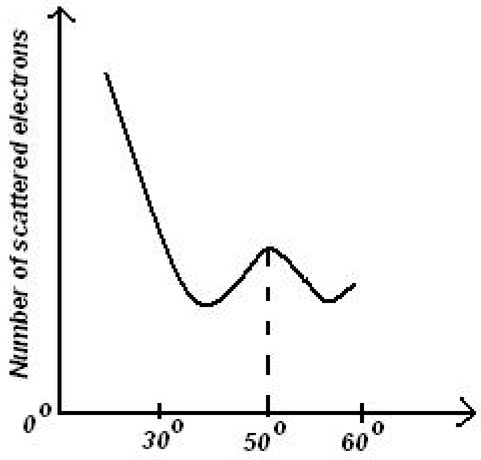


Figure (2): Angle of Direction of Scattering in Degrees

In an investigation, the electron beam accelerated by 54V was directed to strike the given Nickel crystal are a sharp maximum electron distribution occurred at a angle of 500 with the incident beam. The incident and diffracted beam in the experiment make an angle of 650with the Bragg's planes. The spacing of planes in this Bragg's planes by X-Ray diffraction is 0.91 nm.

New according to Bragg's law $2d \sin\theta = n\lambda$

$2\times0.91\times10-10 \times\sin 650 = \lambda \times1 \ (n = 1)$

$\because \lambda = 1.64\times10-1 \times10-9m$

$\lambda = 0.164nm$

For 54V electron the de Broglie wavelength associated with the electron is given by

$$\lambda = \frac{12.25}{\sqrt{54}} A^0$$

$\lambda = 0.164nm$

This is excellent agreement with experimental value. The Davission-Germer experiment provides a verification of de-Broglie hypothesis of the nature of moving particle.

1.6 G.P. Thompson experiments

G.P Thomson performed experiments in which electrons are accelerated from 10,000 to 50,000 volts. In these experiments the generation of electrons are considered analogous to X-Ray obtained by diffraction pattern. The diffraction pattern is obtained by only when wave is associated with particle. Hence Thomson explains the concept of matter waves. The electrons are emitted from the filament and only some accelerated electrons are passing through cathode ‘C’. Next these electrons are passed through two slits S1 and S2 and a thin pencil beam of electrons is obtained. This electrons beam allowed to fall on a thin foil ‘G’ of gold or Aluminium of order10−6 *cm*. The photograph of electron beam from the foil is recorded on the photographic plate ‘P’. Hence a pattern consists of concentric rings. The complete apparatus is kept in high vacuum chamber so that the electrons may not lose their energy y colliding with molecules of air or any inside the tube.

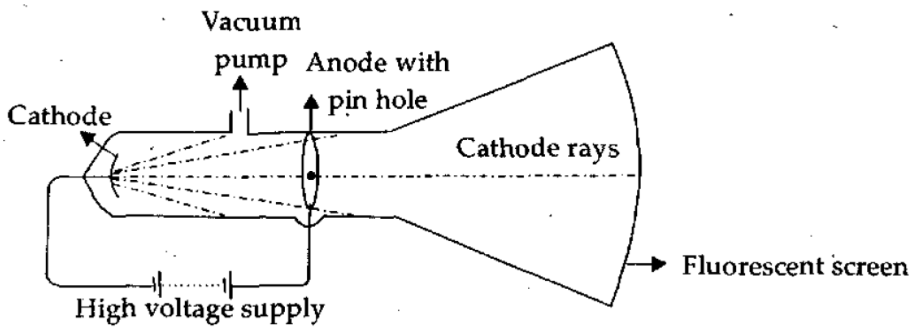
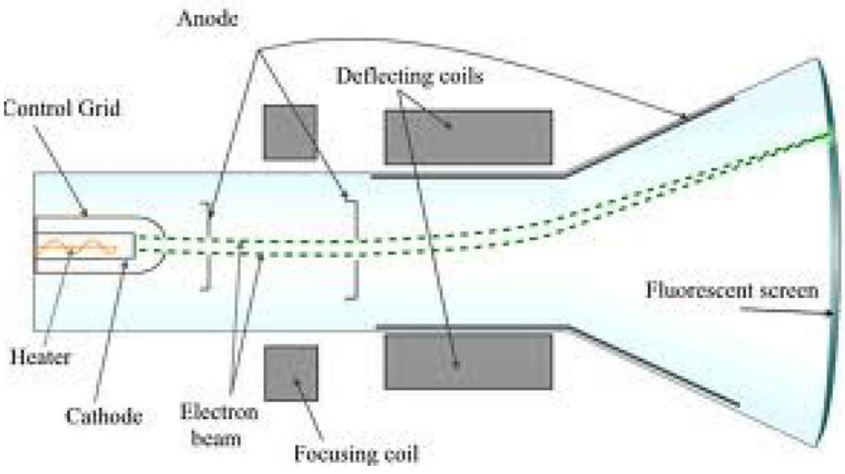


Fig. 1. The discharge tube and the cathode rays



To conclude that, this pattern is due to the electrons and not due the X-Rays. The cathode rays inside the tube are affected by the magnetic fields. The beam shifting considerably along the field is observed. Hence we can conclude that the pattern obtained is due to electrons only since x- Rays are not affected by electric and magnetic fields.

1.7 Heisenberg's Uncertainity Principle

The uncertainty principle is one of the most significant laws of physics, discovered by Werner Heisenberg in 1927 (We cannot know the future because we cannot know the present). Accordingly, if we try to make a simultaneous determination of x and px, our result will show that each is uncertain by respective amounts Δx and Δpx, which are related by the Heisenberg uncertainty relationship.

$\Delta x\ \Delta px \geq \hbar/2$

where $\hbar = h/2\pi = 1.0545\times10^{-34}$ j.sec

This equation states that if we arrange matters so that Δx is small (corresponding to a narrow wave group), Δp will be large. If we reduce Δp in some way, a broad wave group is inevitable and Δx will be large.

The same arguments hold for other form of the uncertainty principle, the other measurements concern energy and time. We might wish to measure the energy emitted during the time interval Δt in an atomic or nuclear process. Since energy $E = h\nu = hc/\lambda$ where ν is the frequency , and $\Delta\nu \geq 1/\Delta t$ also $\Delta E = h\Delta\nu$ then:

$\Delta E\ \Delta t \geq \hbar/2$

If a system live for a time Δt, we cannot determine its energy except to within uncertainty ΔE. A third uncertainty relationship involves the angular momentum Lz, and the azimuthally angle φ, Fig.

$\Delta L_z\ \Delta\varphi \geq \hbar/2$

That is, if we have Lz exactly, we know nothing at all about φ.

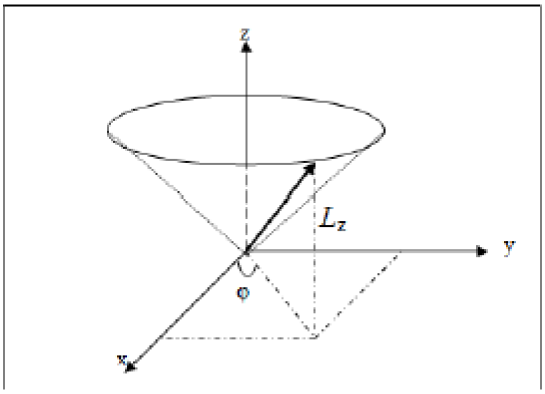


Figure. Bounded particle uncertainty relationships.

de Broglie Wave Descriptions

It is instructive to begin with the classical wave equation that incorporates the concept of de Broglie waves and to review some basic properties of waves and the concept of wave-particle duality.

The wave nature of a moving particle leads to some remarkable consequences when the particle is restricted to a certain region of space instead of being able to move freely, as the electron bounded to the atom or a nucleon in the nucleus, and when it interacts with the nucleus or other particle.

In classical mechanics, the wave equation for a one-dimensional periodic disturbance ξ(x, t) is:

$$\frac{\partial^2 \xi(x,t)}{\partial t^2} = c^2 \frac{\partial^2 \xi(x,t)}{\partial x^2}$$

Solutions of the above equation may be of many kinds, reflecting the variety of waves that can occur, such as a single traveling pulse, a standing wave, a group (train) of waves or of superposed waves, etc…, the general solution is of the form:

$$\xi(x,t) = \xi_0 e^{i(kx-\omega t)}$$

where $\omega = 2\pi\nu$ is the angular frequency, υ the linear frequency, and k is the wavenumber related to the wavelength λ by $k = 2\pi / \lambda$. k and ω must satisfy the relation:

$\omega = ck$

Therefore, our solution has the form of a traveling wave with phase velocity equal to c, which we will denote it by υph. In general, the relation between frequency and wavenumber is called the dispersion relation. We will see that different kinds of particles can be represented as waves, which are characterized by different dispersion relations.

The solution is called a plane wave. In three dimensions, a plane wave is of the form eik.r. It is a wave in space that can be visualized as a series of planes perpendicular to the wavevector k at any spatial point on a given plane, the phase of the wave is the same. That is to say, the perpendicular planes are planes of constant phase. When we include the time variation $e^{(-i\omega t)}$, then $e^{[i(k.r-\omega t)]}$ becomes a traveling plane wave, meaning that the planes of constant phase are now moving in the direction along k at a speed of ω / k , which is the phase velocity of the wave υ_{ph}.

The wave equation also admits solutions of the form:

$\xi\,(x,t) = a o \sin kx \cos\omega t$

These are standing wave solutions. One can tell a standing wave from a traveling wave by the behavior of the nodes, the spatial positions where the wave function is zero. For a standing wave, the nodes do not move, or change with time, whereas for a traveling wave, the nodes are:

$x_n = (n\pi + \omega t)/ k$

Clearly the nodes are positions moving in the (+x) direction with the velocity $dx/dt = \omega/k$.

The choice between traveling and standing wave solutions as we will see depends on the physical solution of interest (which kind of problem one is solving). For the calculation of energy levels of a nucleus, the bound state problem, we will be concerned with standing wave solutions. In contrast, for the discussion of scattering problem, it will be more appropriate to consider traveling wave solutions.

Our interest in the properties of waves lies in the fact that the quantum mechanical description of a nucleus is based on the wave representation of the nucleus. It was first postulated by de Broglie (1924) that one can associate a particle of momentum p and total energy E with a group of waves (wave packet) which are characterized by a wavelength λ and a frequency ν , with the relation, given in Chapter One:

$$\lambda = h / p = h / \gamma m \upsilon, \text{ where } \gamma = (1 - \upsilon^2/c^2)^{-1/2}$$

$$\nu = E / h = \gamma m c^2 / h$$

Here, υ is the particle velocity. Moreover, the motion of the particle is governed by the propagation of the wave packet. This statement is the essence of particle-wave duality, a concept that we adopt throughout our study of nuclear physics. It is important to distinguish between a single wave and a group of waves. This distinction is seen most simply by considering a group of two waves of slightly different wavelengths and frequencies. Suppose we take as the wave packet:

$$\Psi(x,t) = \Psi 1(x,t) + \Psi 2(x,t)$$

With

$$\Psi 1(x,t) = \sin(kx - \omega t)$$

$$\Psi 2(x,t) = \sin[(k + dk)x - (\omega + d\omega)t]$$

Using the identity:

$$\sin A + \sin B = 2\cos[(A - B)/2]\sin[(A + B)/2]$$

We can rewrite $\Psi(x,t)$ as

$$\Psi(x,t) = 2\cos[(dkx - d\omega t)/2]\sin\{[(2k + dk)x - (2\omega + d\omega)]t/2\}$$

$$\approx 2\cos[(dkx - d\omega t)/2]\sin(kx - \omega t)$$

In this approximation, terms of higher order in dk/k or dω/ω are dropped. Eq.5.14 shows two oscillations, one is the wave packet oscillating in space with a period of 2π/k, while its amplitude oscillates with a period of 2π/dk (Fig. 5.2). Notice that the latter oscillation has its own propagation velocity, dω/dk. This velocity is in fact the speed with which the associated particle is moving. Thus, we identify

$$\upsilon_g = d\omega/dk$$

as the group velocity. The group velocity should not be confused with the propagation velocity of the wave packet, called the phase velocity (v_{ph}), given by:

$$\upsilon_{ph} = \nu\lambda = E\,/\,p = c\,[1 + (m_o c\,/\,p)^2]^{1/2}$$

Here, mo is the rest mass of the particle and c the speed of light. We see the wave packet moves with a velocity greater than c, whereas the associated particle speed is necessarily less than c. This means that the phase velocity is greater than or equal to the group velocity.

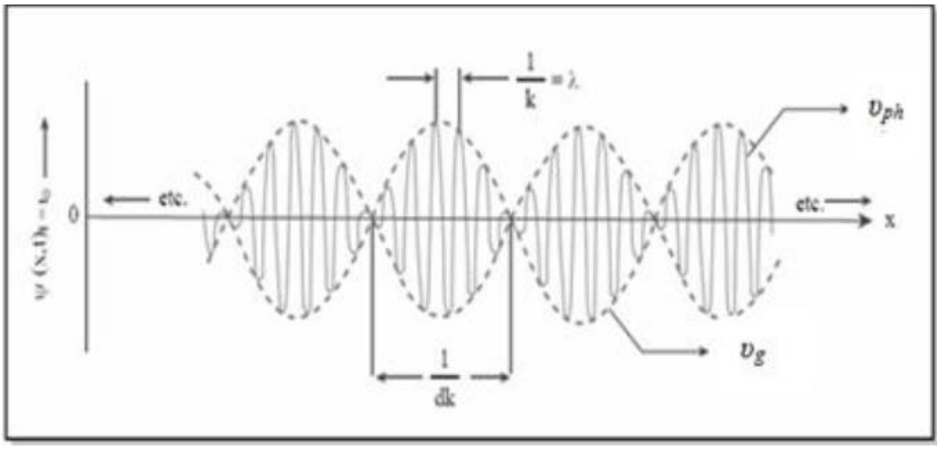


Figure . Spatial variation of a sum of two waves of slightly different frequencies and wave-numbers showing the wave packet moves with velocity υph that is distinct from the propagation (group) velocity υg of the amplitude.

For a non-relativistic particle of mass m moving with momentum p, the associated wavevector k and its kinetic energy is:

$p = \hbar k$ and

$T = p^2/2m = \hbar^2k^2/2m$

This is the “particle view”. The corresponding “wave view” would have the momentum magnitude $p = h/\lambda$, with $\lambda = 2\pi/k$, and energy (usually denoted as E rather than T) as $h\nu = \hbar\omega$. The wavevector, or its magnitude, the wavenumber k, is a useful variable for the discussion of particle scattering since in a beam of such particles, the only energies are kinetic, and both momentum and energy can be specified by giving k.

For electromagnetic waves, the associated particle (the photon), has momentum p, which is also given by ħk, but its energy is:

$E = c\hbar k = cp$

Comparing these two cases, we see that the dispersion relation ω and group velocity νg (Eq.5.15) are:

$$\omega = \hbar k^2/2m, \qquad \upsilon_g = \hbar k/m = p/m$$

for a non-relativistic particle, and $\omega = ck$, $\upsilon_g = c$ for a photon.

For the calculation of energy levels of a nucleus, the bound state problem, such system is like a standing wave in a string stretched between the box's walls, the wave variable (wave function Ψ for the moving particle) must be zero at the walls, i.e., Ψ(0) = 0 and Ψ(a) = 0 for a box of width a.

The possible de Broglie wavelengths of the particle trapped in the box of width (a) is determined by the largest wavelength of λ = 2a and the next λ = a then λ = 2a/3 … for n = 1, 2, 3…. respectively.

So that λ = 2a/n n= 1, 2, 3

Accordingly, the momentum of the particle p and its kinetic energy T will be limited, the particle has no potential energy in this model, the only energies the particle can have are:

$$E_n = \frac{h^2}{2m\lambda^2} = \frac{h^2 n^2}{8ma^2} = \frac{\hbar^2 \pi^2}{2ma^2} n^2$$

N = 1, 2, 3, ….

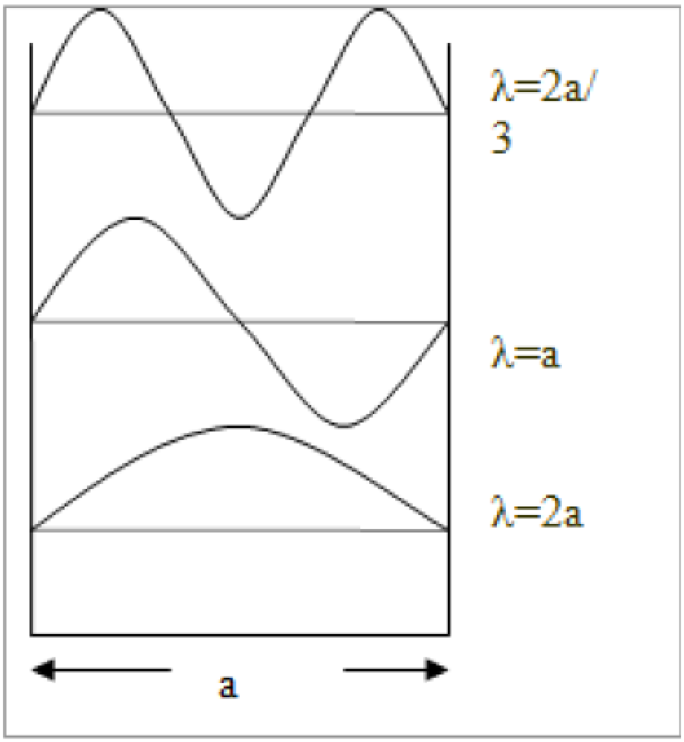


Figure de Broglie wave and principle quantum number)*n*(.

Each permitted energy is called *Energy Level*, and the integer)*n*(that specifies energy levels *En* is called its *quantum number*.

We can draw three general conclusions from the above equation:

1. A bounded particle cannot have an arbitrary energy as a free particle can. These discreet energies depend on the mass of the particle and on the detailed forces acting on the particle to bond it)trapped(.

2. A bounded particle cannot have zero energy, since $\lambda = h/m\upsilon$, a speed $\upsilon = 0$, means an infinite λ, but there is no way to reconcile an infinite wavelength with trapped particle.

3. Because Plank‟s constant is so small, *h* = 6.63×10-34 j.sec, then quantization of energy is conspicuous only when the mass)*m*(and the width)*a*(are also small.

It is convenient to mention that this quantum number)*n*(is precisely the same formula of the principle quantum number determined by the solution of the Schrödinger equation applied to a potential well, as we will see in the next section, where the quantization of the bounded nucleon in the nucleus is therefore described by the principle quantum number)*n*(.

LEARNING ACTIVITY

1. **Discuss in detail about dual nature of matter**
2. **Explain the Davisson and Germer experiment with diagram**
3. **Derive an relation between wave and group velocity**
4. **State and explain Heisenberg's Uncertainity Principle**

Note:

a) Write your answer in the space given below.

b) Check the answer with your academic counsellor.

SUMMARY

In this Unit, we said that dual nature of matter and derive an expression for de brogli wavelength and detail study about Wave and group velocity and also study the experimental setup for Davisson and Germer experiment and GP Thomson method. We also study about Heisenberg’s Uncertainty principle and verification

Block II

Quantum Mechanics

STRUCTURE

Overview

Learning Objectives

2.1 Introduction
2.2 Basic Postulates of wave mechanics –
2.3 Quantum operators -Linear operator, Hermitian operator, Parity operators
2.4 Properties of wave Function
2.5 Orthogonal and normalized wave functions
2.6 Schrodinger's Equations
2.7 Application – Particle in a box.

Summary

OVERVIEW

In this chapter on `Wave mechanics, we shall study the basic of the wave mechanics and quantum operator. We shall study about properties of wave function and schordinger's equations and its applications.

LEARNING OBJECTIVES

After completing this Unit, you should be able to:

- Understand the mathematics needed to understand quantum mechanics.
- Know the inadequacies of classical mechanics, and what was needed to getaround the related difficulties.
- Derive the equation for quantum-mechanical motion.
- Statistically interpret the wavefunction associated with a particle.
- Find the possible states in which a quantum-mechanical particle could be found.
- Get the probability that the quantum-mechanical particle is in any particular state.
- Understand the quantum-mechanical harmonic oscillator.

2.1 Introduction

Quantum Mechanics

Quantum Mechanics began with the work of Planck, Dirac, de Broglie, Heisenberg, Bohr, Schrödinger, and Einstein from 1900 to 1930. It became a necessity as classical mechanics, which had up to that time answered all questions concerning the motion of a body, failed to explain some physical phenomena. It became clear that matter behaves as a particle or as a wave. In other words, a particle behaves as a wave; a wave also behaves like a particle. For example, light waves behave like particles, called photons. On the other hand, with an appropriate „slit,‟ electrons can be diffracted just like any other wave. It became necessary, therefore, to develop a wave equation that gives the dynamics of a particle – the Schroedinger equation. But then, if matter now behaves like a wave, it becomes necessary to give a statistical probabilistic interpretation to the possibility of finding a particle at any particular point, or within a given range of space available for the particle. In other words, it no longer makes sense to say with certainty that a particle is at a particular position, rather, it spreads out over a given range of position: the electron in the hydrogen atom is indeed smeared over the entire sphere outside the nucleus. This is encapsulated in the Heisenberg Uncertainty Principle: It is impossible to measure the position and the linear momentum of a body with infinite accuracy, simultaneously. If the atom is not polarised, it is as if half the electron resides in each hemisphere.

You wonder why we do not realise this in day-to-day experience. This is because the uncertainty in your position is so small, because uncertainty is related to Planck‟s constant, which is of the order of 10^{-34} Js. At the atomic scale, this number is no longer „small.‟ As such, quantum mechanics becomes inevitable at the atomic and subatomic range of distances and masses. Another consequence of the wave nature of matter is that physical quantities can no longer take a continuou range of values. You would recall, for instance, that waves on a string within rigid supports, as well as sound waves in a pipe opened on one or either end can only take a set of frequencies. It then becomes natural for the electron in the hydrogen atom can only occupy a certain set of „allowed energies.‟ That was what Bohr tried to explain with some ad-hoc assumptions of allowed orbits.

With what you have seen in this introduction, it is obvious that quantum mechanics is an interesting, area of physics, that finds application in all life, particularly at the atomic level and below. Quantum mechanics is therefore the present and the future of physics. Solid state devices such as transistors, which are the building blocks of electronics and computers; any material, since all matter is composed of atoms, the ordinary light you deal with everyday, Lasers, elementary particle physics, are just a few of the applications of quantum mechanics.

Basic Postulates of wave mechanics

1: Vector Spaces

1.0 Introduction
2.0 Objectives
- At the end of this Unit, you should be able to:
- Define Vector Spaces
- Give examples of Vector Spaces
- Define linear independence
- Understand Inner or Scalar product of two vectors
- Normalise any given vector

3.0 Main Content
  3.1 Vector Spaces
  3.2 Linear Independence
  3.3 Basis Vector
  3.4 Inner or Scalar Product
  3.5 Norm of a Vector
4.0 Conclusion
5.0 Summary
6.0 Learning Activity
7.0 References/Further Readings

1.0 Introduction

In order to grasp Quantum Mechanics, you need to be conversant with Vector Spaces and other basic ideas of mathematics. The vector space of twice integrable functions enable you to define a set of functions that would form a set of ‘coordinates’ for the vector-like functions, such that as we expand a given vector in 2-dimensional Euclidean space as a linear combination like $a\mathbf{i}+b\mathbf{j}$, we could also expand a given ‘quantum-mechanical function’ as a linear combination of the set of functions. This Unit will teach you how to go about setting up the set of functions, that we shall call an orthonormal set. You shall learn to expand a given function in terms of the orthonormal set, and get to know how to recover the coefficient of expansion of a particular function.

2.0 Objectives

At the end of this Unit, you should be able to:

- Define the term Vector Spaces
- Give examples of Vector Spaces
- Define linear independence
- Understand Inner or Scalar product of two vectors
- Normalise any given vector

3.0 Main Content

3.1 Vector Spaces

No doubt, you are quite familiar with the concept of a vector. With vector spaces, we are generalising this basic idea. In other words, we shall have ‘vectors’ that are no longer just ordinary geometrical vectors, but vectors of a different kind, but all having similar properties. We shall come across matrices that functions that you could give the same treatment as you did geometrical vectors.

Definition

Given a set $\{v_1, v_2, ..., v_n\} = S$. If

(i) $v_i + v_j \in S \quad \forall i, j = 1, 2, ..., n$ 1.1

(ii) $\alpha v_i \in S \ \forall \ i, = 1, 2, .., n;$ 1.2

$\alpha \in K$, where K is a field, e.g., the real number line (R) or the complex plane (C), then, S is called a **vector space** or **linear space**. The vector space is a **real vector space** if $K \equiv R$, and a **complex vector space** if $K \equiv C$.

Condition (i) says that if you add any two vectors of the vector space you will get a member of the space. Condition (ii) shows that a linear multiplication of any two vectors produces a vector also in the vector space. That certainly makes sense, doesn’t it? You don’t want a situation where you add two vectors in your space and get a vector not in the space. Moreover, you avoid a situation where multiplying by a constant takes your vector away from the space. We are now safe to carry out either operation without worrying whether the vector we get is a ‘sensible’ vector, because we are sure it is.

A way to remember these two conditions is: Additivity [condition (i)] + homogeneity [condition (ii)] = linearity.

We now give you some examples of vectors spaces:

Example 1: The set of Cartesian vectors in 3-dimensions, V_3

$\mathbf{a}, \mathbf{b} \in V_3, \ \lambda \in R$.

(i) $\mathbf{a} + \mathbf{b} \in V_3$ 1.3

(ii) $\lambda \mathbf{a} \in V_3$ 1.4

Of course, you know that when two 3-dimensional vectors are added, you also get a 3-dimensional vector. Moreover, multiplying a 3-dimensional vector by a real constant will give you a 3-dimensional vector.

Example 2: $m \times n$ matrices under addition and scalar multiplication, M_{mn}

$A, B \in M_{mn}, \ \lambda \in R$ or C

(i) $A + B \in M_{mn}$ 1.5

(ii) $\lambda A_{mn} \in M_{mn}$ 1.6

You would recall that the addition of two $m \times n$ matrices gives you an $m \times n$ matrix. Similarly, multiplying an $m \times n$ matrix by a real number or a complex number yields an $m \times n$ matrix.

Example 3: A set of functions of x, $\{f(x), g(x), ...\} = F$

$f(x),\ g(x) \in F$, $\lambda \in R$ or C

(i) $f(x) + g(x) \in F$ 1.7

(ii) $\lambda f(x) \in F$ 1.8

Adding two functions of *x* will result in a function of *x*. It just has to be. Also, multiplying a function of *x* by a real number, you get a function of *x*.

3.2 Linear Independence

Given a set $\{\mathbf{v}_i\}_{i=1}^{n}$. If we can write

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n = 0 \qquad 1.9$$

and this implies the constants $a_1 = a_2 = \cdots = a_n = 0$, then we say $\{\mathbf{v}_i\}_{i=1}^{n}$ is a linearly independent set.

If even just one of them is non-zero, then the set is linearly dependent. Think of it: a 3-dimensional Cartesian vector will be a zero vector, **0**, notice the boldface type (not zero scalar), if and only if the three components are independently zero. Thus, for instance, **i**, **j**, and **k**, the traditional unit vectors in 3-dimensional Cartesian space are linearly independent. Mathematically, this means that $\alpha\,\mathbf{i} + \beta\,\mathbf{j} + \gamma\,\mathbf{k} = \mathbf{0}$ if and only if $\alpha = \beta = \gamma = 0$.

Some other examples are in order here:

Example

1. Check if the set $\{\mathbf{i}, 2\mathbf{i}, \mathbf{j}\}$ is linearly independent.

Solution

We form the expression

$c_1\phi_1 + c_2\phi_2 + c_3\phi_3 = 0$

where $\phi_1 = \mathbf{i}$, $\phi_2 = 2\mathbf{i}$ and $\phi_3 = \mathbf{j}$

Thus, $\mathbf{i}c_1 + 2\mathbf{i}c_2 + \mathbf{j}c_3 = 0$

or $\mathbf{i}(c_1 + 2c_2) + \mathbf{j}c_3 = 0$

which implies $c_1 + 2c_2 = 0$ and $c_3 = 0$, since **i** and **j** are non-zero vectors.

We see that $c_1 = -2c_2$, $c_3 = 0$

c_1 and c_2 do not necessarily have to be zero.

Conclusion: The set is not linearly independent.

2. Show that $\{\mathbf{i}, 2\mathbf{k}, \mathbf{j}\}$ is a linearly independent set.

Solution

$\mathbf{i}c_1 + 2\mathbf{k}c_2 + \mathbf{j}c_3 = \mathbf{0}$

$c_1 = 0,\ c_2 = 0,\ c_3 = 0$

The set is linearly independent.

Note that we have made use of the fact that

$x\mathbf{i} + y\mathbf{j} + z\mathbf{k} = 0$ implies $x = 0,\ y = 0,\ z = 0$

3. Show that the set $\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \right\}$ is linearly independent

Solution

$$c_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + c_3 \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = 0$$

from which we obtain

$c_1 + c_2 + c_3 = 0$ (i)

$c_2 + 2c_3 = 0$ (ii)

$c_1 + c_3 = 0$ (iii)

From (iii),

$c_1 = -c_3$ (iv)

and from (ii),

$c_2 = -2c_3$ (v)

Putting (iv) and (v) in (i), gives

$-c_3 - 2c_3 + c_3 = 0$

$-2c_3 = 0$ or $c_3 = 0$

$c_1 = -c_3 = 0,\ c_2 = -2c_3 = 0$

$c_1 = c_2 = c_3 = 0$

Hence, we conclude the set is linearly independent.

Note that we could have written the set of three vectors as $\{\mathbf{i} + \mathbf{k}, \mathbf{i} + \mathbf{j}, \mathbf{i} + \mathbf{j} + \mathbf{k}\}$. Try this out on your own, and be sure you can.

These vectors are not mutually orthogonal, yet, since they are linearly independent, we can write any vector in 3-dimensional Euclidean space as a linear combination of the members of the set.

Now, take the determinant of the matrix formed by each of the set in the examples and convince yourself that there is another way of checking if a set of vectors is linearly independent. We give two examples:

$$\begin{vmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{vmatrix} = 0$$

$$\begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{vmatrix} = 1(1-0) - 1(0-1) + 1(0-1) = 1 + 1 - 1 \neq 0$$

Conclusion: The set is linearly independent if the determinant is not zero, it is linearly dependent if the determinant is zero. Does that sound strange? Look at the two rows or columns of a matrix such that one can be got from the other by a linear combination. The determinant of the matrix must be zero, meaning that the vectors are linearly dependent.

3.3 Basis Vector

Let V be an n- dimensional vector space. Any set of n linearly independent vectors $\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n$ forms a basis for V. Thus, any vector $\mathbf{v} \in V$ can be expressed as a linear combination of the vectors $\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n$, i.e.,

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots x_n\mathbf{e}_n \qquad 1.10$$

Then we say that the vector space V is **spanned** by the set of vectors $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\}$. $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\}$ is said to be a **basis** for V.

If we wish to write any vector in 1 (say, x) direction, we need only one (if possible, a unit) vector. Any two vectors in the x direction must be linearly dependent, for we can write one as $a_1\mathbf{i}$ and the other $a_2\mathbf{i}$, where a_1 and a_2 are scalars.
We form the linear combination

$$c_1(a_1\mathbf{i}) + c_2(a_2\mathbf{i}) = 0 \qquad 1.11$$

where a_1 and a_2 are scalar constants.

Obviously, c_1 and c_2 need not be zero for the expression to hold, for $c_1 = -c_2\dfrac{a_2}{a_1}$ would also satisfy expression (1.11).
We conclude therefore that the vectors must be linearly dependent.

Can you then see that we can say that in general, any $n+1$ vectors in an n-dimensional space must be linearly dependent?

Example

You are quite familiar with the set of vectors (**i**, **j**) as the normal basis vectors in 2-dimensional space or a plane. Show that $(\mathbf{i}+\mathbf{j}, \mathbf{i}-\mathbf{j})$ is also a set of basis vectors for the plane.

Solution

We check for linear independence.

$$\alpha(\mathbf{i}+\mathbf{j})+\beta(\mathbf{i}-\mathbf{j})=\mathbf{0}$$

Then,

$$\mathbf{i}(\alpha+\beta)+\mathbf{j}(\alpha-\beta)=\mathbf{0}$$

This means that

$$\alpha+\beta=0$$

and

$$\alpha-\beta=0$$

Adding the last two equations makes us conclude that $\alpha=0$. Consequently, β is also 0. We conclude that the two vectors are linearly independent. Since these are two linearly independent vectors in two dimensional (Euclidean) space (a plane), they form a basis for the plane.

3.4 Inner or Scalar Product

Here, we shall expand your idea of the inner product of two vectors. In your first year in the University, you came across the dot or inner product of two vectors. In this section, we shall extend that idea, as mathematicians do, to other vector-like quantities. But first, let us take a look at the properties of an inner product.

Properties of the Inner Product

Let V be a vector space, real or complex. Then, the inner product of $\mathbf{v}$, $\mathbf{w} \in V$, written as ($\mathbf{v}$,$\mathbf{w}$), has the following properties:

(i) $(\mathbf{v},\mathbf{v}) \ge 0$ 1.11
(ii) $(\mathbf{v},\mathbf{v}) = 0$ if and only if $\mathbf{v} = \mathbf{0}$ 1.12
(iii) $(\mathbf{v},\mathbf{w}) = (\mathbf{w},\mathbf{v})$ (Symmetry) 1.13
(iv) $(c\mathbf{v},\mathbf{w}) = c^*(\mathbf{v},\mathbf{w})$; $(\mathbf{v},c\mathbf{w}) = c(\mathbf{v},\mathbf{w})$ 1.14
(v) $(\mathbf{v},\mathbf{w}+\mathbf{z}) = (\mathbf{v},\mathbf{w})+(\mathbf{v},\mathbf{z})$ 1.15
(vi) $(\mathbf{v},\mathbf{w}) \le \|\mathbf{v}\|\|\mathbf{w}\|$ 1.16

where c^* is the complex conjugate of the scalar c.

Example 1: Given the vectors **a** and **b** in 3-dimensions, i.e., V_3, we define the inner product as

$$(\mathbf{a},\mathbf{b}) = \mathbf{a}^T\mathbf{b}$$

where $\mathbf{a}^T$ is the transpose of the column matrix representing **a**. This is the dot product you have always been familiar with.

$\mathbf{a} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, and $\mathbf{b} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$. $\mathbf{a}^T = [1 \quad 0 \quad 1]$

$$(\mathbf{a},\mathbf{b}) = \mathbf{a}^T\mathbf{b} = [1 \quad 0 \quad 1]\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = 3$$

Do not mix this up

$$(\mathbf{c},\mathbf{d}) = \mathbf{c}^T\mathbf{d} = [c_x \quad c_y \quad c_z]\begin{bmatrix} d_x \\ d_y \\ d_z \end{bmatrix} = c_x d_x + c_y d_y + c_z d_z = [d_x \quad d_y \quad d_z]\begin{bmatrix} c_x \\ c_y \\ c_z \end{bmatrix} = (\mathbf{d},\mathbf{c}), \text{ with}$$

$$\mathbf{cd}^T = \begin{bmatrix} c_x \\ c_y \\ c_z \end{bmatrix}[d_x \quad d_y \quad d_z] = \begin{bmatrix} c_x d_x & c_x d_y & c_x d_z \\ c_y d_x & c_y d_y & c_y d_z \\ c_z d_x & c_z d_y & c_z d_z \end{bmatrix} \neq \mathbf{dc}^T, \text{ generally.}$$

Example 2: The space of $m \times n$ matrices, M_{mn}:

The inner product of **A** and $\mathbf{B} \in M_{mn}$ is defined as

$$(\mathbf{A},\mathbf{B}) = Tr(\mathbf{A}^{+}\mathbf{B}) \qquad 1.17$$

where $\mathbf{A}^{+} = \overline{\mathbf{A}^T}$, the complex conjugate of the transpose of **A**. Indeed, it does not matter in what order, so it could also be the transpose of the complex conjugate of **A**. If **A** is a real matrix, then there is no need taking the complex conjugate. In that case, $\mathbf{A}^{+} = \mathbf{A}^T$. $Tr(\mathbf{P})$ is the trace of the matrix **P**, the sum of the main diagonal elements of **P**.

e.g., let $\mathbf{A} = \begin{bmatrix} i & 0 \\ 1 & 1 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 0 & -i \\ 1 & 0 \end{bmatrix}$

$$\mathbf{A}^T = \begin{bmatrix} i & 1 \\ 0 & 1 \end{bmatrix}; \; \mathbf{A}^{+} = \overline{\mathbf{A}^T} = \begin{bmatrix} -i & 1 \\ 0 & 1 \end{bmatrix}$$

$$\mathbf{A}^{+}\mathbf{B} = \overline{\mathbf{A}^T}\mathbf{B} = \begin{bmatrix} -i & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & -i \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$$

$$(\mathbf{A},\mathbf{B}) = Tr(\mathbf{A}^{+}\mathbf{B}) = 1 + 0 = 1$$

Example 3: The space of square integrable complex valued functions, F_s, over the interval $[a,b]$, i.e., $f(x) \in F_s$ implies that $\int_a^b |f(x)|^2 dx < \infty$.

We define the inner product on this space by

$$(f,g) = \int_a^b f^*(x) g(x) dx \qquad 1.18$$

where $f^*(x)$ is the complex conjugate of $f(x)$.

Later, you shall see that this space is of utmost importance in Quantum Mechanics.

3.5 Norm of a Vector

Let X be a vector space over K, the real or complex number field. A real valued function $\|\cdot\|$ on X is a norm on X (i.e., $\|\cdot\|: X \to R$) if and only if the following conditions are satisfied:

(i) $\|\mathbf{x}\| \geq 0$ 1.19

(ii) $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = \mathbf{0}$ 1.20

(iii) $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ $\forall \mathbf{x}, \mathbf{y} \in X$ (Triangle inequality) 1.21

(iv) $\|\alpha\, \mathbf{x}\| = |\alpha|\|\mathbf{x}\|$ $\forall \mathbf{x} \in X$ and $\alpha \in C$ (Absolute homogeneity) 1.22

The norm of a vector is its “distance” from the origin. Once again, you can see the basic idea of the distance of a point from the origin being generalised to the case of the vectors in any vector space.

$\|\mathbf{x}\|$ is called the norm of **x**.

In the case where $X = R$, the real number line, the norm is the absolute value, $|\mathbf{x}|$.

If the norm of **v** in the vector space V is unity, such a vector is said to be normalised. In any case, even if a vector is not normalised, we can normalise it by dividing by the norm.

Example 1: Given the vector **a** in V_3, the norm of **a** is

$$\|\mathbf{a}\| = \sqrt{(\mathbf{a},\mathbf{a})} \qquad 1.23$$

Thus, if **a** = $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, then

$$(\mathbf{a},\mathbf{a}) = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 2$$

$$\|\mathbf{a}\| = \sqrt{(\mathbf{a},\mathbf{a})} = \sqrt{2}$$

We see that **a** is not normalised.

However, $\mathbf{c} = \dfrac{\mathbf{a}}{\|\mathbf{a}\|} = \dfrac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ is normalised.

Example 2: The space of $m \times n$ matrices:
Given the $m \times n$ matrix **A**, then the norm of **A** is defined as

$$\|\mathbf{A}\| = \sqrt{Tr(\mathbf{A},\mathbf{A})} \qquad 1.24$$

e.g., let $\mathbf{A} = \begin{bmatrix} i & 0 \\ 1 & 1 \end{bmatrix}$

$$\mathbf{A}^T = \begin{bmatrix} i & 1 \\ 0 & 1 \end{bmatrix}; \ \overline{\mathbf{A}^T} = \begin{bmatrix} -i & 1 \\ 0 & 1 \end{bmatrix}$$

$$\overline{\mathbf{A}^T}\mathbf{A} = \begin{bmatrix} -i & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} i & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

$$Tr(\mathbf{A}^{+}\mathbf{A}) = 2 + 1 = 3$$

Therefore,

$$\|\mathbf{A}\| = \sqrt{3}$$

$\mathbf{A}$ is not normalised, but $\mathbf{C} = \dfrac{\mathbf{A}}{\|\mathbf{A}\|} = \dfrac{1}{\sqrt{3}}\begin{bmatrix} i & 0 \\ 1 & 1 \end{bmatrix}$ is normalised.

Example 3: The space of square integrable complex valued functions, F_s, over the interval $[a,b]$,

Let $f(x) \in F_s$, then we define

$$\|f\| = \sqrt{(f, f)} \qquad 1.25$$

where $(f, f) = \int_a^b |f(x)|^2 \, dx$

f might not be normalised, but $h = \dfrac{f}{\sqrt{(f, f)}}$ is normalised.

It is now obvious that we have to deal with a square integrable set of functions. We want to deal with only functions that we can normalise.

EXAMPLES

Exercise

(i) Normalise each member of the set, and hence expand the vector $\{4\mathbf{i} + 3\mathbf{j} - 4\mathbf{k}\}$ as a linear combination of the normalised set.

(ii) Is the set $\left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \right\}$ linearly dependent or independent? Normalise each vector.

4.0 Conclusion

In this unit, you have learnt about vector spaces, a generalisation of the idea of vectors you have all along been familiar with, expanded to cover matrices, certain functions and all mathematical structures that satisfy the basic laws of vector spaces. You also came across linear independence, and saw the example of the vectors **i** and **j** in two-dimensional Euclidean space, and with the help of linearly independent vectors, we were able to define a basis with which we could specify any vector in any given vector space.

Then, you were introduced to the idea of the norm, a generalisation of the idea of the distance of a vector from the origin. Finally, you learnt how to normalise a vector.

5.0 Summary

In this Unit, you learnt the following:

- Vector spaces are sets that contain some vector-like quantities that satisfy certain conditions.
- How to check whether a set of vectors is linearly independent.
- A set of linearly independent vectors is necessary to span a space.
- *n*-dimensional vector space V is **spanned** by the set of *n* vectors.
- The norm of a vector is its distance from the 'origin.'
- Dividing a vector by its norm normalises it, so that its length is unity.

6.0 Learning Activity

1. Show that the following are vector spaces over the indicated field:
   (i) The set of real numbers over the field of real numbers.
   (ii) The set of complex numbers over the field of real numbers.
   (iii) The set of quadratic polynomials over the complex field.

2. Check whether the following vectors are linearly independent.
   (i) $2i+3j-k$, $-i+j+3k$ and $-3i+2j+k$

3. Show whether or not the set $\left\{\begin{pmatrix}1\\1\end{pmatrix},\begin{pmatrix}-1\\-1\end{pmatrix}\right\}$ is a basis for the two-dimensional Euclidean space.

4. Find the coordinates of the vector $\begin{bmatrix}1 & 2\\-2 & i\end{bmatrix}$ with respect to the basis

$$\left\{\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix},\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix},\begin{bmatrix}0 & -i\\i & 0\end{bmatrix},\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}\right\}.$$

5. Find the inner product of the following vectors:
   (i) $\begin{pmatrix}i\\-2\\2\end{pmatrix}$ and $\begin{pmatrix}2\\-1\\3\end{pmatrix}$
   (ii) ix^2+2 and $2x-3i \quad 0\le x\le 2$.
   (iii) $\mathbf{A}$, $\mathbf{B}\in M_{mn}$ if $\mathbf{A}=\begin{bmatrix}2 & 1 & 1\\-1 & 3 & 1\end{bmatrix}$ and $\mathbf{B}=\begin{bmatrix}-1 & -1 & 2\\1 & 3 & 1\end{bmatrix}$.

6. Find the norm of the following:

(i) $\begin{pmatrix} 2i \\ -1 \\ 3 \end{pmatrix}$ (ii) $ix^2 + 2,\ 0 \le x \le 1$ (iii) $\mathbf{D} = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 1 & 3 \\ 3 & 1 & 2 \end{bmatrix}$

7. Normalise each vector in the set $\left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \right\}$.

7.0 References/Further Readings

1. Mathematical Physics – Butkov, E.
2. Mathematical Methods for Physics and Engineering – Riley, K. F., Hobson, M. P. and Bence, S. J.

Solutions to Learning Activity

1. Show that the following are vector spaces over the indicated field:

(i) The set of real numbers over the field of real numbers.

Let the set be R be the set of real numbers, then,

$$a + b \in R \qquad \forall \ a, b \in R$$

and $$\lambda\, a \in R \qquad \forall \ a \in R,\ \lambda \in R$$

(ii) The set of complex numbers over the field of real numbers.

Let the set be C be the set of complex numbers, then,

$$c_1 + c_2 \in C \qquad \forall \ a, b \in C$$

and $$\alpha\, c \in R \qquad \forall \ c \in C,\ \alpha \in C$$

(iii) The set of quadratic polynomials over the complex field.

Let this set be P. Then $P_1 = a_1x^2 + b_1x^2 + c_1$ and $P_2 = a_2x^2 + b_2x^2 + c_2$ are in P, where $a_1, a_2, b_1, b_2, c_1, c_2$ are constants.

$$a_1x^2 + b_1x^2 + c_1 + a_2x^2 + b_2x^2 + c_2$$

$$= (a_1 + a_2)x^2 + (b_1 + b_2)x + (c_1 + c_2) \in \text{P} \ \forall\ P_1,\ P_2 \in \text{P}$$

$\lambda\,(a_1x^2 + b_1x + c_1) \in$ P $\forall$ $P_1 \in$ P, $\lambda \in$ the complex field.

2. Check whether the following vectors are linearly independent.

(i) $2i + 3j - k,\ -i + j + 3k$ and $-3i + 2j + k$

$$a\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} + b\begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix} + c\begin{pmatrix} -3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$2a - b - 3c = 0$$
$$3a + b + 2c = 0$$
$$-a + 3b + c = 0$$

The solution set is (0, 0, 0), i.e., $a = b = c = 0$.

The vectors are linearly independent.

Alternatively,

$$\begin{vmatrix} 2 & -1 & -3 \\ 3 & 1 & 2 \\ -1 & 3 & 1 \end{vmatrix} = 2(1-6) + 1(3+2) - 3(9+1)$$

$$= 2(-5) + 5 - 30 = -35 \neq 0$$

(ii) $$\begin{bmatrix} i & 1 \\ -2 & 2i \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ -i & 2i \end{bmatrix}, \begin{bmatrix} -1 & 2 \\ 3 & -i \end{bmatrix} \text{ and } \begin{bmatrix} -i & 2i \\ i & -2 \end{bmatrix}$$

$$a\begin{bmatrix} i & 1 \\ -2 & 2i \end{bmatrix}+b\begin{bmatrix} 2 & 1 \\ -i & 2i \end{bmatrix}+c\begin{bmatrix} -1 & 2 \\ 3 & -i \end{bmatrix}+d\begin{bmatrix} -i & 2i \\ i & -2 \end{bmatrix}=\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Expanding,

$$ai+2b-c-id=0 \qquad \text{(i)}$$

$$a+b+2c+2id=0 \qquad \text{(ii)}$$

$$-2a-ib+3c+id=0 \qquad \text{(iii)}$$

$$2ia+2ib-ic-2d=0 \qquad \text{(iv)}$$

Multiplying (i) by 2 and adding to (ii),

$$a(1+2i)+5b=0 \qquad \text{(v)}$$

Multiplying (iii) by i and adding to (iv) gives

$$(1+2i)b+2ic-3d=0 \qquad \text{(vi)}$$

Multiplying (ii) by 2 and adding to (iii),

$$(2-i)b+7c+5id=0 \qquad \text{(vii)}$$

Multiplying (vi) by 5i and (vii) by 3 and adding,

$$5i(1+2i)b-10c-15id=0 \qquad \text{(vi)}$$

$$3(2-i)b+21c+15id=0 \qquad \text{(vii)}$$

$$(5i-10)b-10c-15id=0$$

$$(6-3i)b+21c+15id=0$$

$$(-4+2i)b+11c=0 \qquad \text{(viii)}$$

From (v) and (viii), $b=-\frac{a(1+2i)}{5}=\frac{11}{2i-4}c$

Hence,

$$c=\frac{a(1+2i)(2i-4)}{55}=-\frac{8+6i}{55}a$$

Substituting for b and c in equation (vi),

$$(1+2i)\left(-\frac{(1+2i)}{5}a\right)+2i\left(-\frac{8+6i}{55}a\right)-3d=0$$

$$\frac{(3-4i)}{5}a+-\frac{16i-12}{55}a=3d$$

$$\frac{33-44i-16i+12}{165}a=\frac{45-60i}{165}a=\frac{9-12i}{55}a=d \qquad \text{(ix)}$$

Putting b, c, d in (i),

$$ai+2\left(-\frac{a(1+2i)}{5}\right)+\frac{8+6i}{55}a+i\frac{12i-9}{55}a=0$$

$$ai-\frac{2a}{5}-\frac{4ai}{5}+\frac{8a}{55}+\frac{6ai}{55}-\frac{12a}{55}-\frac{9ia}{55}=0$$

$$ai\left(1-\frac{4}{5}+\frac{6}{55}-\frac{9}{55}\right)+a\left(\frac{8}{55}-\frac{2}{5}-\frac{12}{55}\right)=0$$

$$\frac{55-44+6-9}{55}ai+a\frac{8-22-12}{55}=0$$

$a\left(\frac{18}{55}i - \frac{26}{55}\right) = 0$. Hence, a = 0, meaning that *b*, *c*, and *d* are also zero.

$ai + 2b - c - id = 0$ (i)
$a + b + 2c + 2id = 0$ (ii)
$-2a - ib + 3c + id = 0$ (iii)
$2ia + 2ib - ic - 2d = 0$ (iv)

$$\text{Check if } \begin{vmatrix} i & 2 & -1 & -i \\ 1 & 1 & 2 & 2i \\ -2 & -i & 3 & i \\ 2i & 2i & -i & -2 \end{vmatrix} \neq 0$$

3. Show whether or not the set $\left\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}-1\\-1\end{pmatrix}\right\}$ is a basis for the two-dimensional Euclidean space.
For the set to be a basis, the vectors must be linearly independent.

$$a\begin{pmatrix}1\\1\end{pmatrix} + b\begin{pmatrix}-1\\-1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$$
$$a - b = 0,\ a - b = 0$$
$$a = b$$

a and *b* do not have to be zero. Hence, the vectors are not linearly independent. Sketch the vectors and satisfy yourself that they are indeed linearly dependent: one can be got from the other because they degenerate into a line.

Alternately,

$$\begin{vmatrix} 1 & -1 \\ 1 & -1 \end{vmatrix} = 0$$

4. Find the coordinates of the vector $\begin{bmatrix} 1 & 2 \\ -2 & i \end{bmatrix}$ with respect to the basis

$$\left\{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\right\}.$$

$$\begin{bmatrix} 1 & 2 \\ -2 & i \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} + d\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$1 = a + d$ (i)
$2 = b - ic$ (ii)
$-2 = b + ci$ (iii)
$i = a - d$ (iv)

Adding (i) and (iv):

$$\frac{1+i}{2}=a$$

(i) – (iv):

$$\frac{1-i}{2}=d$$

(ii) + (iii):

$$0=b$$

(iii) – (ii):

$$-\frac{2}{i}=2i=c$$

Hence,

$$\begin{bmatrix}1 & 2\\ -2 & i\end{bmatrix}=\frac{1+i}{2}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}+0\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}+2i\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}+\frac{1-i}{2}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$$

5. Find the inner product of the following vectors:

(i) $\begin{pmatrix} i\\ -2\\ 2\end{pmatrix}$ and $\begin{pmatrix} 2\\ -1\\ 3\end{pmatrix}$

$$\begin{pmatrix}-i & -2 & 2\end{pmatrix}\begin{pmatrix}2\\ -1\\ 3\end{pmatrix}=-2i+2+6=8-2i$$

(ii) ix^2+2 and $2x-3i \quad 0\le x\le 2$.

$$\int_0^2 (ix^2+2)^{*}\times(2x-3i)dx=\int_0^2(-ix^2+2)^{*}\times(2x-3i)dx$$

$$=\int_0^2(2ix^3-3x^2+4x-6i)dx$$

$$=\left[i\frac{x^4}{2}-x^3+2x^2-6ix\right]_0^2$$

$$=8i-8+8-12i$$

$$-4i$$

(iii) $\mathbf{A}\,,\ \mathbf{B}\in M_{mn}$ if $\mathbf{A}=\begin{bmatrix}2 & 1 & 1\\ -1 & 3 & 1\end{bmatrix}$ and $\mathbf{B}=\begin{bmatrix}-1 & -1 & 2\\ 1 & 3 & 1\end{bmatrix}$.

$$(A,B)=Tr(A^{+}B)=Tr\left(\begin{bmatrix}2 & -1\\ 1 & 3\\ 1 & 1\end{bmatrix}\begin{bmatrix}-1 & -1 & 2\\ 1 & 3 & 1\end{bmatrix}\right)=Tr\begin{bmatrix}-2-1 & -2-3 & 4-1\\ -1+3 & -1+9 & 2+3\\ -1+1 & -1+3 & 2+1\end{bmatrix}$$

$$= Tr\begin{bmatrix} -3 & -5 & 3 \\ 2 & 8 & 5 \\ 0 & 2 & 3 \end{bmatrix} = 8$$

6. Find the norm of the following:

(i) $\begin{pmatrix} 2i \\ -1 \\ 3 \end{pmatrix}$ (ii) $ix^2 + 2,\ 0 \le x \le 1$ (iii) $\mathbf{D} = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 1 & 3 \\ 3 & 1 & 2 \end{bmatrix}$

(i)

$$\sqrt{\begin{pmatrix} -2i & -1 & 3 \end{pmatrix}\begin{pmatrix} 2i \\ -1 \\ 3 \end{pmatrix}} = \sqrt{4+1+9} = \sqrt{14}$$

(ii)

$$\int_0^1 (ix^2 + 2)^*(ix^2 + 2)dx = \int_0^1 (-ix^2 + 2)(ix^2 + 2)dx$$

$$= \int_0^1 (4 + x^4)dx = \left[4x + \frac{x^5}{5}\right]_0^1 = 4 + \frac{1}{5} = \frac{21}{5}$$

Norm = $\sqrt{\frac{21}{5}}$

7. Normalise each vector in the set $\left\{\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}\right\}$.

Norm of $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ is $\sqrt{1+4+9} = \sqrt{14}$

The normalised vector is $\frac{1}{\sqrt{14}}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$

Similarly, $\frac{1}{\sqrt{20}}\begin{pmatrix} -2 \\ 0 \\ 4 \end{pmatrix}$ and $\frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ are normalised.

2: ORTHOGONALITY AND ORTHONOMALITY

1.0 Introduction

Orthogonal functions play an important role in Quantum mechanics. This is because they afford us a set of functions 'which do not mix,' just the way you could resolve a vector in two dimensions in the x and y directions, respectively, with the unit vectors **i** and **j**. The dot product of the two unit vectors gives you zero. We would also like to resolve our vectors in some 'directions.' Thus, you need to know about orthogonal and orthonormal functions. The orthonormal functions would form the possible states you can find a system. You know such states should not 'mix.' In this Unit, you will learn about orthonormality and orthogonality; how to create an orthogonal and subsequently, an orthonormal set and expand a given function in terms of an orthonormal set. This would naturally lead to an analysis of the probability of finding a system in any of the states in the orthonormal set. This Unit also gives you an insight into some elements of matrix algebra.

2.0 Objectives

This Unit will equip you with the knowledge of:

- Orthogonal functions
- Orthonormal functions
- Expansion of a given function as a linear combination of a set of orthonormal functions (states).
- Recovering the coefficient of the expansion.
- Finding the probability of finding the system in a given state.
- Some elements of matrix algebra.

3.0 Main Content

3.1 Definitions

(i) We say $\mathbf{v}_1$ and $\mathbf{v}_2$ in a vector space V are orthogonal if their inner product is zero, that is, $(\mathbf{v}_1, \mathbf{v}_2) = 0$.

(ii) Suppose there exists a linearly independent set $\{\phi_i\}_{i=1}^{n}$, i.e., $\{\phi_1, \phi_2, \cdots, \phi_n\}$, such that $(\phi_i, \phi_j) = 0$, $i \neq j$, then, $\{\phi_i\}_{i=1}^{n}$ is an orthogonal set.

(iii) If in addition to condition (ii) above, $(\phi_i, \phi_i) = 1$, then, $\{\phi_i\}_{i=1}^{n}$ is an orthonormal set.

For an orthonormal set, therefore, we can write $(\phi_i, \phi_j) = \delta_{ij}$, where δ_{ij} is the Kronecker delta, equal to 0 if $i \neq j$ and equal to 1 if $i = j$.

As we have seen earlier, if any vector in the vector space, V, can be written as a linear combination

$$\mathbf{v} = a_1\phi_1 + a_2\phi_2 + \cdots + a_n\phi_n = \sum_{i=1}^{n} a_i\phi_i \qquad 2.1$$

then we say the space is spanned by the complete orthonormal basis $\{\phi_i\}_{i=1}^{n}$, where

$$(\phi_m, \phi_n) = \delta_{mn} \qquad 2.2$$

If $\{\phi_i\}_{i=1}^{n}$ is an orthonormal set, It follows that we can recover the coefficient of expansion as follows:

$$(\phi_j, \mathbf{v}) = (\phi_j, \sum_{i=1}^{n} a_i\phi_i) = \sum_{i=1}^{n} a_i(\phi_j, \phi_i) = a_j \qquad 2.3$$

Moreover,

$$(\mathbf{v}, \mathbf{v}) = (\sum_{k=1}^{n} a_k\phi_k, \sum_{i=1}^{n} a_i\phi_i) = \sum_{k=1}^{n} a_k * \sum_{i=1}^{n} a_i(\phi_j, \phi_i) = \sum_{i=1}^{n} |a_i|^2 \qquad 2.4$$

If, in addition, the vector $\mathbf{v}$ is normalised, then

$$\sum_{i=1}^{n} |a_i|^2 = 1 \qquad 2.5$$

Do you remember what you learnt about probability in Statistics? The sum of the probability for various possible events is unity. Thus, we can interpret the $|a_i|^2$ as the probability that the system which has n possible states, assumes state i with probability $|a_i|^2$. In other words, the probability that the system is in state i is $|a_i|^2$.

3.2 Bra and Ket (Dirac) Notation

We have written the inner product in the form $(\cdot,\cdot)$. We could also write it in the form of a bra, $|\cdot>$, and a ket, $<\cdot|$. This is the Dirac notation. Putting the bra and the ket together forms a 'bracket' $<\cdot|\cdot>$. The set of vectors $\{\phi_j\}_{j=1}^{n}$ can be seen as a set of bra vectors (space of vectors) $\{|\phi_j>\}_{j=1}^{n}$. Then, we would need a dual set of vectors (dual space of vectors) $\{<\phi_j|\}_{j=1}^{n}$ to be able to write the inner product. Why? Recall that we needed to change our column vectors to row vectors to be able to take the inner product of two column vectors? If $|B>$ is a column vector, then $<B|$ is the dual vector, the row vector but with the entries being the complex conjugate of what they were as $|B>$.

It follows from the foregoing, that we can write the expansion of a wavefunction

$$\psi = \sum_j c_j \phi_j \text{ as } \psi = \sum_{j=1}^{n} c_j |\phi_j> \qquad 2.6$$

Moreover, $(\phi_j, a\phi_j) = a(\phi_j, \phi_j)$ and $(a\phi_j, \phi_j) = a^*(\phi_j, \phi_j)$. It follows that $a(\phi_j, \phi_j) = (\phi_j, a\phi_j) = (a^*\phi_j, \phi_j) = (a^*)^*(\phi_j, \phi_j)$. We can extract the following rule from this:

$$(\phi_j, a\phi_j) = (a^*\phi_j, \phi_j) \qquad 2.7$$

More generally, a could be an operator A. Then,

$$(\phi_j, A\phi_j) = (A^+\phi_j, \phi_j) \qquad 2.8$$

We can write this in the form,

$$<\phi_j | A | \phi_j> = <A^+\phi_j | \phi_j>$$

Equations 2.3 and 2.4 now become,

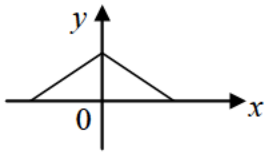
$$<\phi_j, \mathbf{v}> = <\phi_j | \sum_{i=1}^{n} a_i |\phi_i> = \sum_{i=1}^{n} a_i <\phi_j | \phi_i> = a_j \qquad 2.9$$

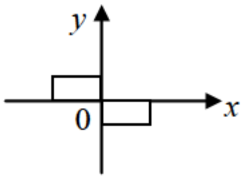
$$<\mathbf{v}|\mathbf{v}> = <\sum_{k=1}^{n} a_k \phi_k | \sum_{i=1}^{n} a_i \phi_i> = \sum_{k=1}^{n} a_k * \sum_{i=1}^{n} a_i <\phi_j | \phi_i> = \sum_{i=1}^{n} |a_i|^2 \qquad 2.10$$

3.3 Orthogonal Functions

An even function is symmetrical about the y axis. In other words, a plane mirror placed on the axis will produce an image that is exactly the function across the axis. An example is shown in Fig …a. An odd function will need to be mirrored twice, once along the y axis, and once along the x axis to achieve the same effect. Fig. … b is an example of an odd function.

A function $f(x)$ of x is said to be an odd function if $f(-x) = f(x)$, e.g., $\sin x$, x^{2n+1}, and a function $f(x)$ of x is said to be an even function if $f(-x) = f(x)$, e.g., $\cos x$, x^{2n} where $n = 0, 1, 2, \ldots$.





Odd function Even function

Fig. …

Some real-valued functions are odd, some are even; the rest are neither odd nor even. However, we can write any real-valued function as a sum of an odd and an even function.

Let the function be $h(x)$, then we can write

$$h(x) = f(x) + g(x) \qquad 2.11$$

where $f(x)$ is odd and $g(x)$ is even. Then, $f(-x) = -f(x)$ and $g(-x) = g(x)$

$$h(-x) = f(-x) + g(-x) = -f(x) + g(x) \qquad 2.12$$

Adding equations (2.11) and (2.12) gives

$$h(x) + h(-x) = 2g(x)$$

Subtracting equation (2.12) from equation (2.11) gives

$$h(x) - h(-x) = 2f(x)$$

It follows, therefore, that

$$f(x) = \frac{h(x) - h(-x)}{2} \qquad 2.13$$

and

$$g(x) = \frac{h(x) + h(-x)}{2} \qquad 2.14$$

Example

Write the function $h(x) = e^{2x}\sin x$ as a sum of odd and even functions.

Solution

$h(x) = e^{2x}\sin x$, $h(-x) = e^{-2x}\sin(-x) = -e^{-2x}\sin x$

Therefore, the odd function is

$$f(x) = \frac{h(x) - h(-x)}{2} = \frac{e^{2x}\sin x + e^{-2x}\sin x}{2} = \frac{e^{2x} + e^{-2x}}{2}\sin x$$

$$= \cosh 2x \sin x$$

The even function is

$$g(x) = \frac{h(x) + h(-x)}{2} = \frac{e^{2x}\sin x - e^{-2x}\sin x}{2} = \frac{e^{2x} - e^{-2x}}{2}\sin x$$

$$= \sinh 2x \sin x$$

It is obvious that the odd function is a product of an odd function and an even function. Likewise, the even function is a product of two odd functions. We conclude, therefore, that the following rules apply:

Even × Even = Even 2.15
Even × Odd = Odd 2.16

$$\text{Odd} \times \text{Odd} = \text{Even} \qquad 2.17$$

The integral

$$\int_{-a}^{a} f(x)dx = 0 \text{ if } f(x) \text{ is odd} \qquad 2.18$$

$$\int_{-a}^{a} f(x)dx = 2\int_{0}^{a} f(x)dx \text{ if } f(x) \text{ is even} \qquad 2.19$$

Recall that the inner product in the space of twice integrable complex valued functions of two complex valued functions $f(x)$ and $g(x)$ over the interval $a \le x \le b$ is defined as $(f, g) = \int_{a}^{b} f^*(x)g(x)dx$.

Two functions $f(x)$ and $g(x)$ are said to be **orthogonal** over an interval $a \le x \le b$ if their inner product is zero.

Example
Show that $\sin mx$ and $\sin nx$ are orthogonal, $m \ne n$, $-\pi \le x \le \pi$.

Solution

The inner product is $\int_{-\pi}^{\pi} \sin mx \sin nx dx = \frac{1}{2}\int_{-\pi}^{\pi} [\cos(m-n)x - \cos(m+n)x]dx$

$$= \frac{1}{2}\left[\frac{1}{m-n}\sin(m-n)x\Big|_{-\pi}^{\pi} + \frac{1}{m+n}\sin(m+n)x\Big|_{-\pi}^{\pi} \right] = 0$$

3.4 Gram-Schmidt Orthogonalisation Procedure

This provides a method of constructing an orthogonal set from a given set. Normalising each member of the set then provides an orthonormal set. The method entails setting up the first vector, and then constructing the next member of the orthogonal set by making it orthogonal to the first member of the set under construction. Then the next member of the set is constructed in a way to be orthogonal to the two preceding members. This procedure can be continued until the last member of the set is constructed.

3.4.1 Example from function vector space

Construct an orthonormal set from the set $\{1, x, x^2, ...\}$ over the interval $-1 \le x \le 1$. Thus, given the set $\{f_1, f_2, f_3, ...\}$, we want to construct an orthogonal set $\{\phi_1, \phi_2, \phi_3, ...\}$, i.e., $\int_{-1}^{1} \phi_i(x)\phi_j(x)dx = 0$, if $i \ne j$, then we normalise each member of the set.

Let $\phi_1 = f_1 = 1$, and $\phi_2 = f_2 + \alpha\phi_1 = x + \alpha$
Then, we determine α, subject to
$(\phi_1, \phi_2) = 0$

$$\int_{-1}^{1} 1\cdot(x+\alpha)dx = \left.\frac{x^2}{2}\right|_{-1}^{1} + \left.\alpha x\right|_{-1}^{1} = 0$$ 2.20

$\Rightarrow \alpha = 0$

Thus, $\phi_2 = x$

Let $\phi_3 = f_3 + \alpha\phi_2 + \beta\phi_1 = x^2 + \alpha x + \beta$

subject to $(\phi_1, \phi_3) = 0$ and $(\phi_2, \phi_3) = 0$

The first condition gives:

$$\int_{-1}^{1} 1\cdot(x^2 + \alpha x + \beta)dx = 0$$ 2.21

or

$$\left.\frac{x^3}{3}\right|_{-1}^{1} + \left.\frac{\alpha x^2}{2}\right|_{-1}^{1} + \beta x\Big|_{-1}^{1} = \left.\frac{2x^2}{3} + 2\beta x\right|_{0}^{1} = 0$$ 2.22

$$\frac{2}{3} + 2\beta = 0$$ 2.23

or

$$\beta = -\frac{1}{3}$$ 2.24

The second condition gives

$$\int_{-1}^{1} x\cdot(x^2 + \alpha x + \beta)dx = \int_{-1}^{1}(x^3 + \alpha x^2 + \beta x)dx$$

or

$$\left.\frac{x^4}{4}\right|_{-1}^{1} + \left.\frac{\alpha x^3}{3}\right|_{-1}^{1} + \left.\frac{\beta x^2}{2}\right|_{-1}^{1} = 0$$ 2.25

$$\left.\frac{2\alpha x^3}{3}\right|_{0}^{1} = 0$$

or

$$\alpha = 0$$ 2.26

Putting the values of α and β from equations (2.24) and (2.26) into the expression $\phi_3 = f_3 + \alpha\phi_2 + \beta\phi_1 = x^2 + \alpha x + \beta$, we arrive at

$$\phi_3 = x^2 - \frac{1}{3}$$ 2.27

ϕ_4, ϕ_5, etc., can be got in a similar fashion.

To normalise ϕ_j, we multiply the function by a normalisation constant, A, say, and invoke the relation

$$\int_{-1}^{1} A^2 \phi_j^{\,2}(x)dx = 1$$ 2.28

For ϕ_1, this becomes

$$\int_{-1}^{1} A^2 1^2 \, dx = 2\int_0^1 A^2 dx = 1$$

from which

$$2A^2 = 1$$

or

$$A = \sqrt{\frac{1}{2}}$$

The normalised function

$$\psi_1 = \frac{1}{\sqrt{2}} \tag{2.29}$$

Similarly,

$$\int_{-1}^{1} A^2 x^2 dx = 1$$

$$A^2 \left.\frac{x^3}{3}\right|_{-1}^{1} = A^2\left[\frac{1}{3}+\frac{1}{3}\right] = 2\frac{A^2}{3} = 1$$

Thus, $A^2 = \sqrt{\frac{3}{2}}$.

Hence, the normalised function,

$$\psi_2 = \sqrt{\frac{3}{2}}x$$

In like manner,

$$\int_{-1}^{1} A^2\left(x^2 - \frac{1}{3}\right)^2 dx = 2\int_0^1 A^2\left(x^4 - \frac{2}{3}x^2 + \frac{1}{9}\right)dx = 1$$

from which

$$2A^2 \left.\frac{x^5}{5} - \frac{2x^3}{9} + \frac{x}{9}\right|_0^1 = 1$$

or

$$2A^2\left(\frac{1}{5} - \frac{2}{9} + \frac{1}{9}\right) = 1$$

Therefore, $\frac{80}{45}A^2 = 1$

The normalised function $\psi_2 = \sqrt{\frac{45}{8}}\left(x^2 - \frac{1}{3}\right)$

3.4.2 Example from R^n

We define the projection operator

$$\mathrm{Pr}\,oj_{\mathbf{u}}\mathbf{v} = \frac{<\mathbf{u},\mathbf{v}>}{<\mathbf{u},\mathbf{u}>}\mathbf{u} \tag{2.30}$$

$$\mathbf{u}_1 = \mathbf{v}_1 \tag{2.31}$$

$$\mathbf{u}_2 = \mathbf{v}_2 - \mathrm{Pr}_{\mathbf{u}_1}\,\mathbf{v}_2 \tag{2.32}$$

$$\mathbf{u}_3 = \mathbf{v}_3 - \mathrm{Pr}_{\mathbf{u}_1}\,\mathbf{v}_3 - \mathrm{Pr}_{\mathbf{u}_2}\,\mathbf{v}_3 \tag{2.33}$$

.

.

$$\mathbf{u}_n = \mathbf{v}_n - \sum_{i=1}^{n-1}\mathrm{Pr}_{\mathbf{u}_i}\,\mathbf{v}_n \tag{2.34}$$

We can then normalise each vector

$$\mathbf{e}_k = \frac{\mathbf{u}_k}{\|\,\mathbf{u}_k\,\|} \tag{2.35}$$

Note that $\mathrm{Pr}_{\mathbf{u}}\,\mathbf{v}$ projects vector **v** orthogonally onto vector **u**.

3.5 Some Useful Mathematics on Matrices

You shall be needing the following because we often represent an operator in quantum mechanics by a matrix. We shall take as the usual basis in 3-dimensional space, $\{\mathbf{e}_1,\ \mathbf{e}_2,\ \mathbf{e}_3\}$. You may also see this basis as {**i**, **j**, **k**}.

3.5.1 Orthogonal Matrices

A tensor Q such that $(Q\underline{a})\cdot(Q\underline{b}) = \underline{a}\cdot\underline{b} \quad \forall \quad \underline{a},\underline{b} \in E$ is called an **orthogonal matrix**. Since $(Q\underline{a})\cdot(Q\underline{b}) = \underline{b}\cdot\{Q^T(Q\underline{a})\} = \underline{b}\cdot\{(Q^TQ)\underline{a}\}$, a necessary and sufficient condition for Q to be orthogonal is

$$QQ^T = I \tag{2.36}$$

or equivalently,

$$Q^{-1} = Q^T \tag{2.37}$$

Note that

$$\begin{aligned}\det(QQ^T) &= \det(Q)\det(Q^T)\\ &= \det(Q)\det(Q)\\ &= \left(\det(Q)\right)^2 = 1\end{aligned}$$

$$\Rightarrow \det(Q) = \pm 1 \tag{2.38}$$

Q is said to be a **proper orthogonal matrix** if $\det(Q)=1$ and an **improper orthogonal matrix** if $\det(Q) = -1$.

If $\det(Q)=1$, then

$$\begin{aligned}\det(Q-I) &= \det(Q-I)\det(Q^T)\\ &= \det(QQ^T - Q^T)\quad (\det(A)\det(B) = \det(AB) \text{ for any 2 square matrices})\end{aligned}$$

$$
\begin{aligned}
&= \det(I - Q^T) && (QQ^T = I \text{ for an orthogonal matrix } Q)\\
&= \det(I^T - Q^{TT}) && (\det A = \det A^T \text{ for any square matrix } A.)\\
&= +\det(I - Q) && (I^T = I \text{ and } Q^{TT} = I)\\
&= -\det(Q - I) && (\det(-A) = -\det(A) \text{ for any square matrix } A.)\\
&= 0 && (\text{if a number is equal to its negative, it must be zero})
\end{aligned}
$$

Therefore, 1 is an eigenvalue so that $\exists\ \mathbf{e}_3 \ni Q\mathbf{e}_3 = \mathbf{e}_3$.

Choose $\mathbf{e}_1$, $\mathbf{e}_2$ to be orthonormal to $\mathbf{e}_3$. In terms of this basis,

$$Q = \begin{bmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.39}$$

$$Q^T = \begin{bmatrix} a & c & 0 \\ b & d & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.40}$$

$$QQ^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a^2 + b^2 & ac + bd & 0 \\ ca + bd & c^2 + d^2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.41}$$

$$a^2 + b^2 = 1 = c^2 + d^2 \tag{2.42}$$

$$ac + bd = 0 = ca + bd \tag{2.43}$$

Also,

$$\det(Q) = 1 = ad - bc \tag{2.44}$$

From equation 2.43, $b = -\dfrac{ac}{d}$

Putting this in 2.43 gives

$$ad + \frac{ac^2}{d} = 1 \tag{2.45}$$

$\Rightarrow a(c^2 + d^2) = d \Rightarrow a = d$ Use equation 2.43 in equation 2.42 to get $c = -b$.

Therefore,

$$Q = \begin{bmatrix} a & b & 0 \\ -b & a & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.46}$$

with $a^2 + b^2 = 1$.

Thus, $\exists\ \theta$, $\ni\ a = \cos\theta$, $b = \sin\theta$,

so

$$Q = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad 2.47$$

If you represent the three unit vectors in 3-dimensional Euclidean space by **i**, **j**, **k**, this corresponds to a rotation about an axis perpendicular to $\mathbf{k}$.

3.5.2 Symmetric Matrices

For a symmetric matrix A, $A = A^T$

Choose $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ as eigenvectors of A, with eigenvalues $\lambda_1, \lambda_2, \lambda_3$.

$$A\mathbf{e}_k = \lambda_k \mathbf{e}_k \quad 2.48$$

$$\lambda_k (\mathbf{e}_k \cdot \mathbf{e}_j) = A\mathbf{e}_k \cdot \mathbf{e}_j \quad 2.49$$

$$= \mathbf{e}_k \cdot A^T \mathbf{e}_j$$

$$= \mathbf{e}_k \cdot A\mathbf{e}_j$$

$$= \lambda_j (\mathbf{e}_k \cdot \mathbf{e}_j).$$

This means that if $\lambda_j \neq \lambda_k$, then $\mathbf{e}_k \cdot \mathbf{e}_j = 0$

Choose e_1, e_2, e_3 to be unit vectors, then, $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$.

This means that we could represent a symmetric matrix as a diagonal matrix with only the entries $A_{ii} = \lambda_i$:

$$A = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \quad 2.50$$

This result is referred to as the **spectral representation** of a symmetric matrix.

3.5.3 Hermitian Matrices

The **Adjoint** (or **Hermitian conjugate**) of a matrix A is given by

$$Adj(A) = A^+ = \left((A)^T\right)* \quad 2.51$$

A Hermitian matrix is the complex equivalent of a real symmetric matrix, satisfying

$$A^+ = A \quad 2.52$$

3.5.4 Unitary Matrices

The complex analogue of a real orthogonal matrix is a **unitary** matrix, i.e., $AA^+ = I$ or, equivalently,

$$A^+ = A^{-1} \quad 2.53$$

3.5.5 Normal Matrices

A **normal** matrix is one that commutes with its Hermitian conjugate.

i.e.,

$$AA^+ = A^+ A \quad 2.54$$

4.0 Conclusion

This Unit introduced you to the concepts of orthogonality and orthonormality. They are so important in Quantum mechanics in that when in place, they guarantee that different vectors lie in specific directions that do not 'mix up' just the way the traditional unit vectors in 3-dimensional space do not 'mix up' when resolving them. You also came across the bra and ket or Dirac notation, another way of dealing with vectors and their inner products. Odd and even functions were brought in to make it easier for you to integrate functions within symmetric intervals. You also learnt about different types of matrices. With Gram-Schmidt orthonormalisation you have a way of creating an orthonormal set of vectors. With an orthonormal set, we can proceed to define the statistical probability with which a measurement of a physical quantity would result in a certain value. You also learnt about certain kinds of matrices.

5.0 Summary

- The inner product of a pair orthogonal vectors is zero.
- A basis that consists of orthogonal vectors only is an orthogonal basis.
- With an orthogonal basis, we can define the probabilities of measurement.
- The Gram- Schmidt orthonormalisation scheme can be used to create an orthogonal basis.

6.0 Learning Activity

1. Which of the following functions are even and which ones are odd?
   (i) $x^2 \sin x \cosh x$ (ii) $|e^x| \cosh 2x$ (iii) $\sec x$
2. Write the following as a sum of odd and even functions.
   (i) $e^{-x} \cosh x$ (ii) $x \ln x$
3. Evaluate the following integrals
   (i) $\int_{-a}^{a} x^{2n+1} dx$, $n = 0, 1, 2, ...$ (ii) $\int_{-a}^{a} x^{2n} dx$, $n = 0, 1, 2, ...$
4. Show that
   (i) $\sin mx$ and $\cos nx$ are orthogonal, $-\pi \leq x \leq \pi$.
   (ii) $\sin mx$ and $\sin nx$ are orthogonal, $m \neq n$, $-\pi \leq x \leq \pi$.
5. If the matrix $\begin{bmatrix} 3 & x \\ 1 & 2 \end{bmatrix}$ is a proper orthogonal matrix, find x.
6. If the matrix $\begin{bmatrix} y & i \\ -i & 2 \end{bmatrix}$ is Hermitian, find the value of y.

7.0 References/Further Readings

1. Mathematical Physics – Butkov, E.
2. Mathematical Methods for Physics and Engineering – Riley, K. F., Hobson, M. P. and Bence, S. J.

Solutions to Learning Activity

1. Which of the following functions are even and which ones are odd?

(i) $x^2 \sin x \cosh x$ (ii) $|e^x| \cosh 2x$ (iii) $\sec x$

(i) is odd, being the product of two even functions and an odd function.

(ii) is an even function, a product of two even function.

(iii) is an even function:

$$\sec(-x) = \frac{1}{\cos(-x)} = \frac{1}{\cos x} = \sec x$$

2. Write the following as a sum of odd and even functions.

(i) $e^{-x} \cosh x$ (ii) $x \ln x$

(i) $h(x) = e^{-x}\cosh x$, $h(-x) = e^x \cosh(-x) = e^x \cosh x$

$$f(x) = \frac{1}{2}[h(x) - h(-x)] = \frac{e^{-x}\cosh x - e^x \cosh x}{2} = -\cosh x\left[\frac{e^x - e^{-x}}{2}\right]$$

$$= -\cosh x \sinh x$$

$$g(x) = \frac{1}{2}[h(x) + h(-x)] = \frac{e^{-x}\cosh x + e^x \cosh x}{2} = \cosh x\left[\frac{e^x + e^{-x}}{2}\right]$$

$$= \cosh^2 x$$

3. Evaluate the following integrals

(i) $\int_{-a}^{a} x^{2n+1} dx$, $n = 0,1,2,\ldots.$ (ii) $\int_{-a}^{a} x^{2n} dx$, $n = 0,1,2,\ldots.$

(i) $\int_{-a}^{a} x^{2n+1} dx = 0$, the integrand being an odd function.

(ii) $\int_{-a}^{a} x^{2n} dx = 2\int_{0}^{a} x^{2n} dx = 2\left.\frac{x^{2n+1}}{2n+1}\right|_{0}^{a} = 2\frac{a^{2n+1}}{2n+1}$

4. Show that

(i) $\sin mx$ and $\cos nx$ are orthogonal, $-\pi \le x \le \pi$.

(ii) $\sin mx$ and $\sin nx$ are orthogonal, $m \ne n$, $-\pi \le x \le \pi$.

(i) $\int_{-\pi}^{\pi} \sin mx \cos nx dx = 0$, the integrand is an odd function

(ii) $\int_{-\pi}^{\pi} \sin mx \sin nx dx = \int_{-\pi}^{\pi} \frac{1}{2}[\cos(m-n)x - \cos(m+n)x]dx$

$$= \frac{1}{2}\left[\frac{1}{m-n}\sin(m-n)x\Big|_{-\pi}^{\pi} - \frac{1}{m+n}\sin(m+n)x\Big|_{-\pi}^{\pi}\right] = 0$$

5. If the matrix $\begin{bmatrix} 3 & x \\ 1 & 2 \end{bmatrix}$ is a proper orthogonal matrix, find x.

$$\det\begin{bmatrix}3 & x\\ 1 & 2\end{bmatrix} = 6 - x = 1, \text{ or } x = 5$$

6. If the matrix $\begin{bmatrix}y & i\\ -i & 2\end{bmatrix}$ is Hermitian, find the value of y.

The matrix is Hermitian if it is equal to its Hermitian adjoint, i.e.,

is $\left(\begin{bmatrix}y & i\\ -i & 2\end{bmatrix}^T\right)^*$ equal to $\begin{bmatrix}y & i\\ -i & 2\end{bmatrix}$

$$\left(\begin{bmatrix}y & i\\ -i & 2\end{bmatrix}^T\right)^* = \begin{bmatrix}y & -i\\ i & 2\end{bmatrix}^* = \begin{bmatrix}y & i\\ -i & 2\end{bmatrix}$$

The matrix is Hermitian.

UNIT 3: OPERATORS AND RELATED TOPICS

1.0 Introduction

Operators are quite important in Quantum mechanics because every observable is represented by a Hermitian operator. The eigenvalues of the operator are the possible values the physical observable can take, and the expectation value of the observable in any particular state is the average value it takes in that particular state. Commuting operators indicate that the corresponding physical observables can have the same eigenstates, or equivalently, they can both be measured simultaneously with infinite accuracy. You shall get to learn about all these in this Unit.

2.0 Objectives

At the end of this Unit, you should be able to do the following:

- Define a linear operator.
- Find the eigenvalues of a linear operator.
- Calculate the expectation value of a physical observable in a given state.
- Do commutator algebra.
- Find the matrix elements of a linear operator.
- Write the matrix for a change from one basis to another.

3.0 Main Content

3.1 Linear Operators

A linear map, or linear transformation or linear operator, is a function $f : X \to Y$ between vector spaces X and Y which preserves vector addition and scalar multiplication, i.e.,

$$f(x_1 + x_2) = f(x_1) + f(x_2)$$

$$f(\lambda x) = \lambda f(x) \text{ for } \lambda \in K \text{, a constant, and } x_1, x_2 \in X$$

Equivalently, $f(ax_1 + bx_2) = af(x_1) + bf(x_2)$.

As am example, the differential operator is a linear operator.

$$\frac{d}{dx}(\alpha\, f_1(x)+\beta\, f_2(x))=\alpha\frac{d}{dx}\, f_1(x)+\beta\frac{d}{dx}\, f_2(x)$$

where α and β are constants (scalars) in the underlying field.

3.1.1 Eigenvalues of a Linear Operator

Let A be an operator and λ the associated eigenvalue corresponding to an eigenvector ψ. Then, we can write

$$A\psi=\lambda\psi \qquad 3.1$$

Frequently, the operator A is a matrix, and the eigenvector ψ a column matrix. It follows that

$$(A-\lambda I)\psi=0 \qquad 3.2$$

where I is the appropriate identity matrix, that is, a square matrix that has 1 along its main diagonal and zero elsewhere.

For a non-trivial solution, we require that the determinant vanish, that is,

$$|A-\lambda I|=0 \qquad 3.3$$

Solving the resulting characteristic (or secular) equation, we obtain the possible values of λ, called the eigenvalues. Then armed with the eigenvalues, we can then obtain the associated eigenfunctions.

Example

Given the matrix $\begin{bmatrix}3 & -2\\ 1 & 2\end{bmatrix}$, find the corresponding eigenvectors and the eigenvalues.

Solution

Let the eigenvector be $u=\begin{bmatrix}u_1\\ u_2\end{bmatrix}$, and the corresponding eigenvalue be λ. Then,

$$\begin{bmatrix}3 & -2\\ 1 & 2\end{bmatrix}\begin{bmatrix}u_1\\ u_2\end{bmatrix}=\lambda\begin{bmatrix}u_1\\ u_2\end{bmatrix}$$

or

$$\left(\begin{bmatrix}3 & -2\\ 1 & 2\end{bmatrix}-\lambda\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\right)\begin{bmatrix}u_1\\ u_2\end{bmatrix}=0$$

which implies

$$\begin{vmatrix}3-\lambda & -2\\ 1 & 2-\lambda\end{vmatrix}=0$$

or $\lambda^2-5\lambda+6+2=0$

$$\lambda^2-5\lambda+8=0$$

$$\lambda = \frac{5+\sqrt{25-32}}{2} = \frac{5}{2} \pm \frac{i\sqrt{7}}{2}$$

Let $\lambda_1 = \frac{5}{2} + i\frac{\sqrt{7}}{2}$. Then, the corresponding eigenvector can be found:

$$\begin{bmatrix} 3 & -2 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \left(\frac{5}{2} + i\frac{\sqrt{7}}{2}\right)\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$3u_1 - 2u_2 = \left(\frac{5}{2} + i\frac{\sqrt{7}}{2}\right)u_1 \qquad \text{(i)}$$

$$u_1 + 2u_2 = \left(\frac{5}{2} + i\frac{\sqrt{7}}{2}\right)u_2$$

From (i), $2u_2 = 3u_1 - \left(\frac{5}{2} + i\frac{\sqrt{7}}{2}\right)u_1 = \left(\frac{1}{2} - i\frac{\sqrt{7}}{2}\right)u_1$

Thus, choosing $u_1 = 1$, we get $u_2 = \frac{1}{2}\left(\frac{1}{2} - i\frac{\sqrt{7}}{2}\right)$

Hence, an eigenfunction for the matrix is $\begin{pmatrix} 1 \\ 1-i\sqrt{7} \end{pmatrix}$

Similarly, choosing $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ as the other eigenvector with a corresponding eigenvalue $\lambda_2 = \frac{5}{2} - i\frac{\sqrt{7}}{2}$, we can get the eigenvector $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$.

Central to the theory of quantum mechanics is the idea of an operator (as we have seen earlier). We have indeed come across some operators. Recall

$$\hat{H}\psi(x) = E\psi(x) \qquad 3.4$$

where $\hat{H}$ is an operator. For the time-independent Schroedinger equation:

$$\hat{H}\psi(x) = -\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) \qquad 3.5$$

$\hat{H}$ is the *total energy operator* or *Hamiltonian*.

We identify some other operators:

(i) The kinetic energy operator $\hat{T}$

$$\hat{T} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} \qquad 3.6$$

(ii) The linear momentum operator $\hat{p}$

$$\hat{p} = -i\hbar \frac{d}{dx} \qquad 3.7$$

(iii) The position operator $\hat{x}$

$$\hat{x} = x \qquad 3.8$$

3.2 Expectation value

The expectation value of a quantity is the statistical predicted mean value of all measurements.

The (statistical) average value of the numbers x_1, x_2, …, x_n is $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$. However, if there is a distribution, such that there are f_i of the value x_i, $i =$ 1, 2, …, n, then the average becomes

$$\bar{x} = \frac{\sum_{i=1}^{m} f_i x_i}{\sum_{i=1}^{m} f_i} = \frac{1}{n}\sum_{i=1}^{n} f_i x_i \text{, since } \sum_{i=1}^{m} f_i = n \qquad 3.9$$

since n is the total number of observations.

In the case of quantum mechanics, the average value, or expectation value, of an operator is

$$< \cdot >= \int_{-\infty}^{\infty} \psi^*(x)(\cdot)\psi(x)dx \qquad 3.10$$

Thus, the expectation value of x is

$$< x >= \int_{-\infty}^{\infty} \psi^*(x)x\psi(x)dx \qquad 3.11$$

Thus, if $\psi = \sqrt{\frac{2}{L}}\sin\left(\frac{2\pi x}{L}\right)$, with $n = 2$, and $0 < x < L$,

$$< x > = \frac{2}{L}\int_0^L \sin\frac{2\pi x}{L} x \sin\frac{2\pi x}{L} dx \qquad 3.12$$

$$= \frac{2}{L}\int_0^L x\sin^2\frac{2\pi x}{L} dx$$

$$= \frac{2}{L} \times \frac{L^2}{4} = \frac{L}{2}$$

The expectation value of the momentum for the same case above is

$$< p >= \int_{-\infty}^{\infty} \psi^*(x)\left(-i\hbar\frac{d}{dx}\right)\psi(x)dx \qquad 3.13$$

$$= \frac{2i\hbar}{L}\int_0^L \sin\frac{2\pi x}{L}\frac{d}{dx}\left(\sin\frac{2\pi x}{L}\right)dx$$

$$= \frac{2i\hbar}{L} \times \frac{2\pi}{L} \int_0^L \sin\frac{2\pi x}{L} \cos\frac{2\pi x}{L}\, dx$$

$$= \frac{2i\hbar}{L} \times \frac{2\pi}{L} \int_0^L \frac{1}{2} \sin\frac{4\pi x}{L}\, dx$$

$$= \frac{2\pi i\hbar}{L^2} \times \frac{L}{4\pi} \cos\frac{4\pi x}{L}\Big|_0^L = 0$$

The energy expectation value of for the ground state of the simple harmonic oscillator:

$$< E >= \int_{-\infty}^{\infty} \psi_0^* \hat{H} \psi_0 dx = \int_{-\infty}^{\infty} \psi_0^* \left(\frac{1}{2}\hbar\omega_0\right)\psi_0 = \frac{1}{2}\hbar\omega_0 \int_{-\infty}^{\infty} \psi_0^* \psi_0 dx = \frac{1}{2}\hbar\omega_0 \qquad 3.14$$

since ψ_0 is normalised.

This is a special case of the general result

$$< \alpha >= \int_{-\infty}^{\infty} \psi * \hat{\alpha}\psi\, dx = \lambda \int_{-\infty}^{\infty} \psi * \psi dx = \lambda \qquad 3.15$$

Thus, we see that for any eigenstate of an operator, the expectation value of the observable represented by that operator is the eigenvalue.

More generally, we would write the expectation value of an operator, *A*, in a certain state ψ, as

$$< \psi \,|\, A \,|\, \psi > .$$

Example

The expectation value of a matrix operator, $\begin{pmatrix} 1 & 2 & -1 \\ 2 & -1 & 1 \\ 1 & 3 & 2 \end{pmatrix}$ in state $\begin{pmatrix} 2i \\ 1 \\ -1 \end{pmatrix}$ is

$$\begin{pmatrix} -2i & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 & -1 \\ 2 & -1 & 1 \\ 1 & 3 & 2 \end{pmatrix} \begin{pmatrix} 2i \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -2i & 1 & -1 \end{pmatrix} \begin{pmatrix} 2i+3 \\ 4i-2 \\ 2i+1 \end{pmatrix} = \begin{pmatrix} 4-6i \\ 4i-2 \\ -2i-1 \end{pmatrix}$$

3.3 Commutators and simultaneous eigenstates

Consider an operator $\hat{P}$ that represents a physical observable of a system, e.g., energy or momentum. Suppose that the state ψ has a particular value p of this observable, i.e., $\hat{P}\psi = p\psi$. Suppose further that the same state also has the value q of a second observable represented by the operator $\hat{Q}$, i.e., $\hat{Q}\psi = q\psi$. Then p and q are called simultaneous eigenvalues. Then,

$$\hat{Q}\hat{P}\psi = \hat{Q}p\psi = p\hat{Q}\psi = pq\psi \qquad 3.16$$

Similarly,

$$\hat{P}\hat{Q}\psi = \hat{P}q\psi = q\hat{P}\psi = qp\psi \qquad 3.17$$

Since p and q are just real numbers, then $qp = pq$. Thus, the condition for simultaneous eigenstates is that $\hat{P}\hat{Q}\psi = \hat{Q}\hat{P}\psi$ or

$$\hat{P}\hat{Q} - \hat{Q}\hat{P} = 0 \qquad 3.18$$

$\hat{P}\hat{Q} - \hat{Q}\hat{P}$ is said to be the *commutator* of $\hat{P}$ and $\hat{Q}$ and operators that satisfy the condition $\hat{P}\hat{Q} - \hat{Q}\hat{P} = 0$ are said to commute. The commutator is normally written $[\hat{P}, \hat{Q}]$.

Examples

1. Show that $[\hat{T}, \hat{p}]$.

$$[\hat{T}, \hat{p}]\psi = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\left(-i\hbar\frac{d}{dx}\right)\psi - \left(-i\hbar\frac{d}{dx}\right)\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\right)\psi$$

$$= \frac{i\hbar^3}{2m}\frac{d^3\psi}{dx^3} - \frac{i\hbar^3}{2m}\frac{d^3\psi}{dx^3} = 0$$

2. Calculate $[\hat{x}, \hat{p}]$.

$$[\hat{x}, \hat{p}]\psi = x\left(-i\hbar\frac{d}{dx}\right)\psi - \left(-i\hbar\frac{d}{dx}\right)(x\psi)$$

$$= -i\hbar x\frac{d\psi}{dx} + i\hbar\psi + i\hbar x\frac{d\psi}{dx}$$

$$= i\hbar\psi$$

Thus, we can write $[\hat{x}, \hat{p}] = i\hbar$

The fact that $\hat{x}$ and $\hat{p}$ do not commute lead to the uncertainty relation $\Delta x\Delta p \geq \hbar$.

Indeed, when two operators do not commute, it means that the two associated observables cannot be measured with infinite accuracy simultaneously. Thus, an attempt to measure the momentum of a particle with infinite accuracy will cause an infinite error in the position as is easily seen in the equation, $\Delta x = \dfrac{\hbar}{\Delta p}$. On the other hand, the momentum and the energy of such a system can be measured simultaneously with infinite accuracy. Other non-commutating operators include $\hat{E}$ and $\hat{t}$, i.e., the energy operator and the time operator.

The potential operator is just $\hat{V} = V$, just as $\hat{x} = x$.

3.4 Matrix Elements of a Linear Operator

We can represent any operator A by a square $n \times n$ matrix

$$A_{ij} = <\phi_i \,|\, A \,|\, \phi_j>, \; i, j = 1, n \tag{3.19}$$

Examples

1. For the identity operator I,

$$I \,|\, \phi_i> = |\phi_i>$$

$$A_{ij} = <\phi_i \,|\, A \,|\, \phi_j> = A_{ij} = <\phi_i \,|\, \phi_j> = \delta_{ij} \tag{3.20}$$

Hence,

$$I = \begin{pmatrix} 1 & 0 & . & . & 0 \\ 0 & 1 & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0. & . & . & 1 \end{pmatrix} \tag{3.21}$$

2. Consider the basis $B = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}$. Suppose we want to change to $C = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$. Then, the matrix of transformation is

$$A_{ij} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Note that $A_{ij} = \begin{pmatrix} <B_1 | C_1> & <B_1 | C_2> & <B_1 | C_3> \\ <B_2 | C_1> & <B_2 | C_2> & <B_2 | C_3> \\ <B_3 | C_1> & <B_3 | C_2> & <B_3 | C_3> \end{pmatrix}$

3.5 Change of Basis

The basis for a vector space is not unique. We can easily construct a linear map (matrix) that takes a basis vector in one basis to another, as seen in example 2 above. Let us consider R^n as a vector space.

Let $\{\mathbf{u}_i\}_{i=1}^{n}$ be a basis in the vector space. We can write any vector **a** in the vector space as $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ . \\ . \\ a_n \end{pmatrix}$. Then, we can write

$$\begin{pmatrix} a_1 \\ a_2 \\ . \\ . \\ a_n \end{pmatrix} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots + c_n\mathbf{u}_n = c_1\begin{pmatrix} u_{11} \\ u_{12} \\ . \\ . \\ u_{1n} \end{pmatrix} + c_2\begin{pmatrix} u_{21} \\ u_{22} \\ . \\ . \\ u_{2n} \end{pmatrix} + \ldots + c_n\begin{pmatrix} u_{n1} \\ u_{n2} \\ . \\ . \\ u_{nn} \end{pmatrix} \qquad 3.22$$

It follows that

$$a_1 = c_1u_{11} + c_2u_{21} + \ldots + c_nu_{n1}$$
$$a_2 = c_1u_{12} + c_2u_{22} + \ldots + c_nu_{n2}$$
$$.$$
$$.$$
$$a_n = c_1u_{1n} + c_2u_{2n} + \ldots + c_nu_{nn}$$

We can write this compactly as

$$\begin{pmatrix} a_1 \\ a_2 \\ . \\ . \\ a_n \end{pmatrix} = \begin{pmatrix} u_{11} & u_{21} & . & . & u_{n1} \\ u_{12} & u_{22} & . & . & u_{n2} \\ . & . & . & . & . \\ . & . & . & . & . \\ u_{1n} & u_{2n} & . & . & u_{nn} \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \\ . \\ . \\ c_n \end{pmatrix} \qquad 3.23$$

or

$$\begin{pmatrix} a_1 \\ a_2 \\ . \\ . \\ a_n \end{pmatrix} = B\begin{pmatrix} c_1 \\ c_2 \\ . \\ . \\ c_n \end{pmatrix} \qquad 3.24$$

where B is a matrix formed by arranging the vectors $\mathbf{u}_1$, $\mathbf{u}_2$, …, $\mathbf{u}_n$ in order.

It follows immediately that we can write

$$\begin{pmatrix} c_1 \\ c_2 \\ . \\ . \\ c_n \end{pmatrix} = B^{-1} \begin{pmatrix} a_1 \\ a_2 \\ . \\ . \\ a_n \end{pmatrix} \qquad 3.25$$

But we might as well have written **a** in another basis $\{\mathbf{v}_j\}_{j=1}^{n}$, as

$$\begin{pmatrix} a_1 \\ a_2 \\ . \\ . \\ a_n \end{pmatrix} = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \ldots + d_n\mathbf{v}_n = \begin{pmatrix} v_{11} & v_{21} & . & . & v_{n1} \\ v_{12} & v_{22} & . & . & v_{n2} \\ . & . & . & . & . \\ . & . & . & . & . \\ v_{1n} & v_{2n} & . & . & v_{nn} \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \\ . \\ . \\ d_n \end{pmatrix} = D \begin{pmatrix} d_1 \\ d_2 \\ . \\ . \\ d_n \end{pmatrix}$$

where $D = \begin{pmatrix} v_{11} & v_{21} & . & . & v_{n1} \\ v_{12} & v_{22} & . & . & v_{n2} \\ . & . & . & . & . \\ . & . & . & . & . \\ v_{1n} & v_{2n} & . & . & v_{nn} \end{pmatrix}$ 3.26

Therefore, equation 3.25 becomes

$$\begin{pmatrix} c_1 \\ c_2 \\ . \\ . \\ c_n \end{pmatrix} = B^{-1} \begin{pmatrix} a_1 \\ a_2 \\ . \\ . \\ a_n \end{pmatrix} = B^{-1}D \begin{pmatrix} d_1 \\ d_2 \\ . \\ . \\ d_n \end{pmatrix} \qquad 3.27$$

Conversely,

$$\begin{pmatrix} d_1 \\ d_2 \\ . \\ . \\ d_n \end{pmatrix} = D^{-1}B \begin{pmatrix} c_1 \\ c_2 \\ . \\ . \\ c_n \end{pmatrix} \qquad 3.28$$

Example 1

Given the basis {(2, 3), (1, 4)}, can we write the expression for a transformation to {(0, 2), (-1, 5)}?

Solution

$$B = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix},\ D = \begin{pmatrix} 0 & -1 \\ 2 & 5 \end{pmatrix},\ B^{-1} = \frac{1}{5}\begin{pmatrix} 4 & -1 \\ -3 & 2 \end{pmatrix},\ D^{-1} = \frac{1}{2}\begin{pmatrix} 5 & 1 \\ -2 & 0 \end{pmatrix}$$

$$B^{-1}D = \frac{1}{5}\begin{pmatrix} 4 & -1 \\ -3 & 2 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 2 & 5 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} -2 & -9 \\ 4 & 13 \end{pmatrix}$$

$$D^{-1}B = \frac{1}{2}\begin{pmatrix} 5 & 1 \\ -2 & 0 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 13 & 9 \\ -4 & -2 \end{pmatrix}$$

$$\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = D^{-1}B\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 13 & 9 \\ -4 & -2 \end{pmatrix}\begin{pmatrix} -2 \\ 6 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 28 \\ -4 \end{pmatrix} = \begin{pmatrix} 14 \\ -2 \end{pmatrix}$$

Check!

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = B^{-1}D\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} -2 & -9 \\ 4 & 13 \end{pmatrix}\begin{pmatrix} 14 \\ -2 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} -10 \\ 30 \end{pmatrix} = \begin{pmatrix} -2 \\ 6 \end{pmatrix}$$

4.0 Conclusion

Linear operators are so important in Quantum mechanics because every observable has an associated linear operator. So, we introduced you to linear operators, and then outlined how to get the eigenvalues and eigenvectors of a given linear operator. The eigenvalues are the possible values a measurement will yield, and the eigenstates are the possible states we can find the system. You also learnt about the expectation value of a physical observable represented by a linear operator. We then went on to discuss commutators and saw that simultaneous eigenstates are possible for a pair of operators if they commutate. You learnt, thereafter, to calculate the matrix elements of a linear operator. You might need to change from one set of basis to another. You also learnt how to do this, so that you might have a picture of what a vector in the space would look like in another basis.

5.0 Summary

- A linear operator is needed for each physically observable physical quantity in Quantum Mechanics.
- The eigenvalues of an operator are the possible values a measurement of the physical observable will yield.
- The eigenstates or eigenvectors of an operator are the possible states in which the system under consideration could be found.
- The matrix representing a linear operator can be determined.
- The basis for a certain vector space is not unique as we can construct more bases as may be needed.

6.0 Learning Activity

1a. Find the eigenvalues and the corresponding eigenfunctions of the matrix.

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

b. If this matrix represents a physically observable attribute of a particle, what is the expectation value of the attribute in each of the possible states. Comment on your results.

2. You are given the set $S_1 = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$.

(a) Are the linearly independent?\
(b) Are they orthogonal?
(c) Are they normalised? If not, normalise them.
(d) Write the vector $\begin{pmatrix} 3 \\ 4 \end{pmatrix}$
   (i) in terms of the usual basis in the Euclidean plane.
   (ii) In terms of the basis $S_U = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$.
(e) Write the matrix of transformation from basis S_U to basis S_1?

3. Find the matrix of transformation between the bases $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ and $\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$. Hence, express the vector $\begin{pmatrix} 3 \\ 4 \end{pmatrix}$ in the two different bases.

4. Write the matrix of transformation between the following bases in R^3, the 3-dimensional Euclidean plane.

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ 5 \end{pmatrix} \right\} \text{ and } \left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \right\}$$

7.0 References/Further Readings

1. Mathematical Physics – Butkov, E.
2. Mathematical Methods for Physics and Engineering – Riley, K. F., Hobson, M. P. and Bence, S. J.

Solutions to Learning Activity

1a. Find the eigenvalues and the corresponding eigenfunctions of the matrix.

$$A=\begin{bmatrix}0&0&1\\0&0&0\\1&0&0\end{bmatrix}$$

b. If this matrix represents a physically observable attribute of a particle, what is the expectation value of the attribute in each of the possible states. Comment on your results.

a. The characteristic equation is formed by $\begin{vmatrix}-\lambda&0&1\\0&-\lambda&0\\1&0&-\lambda\end{vmatrix}=0$

$$-\lambda^3+\lambda=0$$

Eigenvalues are 0, 1 and -1.

For $\lambda = 0$, eigenvector is given by $\begin{bmatrix}0&0&1\\0&0&0\\1&0&0\end{bmatrix}\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}$

Or $\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}=\begin{bmatrix}0\\1\\0\end{bmatrix}$

The normalised eigenfunction is $\psi_1=\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}=\begin{bmatrix}0\\1\\0\end{bmatrix}$

$$\lambda=1:\quad\begin{bmatrix}-1&0&1\\0&-1&0\\1&0&-1\end{bmatrix}\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

$$\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}=\begin{bmatrix}1\\0\\1\end{bmatrix}$$

The normalised wavefunction is $\psi_2=\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}=\frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\1\end{bmatrix}$

$$\lambda=-1:\begin{bmatrix}1&0&1\\0&1&0\\1&0&1\end{bmatrix}\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Normalised wavefunction is $\psi_3 = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

b. The expectation value of A in state $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ is

$$<\psi_1 | A | \psi_1 >= \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = 0$$

$$<\psi_2 | A | \psi_2 >= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{2}\times 2 = 1$$

$$<\psi_3 | A | \psi_3 >= \frac{1}{\sqrt{2}}\begin{bmatrix} -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \frac{1}{2}\times -2 = -1$$

Comment: The expectation values are the eigenvalues we got earlier. This is another way of getting the eigenvalues of an operator.

2. You are given the set $S_1 = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$.

(f) Are the linearly independent?\

(g) Are they orthogonal?

(h) Are they normalised? If not, normalise them.

(i) Write the vector $\begin{pmatrix} 3 \\ 4 \end{pmatrix}$

  (i) in terms of the usual basis in the Euclidean plane.

  (ii) In terms of the basis $S_U = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$.

(j) Write the matrix of transformation from basis S_U to basis S_1?

Solution

(a) Given a set $\{\mathbf{v}_i\}_{i=1}^{n}$, if we can write $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n$ = 0 and this implies $a_1 = a_2 = \cdots = a_n$ = 0, then we say $\{\mathbf{v}_i\}_{i=1}^{n}$ is a linearly independent set.

$$\mathbf{a}= \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \ \mathbf{b}= \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

To check if they are linearly independent.

$$c_1\begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Hence, $c_1 + c_2 = 0$ and $c_1 - c_2 = 0$. From the last equation, $c_1 = c_2$. Putting this in the first equation, $c_1 + c_1 = 0$, or c_1 = 0. Consequently, c_2 = 0. Set is linearly independent.

(b) To check orthogonality, $(\mathbf{a},\mathbf{b}) = \mathbf{a}^T\mathbf{b} = (1 \quad 1)\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 1 - 1 = 0$

(They are orthogonal)

(c) Are they normalised? $(\mathbf{a},\mathbf{a}) = (1 \quad 1)\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2$, or $\| a \| = 2$.

$$(\mathbf{b},\mathbf{b}) = (1 \quad -1)\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 2.$$

They are not normalised.

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \ \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ are normalised.}$$

The set $\{ \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \}$ forms an orthonormal basis for R^2.

In the usual basis S_U,

$$\begin{pmatrix} 3 \\ 4 \end{pmatrix} = 3\mathbf{i} + 4\mathbf{j} = 3\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 4\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

In the basis S_1

$$\begin{pmatrix} 3 \\ 4 \end{pmatrix} = \frac{\alpha}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{\beta}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Hence, $\alpha + \beta = 3\sqrt{2}$ and $\alpha - \beta = 4\sqrt{2}$

$2\alpha = 7\sqrt{2}$ and $2\beta = -\sqrt{2}$

Therefore,

$$\begin{pmatrix} 3 \\ 4 \end{pmatrix} = \frac{7\sqrt{2}}{2\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \frac{\sqrt{2}}{2\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{7}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

3. Find the matrix of transformation between the bases $\left\{\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\right\}$ and $\left\{\begin{pmatrix}1\\1\end{pmatrix},\begin{pmatrix}1\\-1\end{pmatrix}\right\}$. Hence, express the vector $\begin{pmatrix}3\\4\end{pmatrix}$ in the two different bases.

The matrix from basis S_U is $B=\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$, and the matrix from basis S_1 is

$D=\begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}$, $B^{-1}=\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$, $D^{-1}=\frac{1}{-2}\begin{bmatrix}-1 & -1\\-1 & 1\end{bmatrix}=\frac{1}{2}\begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}$

The matrix of transformation from S_U to S_1 is $B^{-1}D=D=\begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}$.

The matrix of transformation from S_1 to S_U is $D^{-1}B=D^{-1}=\frac{1}{2}\begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}$

So, $\begin{bmatrix}3\\4\end{bmatrix}$ in S_U transforms to $D^{-1}B\begin{bmatrix}3\\4\end{bmatrix}=\frac{1}{2}\begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}\begin{bmatrix}3\\4\end{bmatrix}=\frac{1}{2}\begin{bmatrix}7\\-1\end{bmatrix}=\begin{bmatrix}7/2\\-2\end{bmatrix}$ in S_1.

Crosscheck! Does this transform into $\begin{bmatrix}3\\4\end{bmatrix}$ the other way?

$\begin{bmatrix}7/2\\-2\end{bmatrix}$ in S_1 transforms to $B^{-1}D\left(\frac{1}{2}\begin{bmatrix}7\\-1\end{bmatrix}\right)=\frac{1}{2}\begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}\begin{bmatrix}7\\-1\end{bmatrix}=\frac{1}{2}\begin{bmatrix}6\\8\end{bmatrix}=\begin{bmatrix}3\\4\end{bmatrix}$ in S_U.

4. Write the matrix of transformation between the following bases in R^3, the 3-dimensional Euclidean plane.

$$\left\{\begin{pmatrix}1\\0\\2\end{pmatrix}\begin{pmatrix}2\\1\\0\end{pmatrix},\begin{pmatrix}0\\3\\5\end{pmatrix}\right\} \text{ and } \left\{\begin{pmatrix}1\\2\\1\end{pmatrix}\begin{pmatrix}2\\-1\\1\end{pmatrix},\begin{pmatrix}1\\1\\-1\end{pmatrix}\right\}$$

Let $S_a=\left\{\begin{pmatrix}1\\0\\2\end{pmatrix}\begin{pmatrix}2\\1\\0\end{pmatrix},\begin{pmatrix}0\\3\\5\end{pmatrix}\right\}$, and $S_b=\left\{\begin{pmatrix}1\\2\\1\end{pmatrix}\begin{pmatrix}2\\-1\\1\end{pmatrix},\begin{pmatrix}1\\1\\-1\end{pmatrix}\right\}$

The matrix related to S_a is $B=\begin{bmatrix}1 & 2 & 0\\0 & 1 & 3\\2 & 0 & 5\end{bmatrix}$, while the one related to S_b is $\begin{bmatrix}1 & 2 & 1\\2 & -1 & 1\\1 & 1 & -1\end{bmatrix}$.

We need to get the inverse of D, since we need $D^{-1}B$. The inverse of a matrix is the matrix of cofactors divided by the determinant. First, we evaluate the determinant of D.

Determinant of D is

$$1(1-1)-2(-2-1)+1(2+1)=9$$

The inverse of D is the transpose of the matrix of cofactors divided by the determinant:

$$D^{-1}=\frac{1}{9}\begin{bmatrix} \begin{vmatrix}-1 & 1\\ 1 & -1\end{vmatrix} & -\begin{vmatrix}2 & 1\\ 1 & -1\end{vmatrix} & \begin{vmatrix}2 & -1\\ 1 & 1\end{vmatrix} \\ -\begin{vmatrix}2 & 1\\ 1 & -1\end{vmatrix} & \begin{vmatrix}1 & 1\\ 1 & -1\end{vmatrix} & -\begin{vmatrix}1 & 2\\ 1 & 1\end{vmatrix} \\ \begin{vmatrix}2 & 1\\ -1 & 1\end{vmatrix} & -\begin{vmatrix}1 & 1\\ 2 & 1\end{vmatrix} & \begin{vmatrix}1 & 2\\ 2 & -1\end{vmatrix} \end{bmatrix}^T =\frac{1}{9}\begin{bmatrix}0 & 3 & 3\\ 3 & -2 & 1\\ 3 & 1 & -5\end{bmatrix}^T=\frac{1}{9}\begin{bmatrix}0 & 3 & 3\\ 3 & -2 & 1\\ 3 & 1 & -5\end{bmatrix}$$

$$\frac{1}{9}\begin{bmatrix}1 & 2 & 1\\ 2 & -1 & 1\\ 1 & 1 & -1\end{bmatrix}\begin{bmatrix}0 & 3 & 3\\ 3 & -2 & 1\\ 3 & 1 & -5\end{bmatrix}=\frac{1}{9}\begin{bmatrix}9 & 0 & 0\\ 0 & 9 & 0\\ 0 & 0 & 9\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

We have got the inverse right, $DD^{-1}=I$. The matrix of transformation from of transformation from to S_b is, and that of transformation from S_a to S_b is

$$D^{-1}B=\frac{1}{9}\begin{bmatrix}0 & 3 & 3\\ 3 & -2 & 1\\ 3 & 1 & -5\end{bmatrix}\begin{bmatrix}1 & 2 & 0\\ 0 & 1 & 3\\ 2 & 0 & 5\end{bmatrix}=\frac{1}{9}\begin{bmatrix}6 & 3 & 24\\ 5 & 4 & -1\\ -7 & 7 & -12\end{bmatrix}$$

2.6 Schrodinger's Equations

Schrodinger derived a mathematical equation to describe the dual nature of mater waves. Newton's law of motion is applied only to macroscopic particles. But this equation is used for both macroscopic and microscopic particles. The equation in mathematical form is known as Schrodinger's wave equation. There are two forms of Schrodinger wave equations.

(i) Time independent wave equation
(ii) Time dependent wave equation

Schrodinger time independent wave equation

Schrodinger derived the equations of matter waves by making the following assumptions.

(i) deBroglie wavelength can be applied for any particle moving in any field of force.

The total energy for a particle 'E' = P.E + K.E

$$E = V + \frac{1}{2}mv^2$$

$$E - V = \frac{1}{2}mv^2$$

$$2(E - V) = mv^2$$

Multiply by 'm' on both sides, we get

$$2m\,(E\text{-}V) = mv^2 \text{ --------------- } (1)$$

$$p^2 = 2m\,(E\text{-}V) \qquad \because\ p = mv$$

$$p = [2m(E - V)]^{1/2}$$

deBroglie wavelength $\lambda = h/p$

$$\lambda = \frac{h}{[2m(E - V)]^{1/2}}$$

(ii) The wave function associated with the material particles in time‘t’ can be written as,

$$\psi = \psi_0 e^{-iwt} \text{ --------------- (2)}$$

Where, ψ is the amplitude of the wave function

$\omega = 2\pi\nu$ where ν is the frequency of radiation

Consider a particle of mass ‘m’ moving with a velocity ‘v’ whose wave length is given by $\lambda = \frac{h}{mv}$. Let us find the wave nature of the system.

The classical differential equation to this wave system is

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} = \frac{1}{v^2} \cdot \frac{\partial^2\psi}{\partial t^2}$$

here, x, y, z → co-ordinates of the particle.

ψ → wave displacement for the de Broglie wave at any time ‘t’.

V → wave velocity

The above equation can be written as,

$$\nabla^2\psi = \frac{1}{v^2}\frac{\partial^2\psi}{\partial t^2} \text{ ----------------- (3)}$$

where,

$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ is the Laplacian operator.

The solution for the equation (3) is,

$$\psi_{(x,y,z)} = \psi_{0(x,y,z)} e^{-i\omega t} \text{ ---------------- } (4)$$

Differentiating equation (4) twice with respect to time we get,

$$\frac{\partial \psi}{\partial t} = \psi_0 (-i\omega) e^{-i\omega t}$$
$$\frac{\partial^2 \psi}{\partial t^2} = \psi_0 (-i\omega)(-i\omega) e^{-i\omega t}$$
$$\frac{\partial^2 \psi}{\partial t^2} = -\omega^2 \psi_0 e^{-i\omega t}$$
$$\frac{\partial^2 \psi}{\partial t^2} = -\omega^2 \psi \text{ ------------------------- } (5)$$

$$\left[\because i^2 = -1\right]$$
$$\left[\because \psi = \psi_0 e^{-i\omega t}\right]$$

For equations (3) and (5),

$$\nabla^2 = \frac{-\omega^2}{v^2} \psi \text{ -------------------- } (6)$$

But

$$\omega = 2\pi\nu = 2\pi\left(\frac{v}{\lambda}\right)$$

or

$$\frac{\omega}{v} = \frac{2\pi}{\lambda} \text{ ---------------- } (7)$$

Substituting equation (7) in equation (6) we get

$$\nabla^2 \psi = \frac{-4\pi^2}{\lambda^2} \partial \psi$$
$$i.e., \nabla^2 \psi + \frac{4\pi^2}{\lambda^2} \psi = 0 \text{ ---------------------- } (8)$$

Equation (8) is a general equation. For introducing the wave nature, substitute the value $\lambda = \frac{h}{mv}$ in the above equation.

i.e.,

$$\nabla^2\psi + \frac{4\pi^2}{\left(h^2 / m^2v^2\right)}\psi = 0$$

$$\nabla^2\psi + \frac{4\pi^2 m^2 v^2}{h^2}\psi = 0 ------------- (9)$$

Substituting the value of m^2v^2 from equation (1) in equation (9)

$$\nabla^2\psi + 4\pi^2 \frac{2m(E-V)}{h^2}\psi = 0 ------------ (10)$$

$$\nabla^2\psi + \frac{8\pi^2 m}{h^2}(E-V)\psi = 0 ------------- (11)$$

This equation is known as Schrodinger's time independent wave equation.

For one dimensional motion, the above equation becomes,

$$\frac{\partial^2\psi}{\partial x^2} + \frac{8\pi^2 m}{h^2}(E-V)\psi = 0 ----------- (12)$$

Put $\hbar = \frac{h}{2\pi}$

Therefore equation (12) can be written as,

$$\boxed{\nabla^2\psi + \frac{2m}{\hbar^2}(E-V)\psi = 0}$$

The above equation is the Schrodinger time independent wave equation in three dimension.

For a free particle, the potential energy V = 0

$\therefore$ Schrodinger wave equation for a free particle can be written as,

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

<u>Schrodinger time dependent wave equation</u>

When we eliminate 'E' from the Schrodinger time independent wave equation, we can easily get the Schrodinger time dependent wave equation. From Schrodinger second assumption,

$$\psi_{(x,y,z)} = \psi_{0(x,y,z)} e^{-i\omega t} \qquad \text{----------------} \; (1)$$

Differentiating equation (1) with respect to time

$$\frac{\partial\psi}{\partial t} = \psi_0(-i\omega)e^{-i\omega t}$$

Put $\omega = 2\pi\nu$ hence, $\frac{\partial\psi}{\partial t} = -i(2\pi\nu)\psi_0 e^{-i\omega t}$

$$\frac{\partial\psi}{\partial t} = -2\pi i\nu\psi$$

$$\frac{\partial\psi}{\partial t} = -2\pi i\left(\frac{E}{h}\right)\psi \qquad \left[\because \nu = \frac{E}{h}\right]$$

$$\frac{\partial\psi}{\partial t} = -i\frac{E}{h/2\pi}\psi$$

$$\frac{\partial\psi}{\partial t} = -i\frac{E}{\hbar}\psi \qquad \left[\because \hbar = \frac{h}{2\pi}\right]$$

$$E\psi = -\frac{\hbar}{i}\frac{\partial\psi}{\partial t}$$

$$E\psi = i\hbar\frac{\partial\psi}{\partial t}$$

Substituting the value of $E\psi$ in Schrodinger's time independent wave equation i.e.,

$$\nabla^2\psi + \frac{2m}{\hbar^2}(E - V)\psi = 0$$

$$\nabla^2\psi + \frac{2m}{\hbar^2}\left[i\hbar\frac{\partial\psi}{\partial t} - V\psi\right] = 0$$

$$\nabla^2\psi = -\frac{2m}{\hbar^2}\left[i\hbar\frac{\partial\psi}{\partial t} - V\psi\right]$$

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V\psi = i\hbar\frac{\partial\psi}{\partial t}$$

$$\boxed{\left[-\frac{\hbar^2}{2m}\nabla^2 + V\right]\psi = i\hbar\frac{\partial\psi}{\partial t}}$$

The above equation is Schrodinger time dependent wave equation

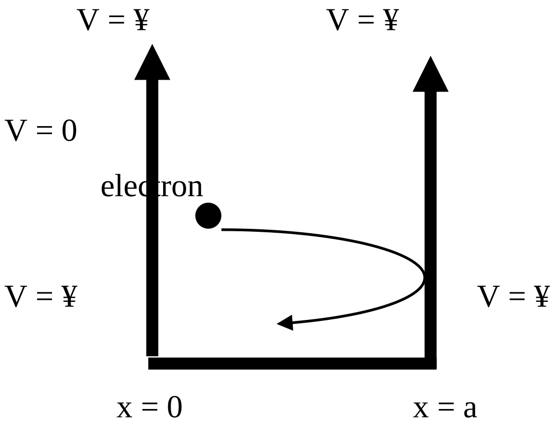
$$H\psi = E\psi$$

Where, $H = \left[-\frac{\hbar^2}{2m}\nabla^2 + V\right]$ is the Hamiltonian operator and

$E = i\hbar\frac{\partial\psi}{\partial t}$ is the energy operator.

2.7 *Application of Schrodinger wave equation using one dimensional potential well*

Particle in one dimensional box



Consider a particle moving inside a box to and fro between two walls at $x = 0$ and $x = a$ along the x-axis direction. When it collides with the wall there is no loss of energy of the electron, hence we can say the collisions are perfectly elastic. So there is no change in potential energy 'V'.

Hence, we can say the value of potential energy inside the box is zero. Here the walls are infinitely high so the particle cannot get out of the box; hence we can say outside to box the potential energy of the particle is infinite.

Therefore, inside the well,

P.E $V = 0$ for $0 < x < a$ and
P.E $V = \alpha$ for $0 \geq x \geq a$

The particle cannot exist outside the box, so the wave function

$$\psi = 0 \quad \text{for} \quad 0 \geq x \geq a$$

The Schrodinger one dimensional time independent equation for a free particle is,

$$\frac{\partial^2\psi}{\partial x^2} + \frac{2mE}{\hbar^2}\psi = 0$$

$$\frac{\partial^2\psi}{\partial x^2} + k^2\psi = 0 \quad \text{-------------} \ (1)$$

Where, $$k^2 = \frac{2mE}{\hbar^2}$$

The general solution for equation (1) is

ψ = A sin kx + B cos kx ---------------- (2)

For evaluating the constants A and B, use boundary conditions

x = 0 and x = a in equation (2)

ψ = 0 at x = 0, i.e., 0 = A sin k (0) + B cos(0)

i.e., B = 0

ψ = 0 at x = a, i.e., 0 = A sin k (a) + 0

i.e., A sin ka = 0

Since the electron is present inside the box, A $\neq$ 0

So we can write, sin (ka) = 0

sin (ka) = 0 only when ka = n π

k = n π/a

$$k = \frac{n\pi}{a} \text{-------------}(3)$$

But from (1) $$k^2 = \frac{2mE}{\hbar^2} \text{-------------}(4)$$

Equating the equations (3) and (4),

$$\frac{2mE}{\hbar^2} = \frac{n^2\pi^2}{a^2}$$

∴ The energy of the particle

$$E_n = \frac{n^2\pi^2\hbar^2}{2ma^2}$$

$$E_n = \frac{n^2h^2}{8ma^2}$$

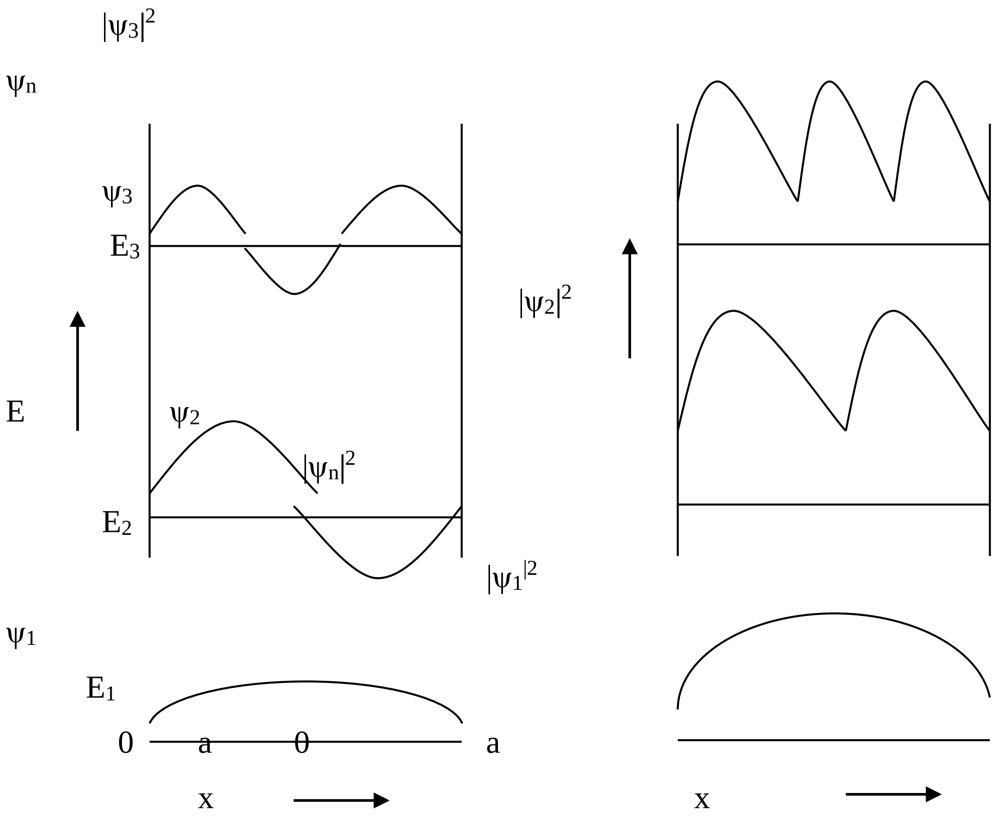
$$\left[\because \hbar = \frac{h}{2\pi}\right]$$

and the wave equation

$$\psi_n = A\sin\left(\frac{n\pi x}{a}\right)$$

Each value of E_n is called Eigen value and the corresponding ψ_n is called Eigen function.

A few energy level diagrams are shown in the fig.



Normalization of the wave function

It is the process by which the probability of finding the particle inside the box can be done.

Let us find the value of A. Since the electron is present inside the box

$$\int_{-\alpha}^{+\alpha} |\psi_n|^2 dx = 1$$

Substituting the value of ψ_n in the above equation we get,

$$\int_0^a A^2 \sin^2\left(\frac{n\pi x}{a}\right) dx = 1$$

$$A^2 \int_0^a \left(\frac{1-\cos\left(\frac{2n\pi x}{a}\right)}{2}\right) dx = 1$$

$$\frac{A^2}{2}\left(\int_0^a dx - \int_0^a \cos\left(\frac{2n\pi x}{a}\right) dx\right) = 1$$

$$\frac{A^2}{2}\left[[x]_0^a - \left[\frac{\sin\left(\frac{2n\pi x}{a}\right)}{\frac{2n\pi}{a}}\right]_0^a\right] = 1$$

$$\frac{A^2}{2}\left[a - \frac{\sin 2n\pi}{\frac{2n\pi}{a}}\right] = 1$$

The second term of the above is zero

$$\therefore \frac{A^2 a}{2} = 1$$

$$A^2 = \frac{2}{a}$$

$$A = \sqrt{\frac{2}{a}}$$

$$\therefore \psi_n = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}\right)x$$

The above equation is called normalized wave function.

LEARNING ACTIVITY

1. **Derive the expression for Schrodinger equation**
2. **Derive the expression for Normalized wave function**
3. **What are the two assumptions of Schrodinger.**

Note:

a) Write your answer in the space given below.

b) Check the answer with your academic counsellor.

SUMMARY

In this chapter, we studied the basic of the wave mechanics and quantum operator also studied about properties of wave function and schordinger's equations and its applications.

Block III

Nuclear Physics

STRUCTURE

OVERVIEW

In this chapter on `Nuclear Physics', we shall study the classification of Nucleus and general properties of nucleus. We shall also study the Term Binding energy, Packing fraction and stability of nucleus and detail study about Liquid drop model and shell model.

LEARNING OBJECTIVES

After completing this Unit, you should be able to:

- Classification and General properties of Nucleus
- Binding energy and packing fraction of nucleus.
- Formation of nuclear models.

3.1 Introduction

The x-ray scattering experiments of Rutherford established the existence of the atomic nucleus. We know make a general review of the constituents and general properties of nuclei. Like the atom, the nucleus may also exist in different quantum states based on energies and angular momenta. The state corresponding to the lowest energy is the ground state in which a nucleus normally exists. We shall also assume that they are in the ground states and properties in such states are known as static characteristics of the neclei. The dynamic characteristics are shown during nuclear reactions, nuclear excitations and nuclear decays. The important static properties are: nuclear charge, nuclear mass, size, shape, binding energy, angular momentum, magnetic moment, statistics, etc.

Constituents of nuclei

With the discovery of neutron by Chadwick, it was recognised that the atomic nuclei are composed of two types of elementary particles – protons and neutrons. Collectively, they are called nucliens. A particular type or a species of nucleus, is called a nuclide.

The proton is identified as the nucleus of the lightest isotope of hydrogen. It carries one electronic but positive charge, +e and has a mass m_p of about 1836 times the electronic mass, m_e. The neutron possesses no charge – electrically neutral and hence the name. Its mass m_n is almost equal to, but slightly more than, the mass of a proton i.e., $m_n \simeq m_p$. The negatively charged to electrons move around the nucleus and the orbits or the effects of the electrons extend to about 10^{-10} m, the radius of the atom. The radius of a typical nucleus however is much smaller, about 10^{-14} m.

The atom is thus electrically neutral with a massive nucleus. The atomic mass distribution is such that most of its space is empty. The clouds of extra-nuclear electrons again are arranged in layers or shells. The outermost or valence electrons dictate the chemical properties of the atom on which the effect of the nucleus is negligible.

By Coulomb's law, the positively charged protons, closely spaced within the nucleus, should repel each other strongly to fly apart. To explain the nuclear stability, one must thus assume that nucleons are held, by a very strong short range attractive force, different from gravitation or electromagnetic force and is called strong interaction.

The number of protons in the nucleus determines the atomic number of the nuclide. It is also called the Z-value or the proton number. The number of neutrons in the nucleus is the neutron number, N. The sum of the numbers of protons (Z) and neutrons (N) inside the nucleus, that is, the total number of nucleons in a nucleus, is known as its mass number A.

$$\therefore \quad \boxed{A = N + Z} \quad \Rightarrow \quad \boxed{N = A - Z}$$

Obviously, the mass number A is an integer just as both N and Z are. For a given Z, a change in neutron number N causes a change in the nuclear mass.

A nucleus of an atom X of atomic number Z and a mass number A, that is, a nuclide is symbolically represented by ${}^{A}_{Z}X$. For instance, ${}^{4}_{2}He$ is the nucleus of a helium atom of atomic number 2 and mass number 4 (an α-particle). The proton and neutron are symbolically written as *p* and *n*.

3.2 Classification of Nuclei

Isotopes, isobars, isotones and mirror nuclei

Isotopes:

Nuclei with the same atomic number Z (i.e., same element), but different mass number A are called isotopes. Thus, a given element of atomic number Z may have isotopes of different mass numbers. Their nuclei have the same number of protons but different number of neutrons. Thus ^{6}Li and ^{7}Li are two stable isotopes of lithium. Protons in the nucleus of each is 3 but their neutron numbers are 3 and 4. Similarly, $^{11}_{6}C, ^{12}_{6}C, ^{13}_{6}C, ^{14}_{6}C, ^{15}_{6}C, ^{16}_{6}C,$ are all isotopes of carbon.

Isotopes were first discovered in naturally radioactive elements. Thomson, while investigating positive rays by parabola method, found two stable isotopes of Neon-^{20}Ne and ^{22}Ne. Elements having stable isotopes contain in natural states a mixture of the isotopes in almost constant proportions – the so – called relative abundances (sec Appendix 2). Hydrogen has two stable isotopes ^{1}H and ^{2}H of relative abundances 99.99% and 0.01% respectively. The isotope ^{2}H is called heavy hydrogen or deuterium (^{2}D) and its nucleus deuteron. Another isotope of hydrogen, which is unstable, ^{3}H is called tritium and its nucleus triton. Some elements like As, F, I and Au have a single isotope only. All the isotopes of an element have identical chemical properties but they slightly differ in physical properties. Isotopes that do not occur naturally can be produced by nuclear reactions.

Isobars:

Nuclei with the same mass number A but different atomic number Z (i.e., different elements) are called isobars. For instance, the two nuclides $^{16}_{8}O$ and $^{16}_{7}N$ are isobaric to each other. Both have the same mass number A = 16 but Z = 8 for $^{16}_{8}O$ and Z = 7 for $^{16}_{7}N$.

Isotones:

Nuclei with the same number of neutrons, that is, having the same N (no matter what the Z-value is), are known as istones. The nuclides $^{23}_{11}Na$ and $^{24}_{12}Mg$ are isotonic to each other, since both have 23-11 = 12 or 24 – 12 = 12 neutrons in the nucleus.

Mirror nuclei :

The pairs of isobaric nuclei where the proton number Z and the neutron number N are interchanged and differ by one unit are known as mirror nuclei. Examples are ($^{11}_{6}C$ and $^{11}_{5}B$), ($^{13}_{7}N$ and $^{13}_{6}C$), etc. Their mass number is A = 2Z – 1 where Z is the higher atomic number of the members of the pair, the other member of the pair has the atomic number Z – 1. The first member is usually β^+ active and changes into the second by β^+ transformation.

3.3 General Properties of Nucleus

Nuclear size

Nuclear radius: It is difficult to define the size of nuclei since their spatial extent is 'fuzzy'. But experiments indicate that the majority of atomic nuclei are spherical or nearly so, in shape. So, its volume is proportional to the total number of nucleons in it or its mass number A.

$$\therefore \frac{4}{3}\pi R^3 \; \alpha \; A$$

$$\text{or,} \quad R \; \alpha \; A^{\frac{1}{3}}$$

$$\boxed{R = r_o A^{\frac{1}{3}}}$$

where R is the radius of the nucleus and ro, a constant, called nuclear radius to parameter. The value of ro ranges from (1.1 to 1.5) x 10^{-15} m, i.e (1.1 – 1.5) fm and can be evaluated by a number of different methods.

- The nuclear radius, as discussed, is the radius of nuclear mass distribution. One may as well talk of the radius of nuclear charge distribution. Since nuclear charge parameter Z α A i.e., is almost linearly proportional to the mass number, and the nuclear charge density pc is nearly the same throughout the nuclear volume, the distribution of nuclear charge + Ze follows the pattern of nuclear mass distribution. Hence charge radius $\simeq$ mass radius, of the nucleus.

- The size of the nucleus was first estimated from Rutherford's x-ray scattering using various atoms. The larger the angle of scattering of α −particle, closer is its approach to the nucleus. If the kinetic energy of α −particle be equal to the repulsive Conlomb energy between α −particle (of charge 2e, mass m and velocity v) and the nucleus (of charge Ze) it would come momentarily to rest such that the distance of closet approach R is given by

$$\frac{1}{2}mv^2 = \frac{2e \times Ze}{4\pi\varepsilon 0R} \Rightarrow \mathrm{R}= \frac{Ze^2}{\pi\varepsilon 0m\, v^2}$$

Substituting for m = 6.64 $\times 10^{-27}$ kg, e = 1.6 $\times 10^{-19}$C and v=10^7 m/s, the value of R for Z = 20 is $\sim$ 1.5 $\times 10^{-14}$m.

The mean squared radius $\langle r^2 \rangle$ of nuclear charge distribution is given by

$$\langle r^2 \rangle = \frac{\int_0^R r^2.4\pi r^2 p(r)dr}{\int_0^R 4\pi r^2 p(r)\, dr} = \frac{\int_0^R r^4\, dr}{\int_0^R r^2\, dr} = \frac{3}{5}\, R^2$$

For a uniformly charged sphere of radius R, ρ = charge density = constant and ρ = 0 for r > R.

$$\boxed{\therefore\ R^2 = \frac{5}{3}\, \langle r^2 \rangle}$$

Estimation of nuclear radius Of the different methods generally used to estimate the nuclear radius, we shall describe here only two of them: (i) the mirror nuclei method and (ii) the muonic x-ray method.

Mirror nuclei method – As already defined, the mirror are pairs of nuclei obtained from each other by the interchange of neutron with proton, e.g. ($^{13}_{6}C$, $^{13}_{7}N$). A mirror nucleus like $^{13}_{7}N$ is unstable and is converted to $^{13}_{6}C$ by positron (β^+) emission.

$$^{13}_{7}N \rightarrow {}^{13}_{6}C + v$$

Where v is the neutrino, a neutral particle of negligible mass.

If Z be the atomic number of the daughter nucleus, then it can be shown that the difference in Coulomb energies of mirror nucleus is

$$\frac{1}{4\pi\varepsilon 0}\left(\frac{6Ze^2}{5R}\right)$$

Where R = the radius of the daughter nucleus.

The above energy is spent is providing (i) the rest mass energy m_ec^2 to produce positron, (ii) the kinetic energy $E_{\beta+}$ of the β^+ particle and (iii) the rest mass $(m_n - m_p)c^2$ required for conversion of a neutron of larger mass m_n into a proton of small mass m_p.

$$\therefore \frac{1}{4\pi\varepsilon 0}\left(\frac{6Ze^2}{5R}\right) = m_ec^2 + E_{\beta+} + (m_n - m_p)c^2$$

Whence R could be readily estimated.

The average value of R estimated by this method is $(1.23 \pm 0.03)\, A^{\frac{1}{3}} \times 10^{-15}$ m.

Muonic *x*-rays – The energies of x-rays emitted by muons are called muonic *x*-rays. When a beam of muons, whose mass in 207 m_e and charge equal to e, is incident on nuclei like graphite, the muons move in Bohr's quantized orbits. Naturally, the orbit has a radius 207 times smaller and energy 207 times greater than that of an electron. Muonic *x*-rays are produced when such μ-mesons are excited and de-excited, the energies of such *x*-rays depend on the value of R, the nuclear radius and may be used to estimate the size of the nucleus.

The average value of R obtained by this method is $(1.20 \pm 0.03)\, A^{\frac{1}{3}} \times 10^{-15}$ m.

- The other important method for estimating R is the electron scattering method into which we are not entering here. We shall highlight some of the important conclusions arrived at by a comparison of these methods in respect of atomic nucleus. These are (i) the distribution of density of protons in the nucleus is slightly different from that of all the nucleons within the nucleus, (ii) the nuclear charge distribution is not spherically symmetric in some nuclei which possess electric quadrupole moment and (iii) the charge radius of a nucleus is not significantly less than the radius of nuclear matter consisting of all the nucleons.

Nuclear density:

Nuclear density – The nuclear density, ρ_N can be estimated from the relation

$$\rho_N = \frac{\text{nuclear mass}}{\text{nuclear volume}}$$

But the nuclear mass M_N is approximately equal to Am_N where A is the mass number and m_N the mass of the nucleon $\simeq 1.67 \times 10^{-27}$ kg.

The nuclear volume, $V_N = \frac{4}{3}\pi R^3 = \frac{4}{3}\pi\,(r0A^{\frac{1}{3}})^3 = \frac{4}{3}\pi r_0^3 A$.

$$\therefore \rho N = \frac{AmN}{\frac{4}{3}\pi r_0^3 A} = \frac{mN}{\frac{4}{3}\pi r_0^3} = \frac{1.67 \; x \; 10^{-27}}{\frac{4}{3}\pi\,(1.3 \; x \; 10^{-15})^3}$$

Using the relation (1.8.3) above, we have

$$\rho N \simeq 1.816 \text{ x } 10^{17} \text{ kg /m}^3$$

an unusually large number. So, the nuclear matter is in a highly compressed state and the nuclear density is extremely high being comparable to that of some stars like the white dwarfs.

- Since the atomic radius = 10^4 x nuclear radius, the atomic density, $\rho N \simeq 2 \text{ x } 10^{17} / (10^4)^3 \simeq 2 \text{ x } 10^5$ kg/m^3 which is much smaller compared to pn. The density of water, pw = 10^3 kg / m^3 which is much smaller compared to pA.
- As the density of nucleus is independent of A (Eq.1.8.3), its value is almost the same for all nuclei.

Nuclear charge

The charge of a nucleus is due to the protons present in it. Each proton has a positive charge equal to 1.6×10^{-19} C.

∴ The nuclear charge = Ze, where Z is the atomic number.

3.4 Angular momentum of nucleus: Nuclear spin

While studying Bohr's theory of atom, we observed that the corresponding atomic states of the isotopic nuclides do not possess the same energy. This is because the reduced mass of the electron depends on the nuclear mass. This isotopic shift apart, each atomic energy level with a given value of J shows a splitting even when a single isotope of an element is taken. This splitting, which is finer than the fine structure splitting, is referred to as hyperfine structure of the level, observed only when a high resolution interferometer is used. The spinning property of an electron was introduced to explain the doublet structure of the alkali spectra. In a like fashion. Pauli suggested in 1924 that the hyperfine structure could be explained if we assume that the nucleus of the atom possesses an angular momentum of spin with an associated magnetic moment.

The methods employed to study the angular momentum of spin and the magnetic moment of nuclei are based on (i) the hyperfine structure of spectral lines, (ii) alternating intensities in homouclear molecular spectra, (iii) microwave spectra, (iv) magnetic resonance and the deflection of atomic and molecular beams, (v) nuclear magnetic resonance (NMR) in bulk and (vi) optical detection of NMR.

The spin of a nucleus is the resultant of the spins of its constituent nucleons – protons and neutrons. It turns out that the spins of protons and neutrons can be represented, like that of an electron, by the same quantum number $\frac{1}{2}$. They have an angular momentum $\frac{1}{2}$ (h/2π) or - $\frac{1}{2}$ (h/2π) respectively.

In addition to the spin angular momentum, the protons and neutrons in the nucleus have orbital angular momentum such that its magnitude in specific Z-direction is an integral multiple of h/2π. Thus the intrinsic angular momentum of a nucleus is a vector $\vec{I}$ such that

$$\vec{I} = \Sigma_n l_n + \Sigma_n l_s = \vec{L} + \vec{S}$$

Where $\Sigma_n l_s$ is the contribution of the spin and $\Sigma_n l_n$ is the contribution of orbital angular momentum of all the nucleons.

The magnitude of the total angular momentum vector $\vec{l}$ is given by $\vec{l} = \sqrt{I(I+1)}$ (h/2$^{\pi}$)

Where I is a quantum number that gives the maximum value of $\vec{l}$ along the specific z-direction. I may be zero, half integral or an integral multiple of h/2$^{\pi}$.

Now, the orbital angular momentum $\vec{L}$ is an integral multiple of h/2$^{\pi}$. The spin angular momentum $\vec{S}$ is either a half-integral multiple or an integral multiple of h/2$^{\pi}$ depending on whether the number of nucleons is odd or even. The nuclei in the ground state with even Z and even N nucleons have zero angular momentum without exception. So for even A-type nuclei having either odd Z, odd N or even Z, even N nucleons, the vector $\vec{I}$ will be zero or an integral multiple of h/2$^{\pi}$. And for odd A-type nuclei having either odd Z, even N or even Z, odd N nucleons, $\vec{l}$ will be an odd half-integral multiple of h/2$^{\pi}$.

$\therefore \vec{I} = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \frac{7}{2}$ . . . for odd A-type nuclei

and $\vec{I}$ = 0,1,2,3 . . . for even A-type nuclei

The total angular momentum $\vec{l}$ of a nucleus is loosely but usually called the spin of the nucleus or nuclear spin, but it is different from the spin angular quantum number. The total angular momentum of a nucleus can be computed from the multiplicity and relative spacings of spectral lines in an applied magnetic field . If a nucleus with total angular momentum $\vec{l}$ be placed in an externally applied magnetic field, the magnetic quantum number m_I have values ranging from + I to –I and thus split the energy levels into 2I + 1 sub-levels. The transitions between these sub-states may be used to estimate I from the multiplicity of spectral lines. We have already discussed earlier, in this section, the different methods used to determine $\vec{l}$.

- The symbol J is also sometimes used to represent the total angular momentum of a nucleus

Nuclear energy levels.

Studies of inelastic scattering and nuclear reactions indicate that nuclei posses a discrete spectrum of excited states – the nuclear energy levels. As suggested in Bohr's theory of H-atom, here also, when the nucleus is in an excited state it may give up the energy and come back to the ground state by an emission of photon of energy exactly equal to the difference of energies of the two states involved. This magnitude of energy however is such that the photon is a γ radiation.

There are definite minimum energies needed to disrupt the nucleus into various components. For instance, the light nucleus ^{12}C may break up into

(i) $^{8}Be + ^{4}He$ (ii) $^{11}B + ^{1}H$
(iii) $^{11}C + ^{1}n$ (iv) $^{10}B + ^{2}H$ etc.

The minimum energy values for the above disruptions of ^{12}C nucleus are all above the ground state energy. It implies that the nucleus in the ground state is stable against these decays. But those excited states which are above the thresholds can break up in various ways indicating that the nucleus does not remain in an excited state indefinitely. The life time of an excited state is finite and may be short or long depending on such factors as the probability of decay, the number of possible decay modes, selection rules, etc. From the uncertainty relation $\Delta E . \Delta t \sim h$, it follows that if the life time is finite, the energy of the state cannot be ideally sharp but would show a width or energy spread, $\Delta E = h/\Delta t$.

Nuclear magnetic dipole moment

The magnetic dipole moment associated with a spinning electron is given by 1 Bohr magneton, μ_B where

$$\mu_B = \frac{eh}{2m_e} = \frac{eh}{4\pi m_e} \left(\because h = h/2\pi\right)$$

$$= 9.27 \times 10^{-24} \text{ J/Tesia}$$

In analogy, the magnetic dipole moment associated with nuclear spin is given by what is called a nuclear magneton, $\mu_{N.}$

$$\mu_N = \frac{eh}{2m_p} = \frac{eh}{4\pi m_p} = 5.051 \text{ x } 10^{-27} \text{ J/Tesla}$$

Where $m_p = 1836\ m_e$, the mass of a proton.

The measured values of magnetic moment of proton and neutron are: $\mu_p = 2.7927\ \mu_N$

$\mu_n = 1.9131\ \mu_N$

The ratio of a nuclear magneton to a Bohr magneton

$$\frac{\mu_N}{\mu_B} = \frac{m_e}{m_p} = \frac{1}{1836}$$

The nuclear magneton is thus 1836 times smaller than the Bohr magneton and is also called Rabi magneton.

The magnetic moments of proton and neutron are intimately related to their intrinsic spin angular momenta,

$$v_p = s_p h, p_n = s_n h$$

Where $8_p = s_n = \frac{1}{2}$.

For all nuclei, the magnetic moment μI is, in amalogy with that of electron, given by

$$\mu I = g\mu_N = g\frac{eh}{2m_p}$$

Where Lande g-factor (also called gyromagnetic ratio) varies from nucleus to nucleus.

For proton, g = 2.792 so that μproton = 2.7927μ_N.

For neutron, g = - 1.913 so that μneutron = 0 1.9131 $\mu_{N.}$

- The magnetic moment of a proton is not 1 nuclear magneton; it has instead a moment +2.7927 $\mu_{N.}$ The positive sign indicates that the direction of the magnetic moment vector, $\vec{\mu}$, coincides with that of the angular momentum vector $\vec{I}$, as if there were circulation of positive charge.

- The neutron has no net electric charge. Bu it has also a magnetic moment equal to – 1.9131μ_N. The negative sign points out that the direction of angular momentum vector $\vec{I}$ is opposite to that of the magnetic moment vector $\vec{\mu}$.

Note 1. Nuclei possess magnetic dipole moments because of the magnetic moments of individual nucleons. In even-even nuclei, the resultant spin is zero and the magnetic moment is also zero. In case it is not zero, it is measured in terms of nuclear magnetron.

Note 2. It is rather hard to understand how the neutral particle, neutron, can have a magnetic moment. It can be attributed to the internal structure of neutron which can be looked upon as a combination of a proton and an electron. These particles, spinning in different directions, have different values for magnetic moments which when vectorially added give a negative magnetic moment. Further, the fractional values of dipole moments of neutrons and protons show that they are not simple in structure.

According to the current view, neutron consists of three quarks (Ch. Elementary particles) of fractional charges. The charge distribution inside neutron is thus not symmetrical leading to a non-zero magnetic moment.

Table 1.1 : Magnetic dipole moment of some nuclides

Nuclide	Spin	Magnetic dipole moment

		(in unit of μ_N)
^1_0n	$\frac{1}{2}$	- 1.9131
^1_1H	$\frac{1}{2}$	2.7927
^2_1H	1	0.8574
^3_2He	$\frac{1}{2}$	-2.1275
^3_1H	$\frac{1}{2}$	2.9789
^4_2He	0	0

The net magnetic moment $\vec{\mu}$ of a nucleus depends on the resultant total angular momentum of the nucleus, $\vec{I}$ and is given by

$$\vec{\mu} = g\mu_N\vec{I} = \gamma\vec{I}$$

Where γ is the product of the gyromagnetic ratio g and the nuclear magneton μ_N.

The values of nuclear magnetic moments can be estimated by nuclear magnetic resonance (NMR) spectrometers, microwave spectrometers, molecular beam magnetic deflection method etc.

Determination of nuclear magnetic moment:

Nuclear magnetic resonance (NMR) method

The nuclear magnetic moment may be directly determind experimentally by the methods of magnetic resonance and nuclear resonance absorption.

Rabi's method – The first accurate measurement of nuclear magnetic moment by magnetic resonance was made by Rabi and co-workers. The method is known as molecular beam magnetic resonance and resembles the Sterm-Gerlach experiment.

It consists essentially in allowing a beam of molecules from an oven (source) into the region of a non-uniform magnetic field of a magnet A. The beam follows a curved path, as shown in Fig.1.3, and passes through the slit S$_1$. The magnet B produces however a uniform magnetic field and as such would exert a net zero deflecting force. But in the region of this uniform field, another magnetic field is created (dotted circle) at right angles to the previous fields by a high-frequency (hf) electric current. Next, the magnet C produces a non-uniform field that bends back the molecular beam such that the beam passes through the slit S$_2$ and is detected by the detector. Note that the magnets A and C give non-uniform fields in the same direction but with opposite gradients d$^{\vec{B}/d}$Z.

The molecules, on passing through the uniform field, are spatially quantized and process like a spinning top about the field direction. As the frequency of electric current is varied and made equal to that of precession, energy gets absorbed and the molecules take up a different quantized space-orientation. The magnet C deflects them to pass through the slit S$_2$. The decrease in the number of molecules through S$_2$ can then be used to find the processional frequency which depends on the magnetic moment and the intensity of the central uniform field. The method thus gives the value of the magnetic moment.

- Purcell, Torrey and others developed another method, known as magnetic nuclear resonance absorption, that uses no molecular beam. The solid (or liquid) sample is placed in a uniform magnetic field on which an hf-alternating field is superposed at right angles. Protons and simple nuclei precessing round the uniform field direction get their quantized angles of space-orientation abruptly changed by absorption of energy when the ac-field is tuned to resonance with precessional frequency. This absorption can be detected and magnetic moments computed.

Parity of nuclei

Parity is purely a quantum mechanical concept having no classical analogue.

First proposed by Eugene Wigner, it describes the kind of spatial symmetry of physical phenomena. A particle moving with a large velocity can be quantum mechanically associated with a wave and the wave motion can be described by a wave function ψ(x, y, z) which depends on the space coordinates (x,y,z). Also, if $\psi*$ be the complex conjugate of ψ, $\psi * \psi = |\psi|^2$ gives the probability of finding the particle at any given point.

Parity is the property of such a wave function representing a quantum mechanical nuclear state, which may or may not chage its sign or inversion of the space coordinates from (x,y,z) to (-x, -y, -z) throughout i.e., on reflection of the coordinate system at the origin. The parity of a nucleus is thus related to the behaviour of nuclear wavefront as a result of reflection.

Definition: If the sign of the wave function ψ does not, as a whole, change with the change in sign of the space coordinates – the so-called process of reflection of nuclear position about the origin of (x,y,z) system of axis – the parity is said to be even or positive. If however the sign of spatial part of wave function changes with change in sign of space co-ordinates, the system is said to have odd or negative parity.

$\psi(-x, -y, -z) = +\psi(x,y,z)$ even (+) parity

$\psi(-x, -y, -z) = -\psi(x, y, z)$ odd (-) parity

Parity, thus defined, depends on the quantum state of motion of the system and for a particle under a central force (i.e. hydrogen-like atoms), the parity P is determined by the orbital quantum number l : p = $(-1)^1$, so that for l = 0 or even, the parity is even and for l = odd, the parity is odd. In the case of a system of particles P = $(-1)^L$ where L = $\sum l$, the orbital angular momentum of the system.

Let us further investigate the consequence of this reflection of the coordinate system, called the parity operation, P. It is equivalent to changing from a right-handed frame of reference to a left-handed one.

Consider a system of *n*-particles so that

$$P\psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n) = \psi(\overrightarrow{-r_1}, -\vec{r}_2, \dots, \overrightarrow{-r_n})$$

Assuming that the Hamiltonian H of the system remains invariant under inversion about the origin,

$$PH(\vec{r}_1, \vec{r}_2, \dots) = H((\vec{r}_1, \vec{r}_2, \dots)$$

But, Schrodinger's wave equation gives

$$H\psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n) = E\,\psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n)$$

$$\therefore PH\psi(\vec{r}_1, \vec{r}_2, \dots) = PE\psi(\vec{r}_1, \vec{r}_2, \dots) = E\,\psi(-\vec{r}_1, -\vec{r}_2), \text{ by (1.12.1)}$$

$$\therefore \mathrm{H}\psi(-\vec{r}_1, \overrightarrow{-r_2}, \dots) = \mathrm{E}\,\psi(-\vec{r}_1, -\vec{r}_2), \text{ by (1.12.4).}$$

$\psi(-\vec{r}_1, \overrightarrow{-r_2}, \dots)$ thus satisfies the same differential equation as $\psi(-\vec{r}_1, \overrightarrow{-r_2}, \dots)$. Assuming non-degeneracy, the two solutions of Schrodinger's equation must be connected by a phase factor (constant).

$$\therefore \psi(-\vec{r}_1, \overrightarrow{-r_2}, \dots) = \mathrm{k}\,\psi(-\vec{r}_1, \overrightarrow{-r_2}, \dots)$$

$$\therefore \mathrm{P}\psi(-\vec{r}_1, \overrightarrow{-r_2}, \dots) = \psi(-\vec{r}_1, \overrightarrow{-r_2}, \dots) = \mathrm{kP}\,\psi(-\vec{r}_1, \overrightarrow{-r_2}, \dots)$$

$$= \mathrm{k}\psi(-\vec{r}_1, \overrightarrow{-r_2}, \dots) = \mathrm{kk}\,\psi(-\vec{r}_1, \overrightarrow{-r_2}, \dots)$$

by virtue of (1.12.6)

$$\boxed{k^2 = 1 \Rightarrow \mathrm{k} = \pm 1}$$

Thus, according to (1.12.6) and (1.12.8), parity is a quantum number implying that all wave functions are either odd or even under space inversion or P-operation.

- A part from orbital parity, elementary particles may also have what is called intrinsic parity referring to inversion of some internal axis of the particle. It is defined in a relative manner. The nucleons are taken to have even parity, purely by convention, and it is then fixed for other particles such that the total parity, defined as the product of orbital and intrinsic parity, is conserved in nuclear reactions involving strong and electromagnetic interactions between the particles.

It was believed, till 1956, that in all nuclear reactions, the parity was conserved. But in 1956, direct experimental evidence was obtained that parity is not conserved in nuclear phenomena involving weak interaction forces. The non-conservation of parity was first suspected theoretically by Li and Yang and was subsequently confirmed experimentally by Wu in 1956.

The weak interaction in β-decay provides an example of non-conservation of parity. The conservation of parity leads to some important selection rules in nuclear, atomic and molecular processes and in the production and decay of elementary particles. Parity is thus an important characteristic of a state describing quantum mechanical systems.

● An interesting consequence of the fact that parity P is a good quantum number is that nuclei can have no permanent electric dipole moment. While we are familiar with the classical definition of electric dipole moment of a charge distribution, quantum mechanically it is defined as

$$\langle D\rangle = \int \sum_j e_j \vec{r}_j |\psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n)|^2 \, d\vec{r}_1 d\vec{r}_2 \dots d\vec{r}_n$$

If now we replace $\vec{r}_1$ by $-\vec{r}_1$, $\vec{r}_2$ by $-\vec{r}_2$, etc., the first factor $\vec{r}_j$ changes sign, but not $|\psi|^2$ because of parity.

$$\boxed{\langle D\rangle = -\langle D\rangle \;\Rightarrow \langle D\rangle = 0}$$

We thus obtain the important result that electric dipole moment of a nucleus in its ground state vanishes. This is also true for all non-degenerate excited states.

Electric quadrupole moments

Many nuclei are not exactly spherical lines. Departure of nuclear charge distribution from spherical symmetry is specified by the nuclear quadrupole moment.

If a localized charge distribution described by $^*\rho(\vec{r})$ be placed in an external potential $\Phi(\vec{r})$, the electrostatic energy of the system is

$$W = \int \rho(\vec{r})\,\Phi(\vec{r})\,d^3r$$

With origin $\vec{r} = 0$ as the centre of mass of the nucleus. Since $\rho(\vec{r})$ is confined to the small nuclear volume, $\Phi(\vec{r})$ may be expanded in a Taylor series to get

$$\Phi(\vec{r}) = \Phi(0) + r\vec{\nabla}\Phi(0) + \frac{1}{2}\sum\sum x_i x_i \frac{\partial^2\Phi(0)}{\partial x_i \partial x_j} + \ldots (i, j = x_1, x_2, x_3).$$

Now, the electric field $\vec{E} = -\vec{\nabla}\Phi$. But since the field is source-free, $\vec{\nabla}.\vec{E} = 0$.

$\therefore \vec{\nabla}.\vec{E} = \vec{\nabla}.(-\vec{\nabla}\Phi) = -\nabla^2\Phi = -\frac{\partial^2\Phi}{\partial x_i \partial x_j}\delta_{ij} = 0$; *i*, *j* being dummy indices.

Substracting $\frac{1}{6}r^2 \frac{\partial^2\Phi}{\partial x_i \partial x_j}$ from the third term of (1.13.2), we obtain

$$\Phi(\vec{r}) = \Phi(0) + r\vec{\nabla}\Phi + \frac{1}{6}\sum_i\sum_j(3x_ix_j - r^2\delta_{ij})\frac{\partial^2\Phi(0)}{\partial x_i \partial x_j} + \ldots$$

$$= \Phi(0) - r\vec{E}(0) - \frac{1}{6}\sum_i\sum_j(3x_ix_j - r^2\delta_{ij})\frac{\partial E_j(0)}{\partial x_i} + \ldots$$

Inserting (1.13.3) in (1.13.1), we obtain

$$W = q\Phi(0) - {}_pE(0) - \frac{1}{6}\sum Q_u \frac{\partial E_j(0)}{\partial x_i}$$

Where $q = \int \rho(\vec{r})^2 d^3r$ - total charge of the nucleus,

$\rho = \int \vec{r}p(\vec{r})d^3r$ - dipole moment of charge distribution,

$Q_{ij} = \int(3x_ix_j - r^2\delta_{ij})\, p(\vec{r})d^2r$

= (i-j) component of quadrupole moment tensor of the charge distribution inside the nucleus.

$Q_{ij} = Q_{ji}$ implies that the tensor is symmetric and of rank 2 in three dimensions. It can be shown that

$$Q_{11} + Q_{22} + Q_{33} = 0$$

By a principal axis transformation, one obtains a diagonal representation of the tensor. As the tensor is, by (1.13.6), traceless it, remains invariant under the above transformation. The number of independent components is two. For a nucleus in the shape of a spheroid (Fig.1.4) – a prolate (the charge distribution is stretched in the z-direction) or an oblate (the charge distribution is stretched perpendicular to z-direction), the axial symmetry gives

$$Q_{11} = Q_{22} \neq Q_{33} \Rightarrow Q_{33} = -2\, Q_{11} = -2\, Q_{22}$$

So, for such a system, $Q_{33} = \int \rho(\vec{r})(3z^2 - r^2)d^3r$

● 1. For a prolate spheroid, $\frac{x^2}{a^2} + \frac{y^2}{a^2} + \frac{z^2}{b^2} = 1$ (a > b)

Now, assuming a uniform charge density, we have

$$\rho(\vec{r}) = q/\frac{4}{3}\pi a^2 b$$

Where Z is the charge number of the nucleus.

●2. For a spherically symmetric charge distribution,

$$Q_{11} = Q_{22} = Q_{33} = Q, \text{ say}$$

Then, by (1.13.6), $Q_{11} = Q_{22} = Q_{33} = Q = 0$

Obviously, the existence of quadrupole moment Q implies deviation from the spherical symmetry, as stated at the very outlet.

It R be the average nuclear radius and Δ R the deviation of R from the direction of symmetry axis, then

$$Q = \frac{6}{5} ZR^2 (\Delta R/R)$$

Q is usually measured in unit called barn, where 1 barn = 10^{-28} m^2.

The magnitude of Q lies mostly in the range 10^{-28} to 10^{-30} m^2 and depends on the radius and charge of the nucleus and the amount of its deviation from spherical symmetry. Quadrupole moment was first discovered in deuteron from the observations of the hyperfine structure of atomic spectral lines. For deuteron, Q = +2.82 milli-barn (mb) which shows that the charge distribution in ^2H nucleus has the shape of a prolate spheroid.

Q can also be estimated from the interaction of electric field gradient with the quadrupole moment of the nucleus from Mossbauer spectroscopy, microwave spectroscopy, paramagnetic resonance spectroscopy, nuclear quadrupole resonance spectroscopy and from optical hyperfine spectra.

While a nucleus like deuteron, ^2H, has the ratio $\Delta R/R$ = 4%, the value of this ratio may be as high as 25-30%.

The above discussion on quadrupole moment has been made from classical considerations. When quantum mechanics is applied, it receives a new definition.

The charge distribution in a nucleus is treated quantum mechanically by defining quadrupole moment operator by

$$Q_{ij} = \sum_{p\,(protons)} \left[3x_{pi}x_{pj} - \delta_{ij} r_p^2\right]$$

Where x_p's are the coordinates of protons.

$$\therefore (Q_{ij}) = \int \sum_{(protons)} \left[3x_{pi}x_{pj} - \delta_{ij} r_p^2\right] |\emptyset(\vec{r})|^2 d^3r$$

$$= C \left\langle \left[\frac{3}{2}\left(J_i J_J + J_j J_I\right) - \delta_{ij} J^2\right] \right\rangle$$

Where C is a constant and < > indicate any matrix element between (2J + 1) – nuclear states of angular momentum quantum number J, labelled by J_z. Thus all the matrix elements of Q_{ij} are determined by a single quantity.

We usually take the expectation value of Q_{33} in the state with J_z equal to the maximum value of J, and define this as the nuclear electric quadrupole moment Q.

$$\therefore Q = C\,[3J^2 - J\,(J+1)]$$

$$= CJ\,(2J - 1)$$

$$\Rightarrow C = \frac{Q}{J\,(2J-1)}$$

Other matrix elements are determined in terms of Q ; Q vanishes for J = $\frac{1}{2}$ or J = 0.

Statistics of nuclei

The statistical behaviour of nuclear particles such as protons, neutrons, electrons and other fundamental particles can only be described quantum mechanically as they are indistinguishable from each other.

The wave function of a system of two identical particles, describing a particular property of the particle, can be either symmetric or antisymmetric by the interchange of spatial coordinates of the two particles. When the wave function does not change in sign by such an interchange, it is said to be symmetric. The corresponding quantum statistics is known as Bose-Einstein (B.E.) statistics. If however the sign of the wave function changes by the interchange of coordinates, the wave function is said to be antisymmetric and the corresponding statistics is called Fermi-Dirac (F.D.) statistics.

All fundamental particles or their assemblies obey either the B.E. statistics or the F.D. statistics. Those parties which have total angular momentum or spin equal to zero or an integral multiple of h/2π follow the Bose Einstein statistics and are called bosons. Such particles do not obey Pauli's exclusion principle so that any number of bosons can occupy the same quantum state. Particles such as photons, π-mesons, K-mesons, and all nuclei having even mass numbers such as $^{2}_{1}\mathrm{D}$, $^{4}_{2}\mathrm{He}$, $^{16}_{8}\mathrm{O}$, … and integral spin follow the Bose's statistics and are bosons.

Particles with an odd half integral spin ($\frac{1}{2}, \frac{3}{2}, \dots$) that is, having total angular momentum equal to an odd half integral multiple of h/2π however follow the Fermi statistics and are called fermions. Fermions, unlike bosons, follow Pauli's exclusion principle so that no two identical particles can have all the quantum numbers the same barring them to occupy the same quantum state. Particles such as e^-, e^+, p, n, μ meson and nuclei of odd mass number like $^{1}_{1}H$, $^{7}_{3}Li$, $^{10}_{9}F$, $^{23}_{11}Na$, $^{31}_{15}P$, $^{35}_{17}CI$, … which possess half integral spin follow the Fermi-Dirac statistics and are fermions.

In nuclear reactions, the statistics of the particles is conserved. If F.D. statistics is represented by symbol – 1 and B.E. by +1, the statistics of a group of particles is determined by the product of the symbols for each. For instance, let us take the case of annihilation of a positron and an electron. Before annihilation, the statistics of the system is (-1) x (-1) - +1. After annihilation, the electron-positron pair gives rise to two photons which are bosons and they have the statistics (+1) x (+1) = +1. So, there is the conservation of statistics in nuclear phenomena.

3.5 Binding Energy

Nuclear mass and binding energy

The mass-spectroscopic measurements give not the masses of the nuclei but those of the atom. The nuclear mass M_{nuc} is obtained from the atomic mass M(A,Z) by subtracting the masses of Z orbital electrons.

$$\therefore \quad \boxed{M_{nue} = M(A,Z) -}$$

The above expression however is not exact in that the binding energies of the electrons have not been taken into consideration. The error however is negligibly small.

The nuclei are very strongly bound and energies of $\sim$ few MeV are needed to break away a nucleon from the nucleus. In contrast, only a few eV is necessary to detach an orbital electron from an atom. So, to break up a nucleus of Z protons and N neutrons completely into separate particles, a minimum amount of energy is to be supplied to the nucleus. This supplied energy is called the binding energy. E_B, of the nucleus.

Conversely, to build, out of Z protons and N neutrons remaining at rest and separate from one another, a nucleus of mass number A(=N + Z) and nuclear charge Z, an amount of energy equal to E_B will be evolved. But what indeed is the source of this energy?

According to special relativity, the energy equivalent E corresponding to a complete conversion of a mass m into energy is E = me^2, where c is the velocity of light in free space. In forming a nucleus out of the constituent particles, a fraction of the total mass of the constituents disappears and the evolution of equivalent energy E_B occurs. It ΔM be the amount of mass disappeared, then the

$$\text{Bind energy, } E_B = \Delta M.c^2$$

If M_H and M_n be the masses of hydrogen atom and the neutron respectively,

$$\Delta M = ZM_H + NH_n - (M(A,Z)$$

Where M(A,Z) is the mass of the atom of mass number A and atomic number Z.

$$\therefore E_B = [ZM_H + NH_n - M(A,Z)]c^2$$

$$= [ZM_p + NM_n + Zm_3 - M_{nuc} - Zm_e]c^2$$

$$\boxed{\therefore E_B = [ZM_p + NM_n - M_{nuc}]c^2}$$

Using equation

So, the mass-loss ΔM for the formation of nucleus is equal to the sum of the masses of Z protons, N-neutrons minus the nuclear mass of the atom.

In energy unit, $\Delta M = ZM_p + NM_n - M_{nue}$, dropping c^2 from the above equation (1.3.4).

The methods of determination of atomic masses have already been described in companion volume: Physics of Atoms in Chapter : Ions and Electrons.

Unit of atomic mass

The unit of atomic mass is presently defined to be one-tweifth of the mass of the atom of carbon isotope ^{12}C taken to be exacity 12 units, and is symbolized by u, the abbreviation for 'unified atomic mass unit'. This comes to usage since 1961 by an international agreement. It is called 'unified' since prior to 1961 there was no unity among the physicists and the chemiss in regard to atomic mass unit. While the physicists' unit was one-sixteenth of the mass of ^{16}O isotope (taken to be exactly 16 unit), called the 'atomic mass unit' (amu), the chemists' atomic mass unit was one-sixteenth of the average atomic weight of the three isotopes of oxygen ^{16}O, ^{17}O and ^{18}O with relative abundance 99.76%, 0.04% and 0.20% respectively. The conversion factor from one scale to the other is

$$1\ u : 1amu = 1.0003172:1$$

Since, 1 mole of ^{12}C has the mass of 12g or 12 x 10^{-3} kg, the unit of atomic mass in ^{12}C scale is

$$1u = \frac{1}{12} \times \frac{12 \times 10^{-3}}{N_A}\ kg = \frac{10^{-3}}{6.02205 \times 10^{23}}\ kg.$$

(N_A = Avogadro number)

$$= 1.660566 \times 10^{-27}\ kg.$$

The energy-equivalent of 1 u is thus

$$1\ u = 1.660566 \times 10^{-27} \times c^2 = 1.660566 \times 10^{-27} \times 8.98755 \times 10^{16}$$

$$= 14.924427 \times 10^{-11}\ J = \frac{14.924427 \times 10^{-11}}{1.60219 \times 10^{-13}}\ MeV$$

$$= 931.502\ MeV = 931.5\ MeV$$

The energy-equivalence of the rest mass of electron, proton and neutron are respectively given as under:

Electron (m_e) = 9.10953 × 10^{-31}kg = 5.48580 × 10^{-4} u 0.51 MeV/c^2

Proton (m_p) = 1.67265 $\times 10^{-27}$ kg = 1.0072765 u = 938.27 MeV/c^2

Neutron (m_n) = 1.67495 10^{-27} kg = 1.0086650 u = 939.57

Importance of accurate determination of atomic masses – Modern mass spectroscopes can measure atomic masses accurately up to one part in a million or still better. Such accuracies are of great importance in correctly determining the nuclear binding energies and nuclear disintegration energies. A simple calculation on the α-disintegration of a heavy element like Ra-226 would indicate that the disintegration energy is less than one part in 40,000 of the mass disintegrating nucleus. So, if the atomic masses are not measured more accurately than that order, correlation of the measured disintegration energy with the change in mass due to disintegration would be a far cry.

Binding energy and stability of nucleus

If $E_B > 0$, i.e. positive, the nucleus is stable and energy from outside is to be supplied to disrupt of nucleus into its constituents separately. If, however, $E_B < 0$, i.e. negative, the nucleus is unstable and will disintegrate of itself. The E_B, more is the stability.

As an example, let us compute the binding energy E_B for an α-particle, i.e. ${}^4_2He$. The belium nucleus is made up of 2 protons and 2 neutrons.

$\therefore ZM_p = 2 \times 1.007276 = 2.014552$ u

$NM_n = 2 \times 1.008665 =$ 2,017330 u

Total = 4.031882 u

Atomic mass of ${}^4_2\text{He}$ = 4.002603 u

Difference = + 0.029279 u

The plus sign indicates that the nucleus is stable. And since 1 u $\simeq$ 931 MeV, the binding energy in MeV is $E_B = 0.029279 \times 931 = 27.16$ MeV.

The binding energy of He-nucleus is 27.16 MeV and this explains why it is a very stable structure, coming out as it does as α-particle in radioactive decay.

Every nuclide has a fixed binding energy meaning that the same amount of energy would always be required to pull all the nucleons completely apart. This binding energy divided by the number of nucleons is called the (average) binding energy per nucleon or binding fraction, $f_B = E_B/A$.

3.6 Packing Fraction

Mass defect and packing fraction

Accurate determination of the atomic masses indicates that they are not exactly whole numbers, although the difference is small. For instance, the atomic mass of ^{12}C is exactly 12u, but the masses of other atoms, although very close to, are not exactly whole numbers (integral). For instance,

^{1}H-=1.007825 u; ^{2}H = 2.014102 u

4He = 4.002603 u; ^{16}O – 15.994915 U; ^{226}Ra = 226.02543 u etc.

The departure of measured atomic mass from M (A,Z) from mass number A is rather significant.

Mass defect – The difference between the measured atomic mass M(A,Z) in u, and the mass number A of a nuclide is called the mass defect, ΔM'.

$$\boxed{\Delta M' = M(A,Z) - A}$$

The mass defect of ^{4}He = 4.002603 – 4 = +0.002603 u and that of ^{16}O – 15.994915 – 16 = -0.005085 u. The mass defect can therefore be both positive and negative. It is found that the mass defect is positive for very light and very heavy atoms, and it is negative for atoms in the intermediate range.

Packing fraction

Packing fraction – Aston expressed the departure of atomic masses from their mass numbers in terms of packing fraction for each nuclide.

The packing fraction f is defined as the mass defect per nucleon in the nucleus, that is, the mass defect of an atom divided by its mass number. So,

$$f = \frac{\Delta M'}{A} = \frac{M(A,Z) - A}{A}$$

$$= \frac{M(A,Z)}{A} - 1$$

$$\boxed{M(A,Z) = A(1 + f)}$$

Note that the f has the same sign as the mass defect $\Delta M'$

It is found that the packing fraction f varies with the mass number in a systematic fashion (the packing fraction curve) as represented graphically in Fig.1.1. It is observed that the packing fraction is positive for very light nuclei and as A increases, f decreases rapidly, becoming negative for A > 20. It attains a minimum value (negative) at $A \sim 60$, whence it starts increasing again but rather slowly. For $A \sim 180$, it becomes positive again.

This systematic variation of f with A can be explained from consideration of nuclear binding energy. The binding fraction.

$$F_B = \frac{E_B}{A} = \frac{ZM_p + NM_n - M(A,Z)}{A}$$, in energy unit.

Computation of f_B's for different nuclides shows that they are highly variable. For instance, it is 1.112 MeV for deuteron. 7.07 MeV for α-particle (^{4}He) and 7.98 MeV for ^{16}O. These difference reflect the relative strengths of their binding. While ^{2}H is weakly bound, ^{4}He or ^{16}O are relatively more strongly bound.

Binding fraction *vs* mass number curve

The nature of variation of f_B with A, for different nuclei is represented graphically in Fig.1.2 and is called the binding fraction curve.

A critical survey of the curve will readily bring out the following points of immense physical importance:

1. f_B is very small for very lightnuclei and goes on increasing rapidly with increasing A and reaches a value $\sim$ 8 MeV for the mass number A$\sim$ 20. Thereafter, the rise of the curve is much slower, reaching of maximum value of 8.7 MeV for A = 56. If A is increased still further, the curve again starts decreasing but slowly.

2. The variation in f_B is very slight in the range of mass number $20 < A < 180$ and in this region f_B may be considered to remain virtually constant having a mean value $\sim$ 8.5 MeV

3. For A > 180, that is, for heavy nuclei, the f_B-value decreaces monotonically with increasing A and is $\sim$ 7.5 MeV for the heaviest nuclei.

4. A rapid fluctuation in f_B is noted for very light nuclei with some peaks in the curve in this region, corresponding to the even-even nuclei, such as ^{4}He, ^{8}Be, ^{12}C, ^{16}O etc. i.e. with mass number A = 4n, where n = 1,2,3,4 . . . etc. Similar peaks but less prominent in the curve are also seen at Z or N equal to 20, 28, 50, 82, 126. These are called magic numbers (to be discussed, in more details, in Chapter: Nuclear models).

The significance of these peaks is that the corresponding nuclei are more stable relative to those in their neighbourhood.

- The phenomena like the energy release in nuclear fission and also the reason for α-decay of heavy nuclei can be qualitatively explained by the binding fraction curve and will be discussed at appropriate places.

Complementarity of binding and packing fraction curves

It is easy to see that the nature of the binding fraction curve (Fig.1.2) is complementary to the packing fraction curve (Fig.1.1). It can be readily understood why it is so.

We have: $E_B = ZM_P + NM_n - M(A,Z)$

$= Z(1 + f_p) + N(1 + f_n) - A(1 + f)$, using (1.6.1)

$= (Z + N) + Z f_p + N f_p - A - Af$

$= Z f_p + N f_n - Af \quad (\because A = Z + N)$

$\therefore \; f_B = \frac{E_B}{A} = \frac{Zf_p + Nf_n}{A} - f$

Now, the first term on the right side of (1.7.1) is nearly constant, particularly for lower A- values when $Z \simeq N \simeq A/2$. Thus, f_B increases or decreases as f decreases or increases respectively. Hence, the curves representing the variation of f_B *and* f appear complementary in character. Where the (f-a) curve shows a minimum, the (f_B-A) curge; and for higher A, the region of positive slope in the first curve corresponds to the region of negative slope in the second.

3.5 Nuclear stability

Nuclear stability means that nucleus is stable meaning that it does not spontaneously emit any kind of radiation. On the other hand, if the nucleus is unstable, it has the tendency of emitting some kind of radiation, which makes it radioactive. Therefore the radioactivity is associated with unstable nucleus:

Nuclear forces :

According to Coulomb's law, the positively charged protons, closely spaced within the nucleus, should repel each other strongly and they should fly apart. It is therefore difficult to explain the stability of nucleus unless one assumes that nucleons are under the influence of some very strong attractive type forces. The forces are essentially equal in magnitude as warranted by experimental evidence and were studied extensively over a long period by the Japanese scientists Hideki Yukawa. In 1935, he described the chief characteristics of nuclear forces and postulated a particle, pion with a rest mass $270m_e$, that played an integral part in the explanation of nuclear forces. Yukawa was awarded Nobel Prize in physics in 1949 for his contributions to the understanding of nuclear forces.

According to Yukawa, the following are the characteristics of nuclear forces:

1. The are short range forces, i.e., effective only at short ranges.
2. They are charge-independent, i.e., they do not seem to depend on the charge of the particle.
3. They are the strongest known forces in nature.
4. They get readily saturated by the surrounding nucleons, and
5. They are spin-dependent.

We shall now discuss the above characteristics of the nuclear forces in somewhat more details.

Short range – The results of scattering experiments: *p-p* scattering n-p scattering etc. show that the nuclear forces operate over extremely short distance inside the nucleus. Between two nucleons, the distance is of the order of 1 fermi ($1fm = 10^{-15}$ m) or less. They are not like the inverse square law forces such as Coulomb force between electric charges. If a nucleus is bombarded with protons and if the range of nuclear force be of the same order of magnitude as Coulomb repulsion, they would be affected by both type of forces. But the scattering of protons will be different from the one corresponding to a pure Coulomb scattering.

The protons that pass not too close to the nucleus are scattered by electric repulsive forces. But if the energy of the incident protons be large enough to overcome Coulomb repulsion, they may pass very close to nucleus, within a distance $_{ro}$ from the centre of the nucleus, and fall in the range of attractive nuclear forces. They would then be captured and fall, as it were, into the potential well of the nucleus. The scattering of protons in this case is mainly due to strong and attractive nuclear forces and the distribution is distinctly different from Coulomb scattering.

There is however some evidence to suggest that at extremely short distance (0.5 fm), the attractive force turns into a repulsion so that in a stable nucleus, the nucleons do not get too close together.

Charge independence – Experimental evidence indicates that the interaction between any two nucleons is independent of the charge. Also the interactions among the nuclear forces between *n-n*, *p-p* and p-n, exclusive of Coulomb forces, have been found to be the same to a high degree of accuracy. This suggests that the proton and neutron can be considered as different charge states of the same particle. Charge independence of nuclear forces means interaction $p\text{-}p \simeq p\text{-}n \simeq n\text{-}n$. This is evident from the tendency of nuclei with small Z to have $Z = N = A/2$.

Strong forces - The strong interactions, the forces between the nucleons, are the strongest forces found in nature. The gravitational and the electromagnetic interaction were known to us long before the nuclear forces, as they were associated with macroscopic bodies, e.g., the gravitational forces between the planets and the sum and the electrical forces between charged bodes. But they are far weaker compared to the nuclear force. For instance, the gravitational force is only $\sim 10^{-40}$ of the strong interaction.

Although, protons repel each other via the electric force, they attract each other via the strong nuclear force that keeps the nucleons together. Neutrons attract other neutrons and protons via this strong nuclear force.

Saturation – Nuclear forces are the only ones in nature that show saturation effect. The ability of nuclear forces to act upon other particles attain a point of saturian meaning that a nucleon can interact strongly with only a limited number of neighbouring nucleons. Addition of nucieons only increases the total binding energy but not the B.E. per nucleon.

Summarising: (i) the forces between nucleons are attractive in nature when they are 0.5 – 25 F part; (ii) these forces are of short range having maximum value at about 2×10^{-15} m and fall off sharply with distance, becoming negligible beyond this range; (iii) they are charge-independent so that the nuclear force between a proton and a neutron or between a neutron and neutron are almost the same; (iv) they have the property of saturation – a particular nucleon interacts with a limited number of nucleons around it and the other surrounding ones remain unaffected. So they become saturated over short distances; (v) the nuclear forces depend on the natural orientation of spins of various nucleons and are different in parallel and antiparallel spins.

A more detailed study of interaction requires the specification of the form of the nuclear potential V(r). Scattering experiments with high energy particles show that there is a repulsive core at the centre of the nucleus but for which all nuclei would have the same radius, but $R = {}_{ro}A^{1/3}$.

● In addition to the strong nuclear force which is far stronger than Coulomb interation, there is, as indicated by experimental evidence, a third type of force which is also a short range force but much weaker than the nuclear force. This is termed weak interaction. It may be as small as 10^{-14} of strong nuclear force. It is also not of gravitational type.

Interestingly, the weaker the force, the larger must be the system in order that it might be of importance. For example, the strong interactions hold the nucleons, the electromagnetic force holds the larger systems of atoms and molecules, while the gravitational force becomes important only in astral systems.

The chief forces of nature are thus of the following four types:

(i) the strong nuclear force,
(ii) the electromagnetic force,
(iii) the weak interaction force and
(iv) the gravitational force.

● According to Yukawa's theory, protons and neutrons do not exist independently within a nucleus but constantly exchange charges by emission and absorption of π-mesons (pions) in themselves. This constant emission and absorption result in an exchange of virtual mesons by nucleons, within the nucleus, in ultra short intervals $\sim 10^{-23}$ to 10^{-24} s. As the exchange occurs in a very short time, the uncertainty principle requires that no visible change in nucleonic mass would be observed. This gives rise to rapid meson exchange or meson field between protons and neutrons in which meson acts as a quantum of nuclear force. The process in analogous to exchange of photons between charged particles in electromagnetic interactions.

1.16 Rutherford's α-ray scattering

Rutherford and his co-workers Geiger and Marsden made a detailed systematic study of the scattering of α-particles obtained from radioactive sources by ultra-thin foils of elements of high atomic weight e.g., gold, platinum etc. The foil is so thin that only single scattering occurs. They observed that the majority of α-particles were scattered through small angles, are expected, but a few however deviated by 90° and a still smaller number surprisingly turned back, being deflected almost by 180°. This large angle scattering was called anomalous scattering as Thomson's atom model could not explain it.

To explain the large angle scattering of α-particles, Rutherford proposed that inside the atom, the positive charge and almost the whole of the atomic mass were concentrated in a very small central region, called the nucleus, round which, in some sort of configuration, the electrons rotate in groups.

Rutherford's scattering formula

We shall now derive the Rutherford's scattering formula here after Max Born.

The nucleus of the atom with charge Ze and an α-particle with charge E(=2e), mass M and distant r, repel one another with a force $F = (Ze)\,E/4\pi\epsilon_0 r^2$ where ϵ_o is the permittivity of free space. Let the heavy nucleus be at rest. The α-particle would then describe under the central inverse square law of repulsion, one branch of a hyperbola PAP' with K, one of its foci (K,K'), as the nucleus.

Let b be the distance of the nucleus K from the asymptote of the hyperbola (Fig.1.5) which would have been described if there were no repulsion. This b is called the collision (or impact) parameter which may be defined as the minimum distance to which the α-particle would approach the nucleus if there were no force between them.

Let q be the distance of K from the vertex A. Then we have, from the equation of a conic (hyperbola)

$$Q = \varepsilon\,(1 + \cos\theta)$$

Where $\varepsilon =$ OK, the linear eccentricity; 0 is the origin and 0 the angle between the axis and the asymptote of the hyperbola.

$$\therefore q = b\,\frac{1+\cos\theta}{\sin\theta} = b\cot\frac{0}{2}\ (\because \sin\theta = b/\varepsilon)$$

Step 1. First, we seek a relation between the impact parameter b and the angle of scattering Φ, defined as the angle between the asymptotic direction of approach of the α-particle and the asymptotic direction in which it moves back, by applying the laws of conservation of energy and momentum. At a great distance from the nucleus, α-particle have only kinetic energy and let v be the velocity there.

From the conservation of energy when it passes the vertex A where the velocity is *vo*, we obtain

$$\frac{1}{2}Mv_0^2 + \frac{ZeE}{4\pi\varepsilon 0q} = \frac{1}{2}Mv^2 ,$$

$$\left(\frac{ZeE}{4\pi\varepsilon 0q} = \text{potential energy}\right)$$

$$\therefore \frac{v_0^2}{v^2} = 1 - \frac{2ZeE}{4\pi\varepsilon 0q\,Mv^2} = 1 - \frac{2k}{q} = 1 - \frac{2k\ \sin\theta}{b(1+cos\theta)} ,$$

Using (1.16) and also by substituting

$$k = \frac{ZeE}{4\pi\varepsilon 0q\,Mv^2}$$

From conservation of angular momentum, again, we have

$$Mvb = Mvoq$$

$$\therefore \frac{v0}{v} = \frac{b}{q} = \frac{sin\theta}{1+cost\theta} \Rightarrow \frac{v_0^2}{v^2} = \frac{sin^2\theta}{(1+\cos\theta)^2} = \frac{1-cos\theta}{1+\cos\theta}$$

Comparing (1.16.2) and (1.16.4), we obtain

$$1 - \frac{2k\sin\theta}{b(1+\cos t\theta)} = \frac{1-\cos\theta}{1+\cos\theta} \Rightarrow \frac{b}{k} = \frac{\sin\theta}{\cos\theta} = \tan\theta,$$ on simplification.

$$\boxed{b = k\tan\theta = k\cot\Phi/2}$$

$\therefore$ $(\because \Phi = \pi - 20)$

Which shows that Φ is b-dependent and smaller the b, larger is Φ. But it is not possible to measure b directly and (1.16.5) cannot be verified experimentally. We therefore switch over to the next step.

Step 2. We shall now find how many α-particle in an incident parallel beam are deviated by a specified angle.

Let a plane P be at right angles to the incident beam (Fig.1.6a) and at a great distance from nucleus K. The foot of the perpendicular from K to the plane is C.

Plainly, all the α-particles that pass through a ring of P formed by two circles of radii b and b + db will be subject to a deviation between Φ and Φ + dΦ. If one particle passes per sec per unit area of P, the number of particles passing through the ring is

$$dn = 2^{\pi} b \times db$$

$$\text{where } db = d(k\cot\Phi/2) = -k d\Phi/2\sin^2(\Phi/2)$$

$$\therefore |dn| = \frac{\pi k^2 \cos(\Phi/2)}{\sin^3(\Phi/2)} |d\Phi|,$$

the number of particles deviated through Φ and Φ + dΦ, i.e., within a solid angle $d^{\Omega} = dS/r^2$, where $dS = 2^{\pi 2} \sin\Phi d\Phi$ (Fig.1.6(b)), the area on the screen impacted by scattered particles. So, W(Φ), the number of α-particles deviated per unit solid angle in $dn/d^{\Omega} = dn/2^{\pi} \sin\Phi d\Phi = db/4^{\pi} \sin(\Phi/2)\cos(\Phi/2) d\Phi$ and is the probability of deviation per unit solid angle. So, using (1.16.6)

$$\boxed{W(\Phi) = \frac{K^2}{4}\frac{1}{\sin^4(\Phi/2)} = \left(\frac{ZeE}{8\pi\varepsilon 0 Mv^2}\right)^2 \frac{1}{\sin^4(\Phi/2)}}$$

This is the α-ray scattering formula of Rutherford. It requires a correction for the motion imparted to scattering nucleus, which is appreciable only in light nuclei.

- Alternative form – If N_i monoenergetic α-particles are incident normally per unit area on a thin foil of thickness t, containing n atoms (hence n nuclei) per unit volume, the number N_Φ of scintillations produced by α-particles falling normally on unit area of the screen along Φ is

$$N_\Phi = N_i nt.W(\Phi) = N_i nt \left(\frac{ZeE}{8\pi\varepsilon 0 M v^2}\right)^2 \frac{1}{\sin^4(\Phi/2)},$$

which is just an alternative form of Rutherford's formula.

Discussion – According to Rutherford's formula, the amount of scattering (= no. of particles per unit area striking a screen at a given distance from the scatterer and normal to the direction of motion of scattered particles) should be proportional to (i) $1/\sin^4$ (Φ/2), (ii) the thickness t of the material, (iii) Z^2 and (iv) v^{-4} , that is, inversely to the square of the kinetic energy.

Each relationship between Z, M, v and Φ contained in Rutherford's equation can be tested experimentally by counting the scattered α-particles. The v-relation can be studied only over a small region as the range of velocity available with naturally occurring α-particles is small. Giger and Marsden checked (1.16.8), factor by factor, changing one factor at a time, keeping the others constant and the agreement, in general, was excellent.

- In deriving the relation (1.16.5) between the impact parameter b and the angle of scattering Φ, the following assumptions have been made.

1. The α-particles and the atomic nucleus are point charges.

2. Compared to α-particles (mass = 4 u), the gold nucleus is quite massive (mass number = 197 u) and the recoil energy energy of α-particles has been ignored.

3. The scattering is due to the electrostatic repulsion between the α-particles and the positive charge Ze of the nucleus.

4. The α-particles cannot penetrate the nuclear region and the strong interaction of nuclear forces are not considered.

- In deriving Butherford's formula, note that the we have reglected to effect of extra-nuclear electrons of the gold atom. But why? This is due to the fact that the effect of the extremely light electrons is quite negligible; the deflection of heavy α-particles due to them is only through angles $\sim$ 1°.

- An important parameter in α-ray scattering is the distance of closet approach. As the α-particles approaches for a head-on collision with the nucleus, it is slowed down due to the repulsive force of the nucleus till at the point A (Fig.1.7), whence it is turned back from the nucleus, all the kinetic energy of the particle is converted into potential energy. The distance d of this point A from the nucleus is the distance of the closet approach.

$$\therefore \frac{1}{2} Mv^2 = \frac{ZeE}{4\pi\varepsilon 0 d} = \frac{2Ze^2}{4\pi\varepsilon 0 d}$$

$$\boxed{d = \frac{Ze^2}{\pi\varepsilon 0 M v^2} = 2k}$$

- Another important parameter is the cross-section σ of scattering. The incident α-particles aimed to strike at the circumference of the circle of radius b, about the nucleus, will all get deflected through Φ (Fig.1.8). The α-particles striking within the shaded area πb^2, will get deflected by an angle > Φ. T $\boxed{\sigma = \pi b^2}$ is called the cross-section σ of scattering.

- Ratio of scattered to incident particles, N_s/N_i – Let the total target area presented by all the nuclei in the foil be A and thickness t, and t be so small that single scattering occurs.

Number of nuclei in the foil = nAt, n being the number of atoms (nuclei) per unit volume. Hence the target area for scattering by at least Φ is $N_s = \sigma(nAt)$. The total area of the foil being large compared to σ, the incident beam cannot be aimed to strike any one nucleus, and the probability that one incident particle may be scattered by an angle greater than Φ is

$$\frac{N_s}{N_i} = \frac{\sigma(nAt)}{A} = \sigma nt = \pi b^2 nt$$

Verification of Rutherford's formula

Rutherford's scattering formula has been verified over a wide range of Φ by Geiger and Marsden using various thin metallic foils.

Arrangement – The apparatus employed by them is shown schematically in Fig.1.9 C is an evacuated metallic chamber housing a radon tube S mounted in a cavity in a metal block D. The α-particles from S are collimated into a fine beam by an aperture in D and is allowed to fall normally upon the foil F. Both S and F are attached to the tube for evacuating the chamber C, and are held fixed in position throughout the experiment. The scintillations are observed by the microscope M fitted with a fluorescent screen P mounted in the focal plane of the objective. M is inserted in the wall of C that rests on a horizontal graduated platform A through greased-cone joints. C can be rotated (Fig.1.10) about a vertical axis through F without affecting the vacuum. With a given foil as scatterer, this vacuum. With a given foil as scatterer, this rotation would provide different scattering angles from 5° to 150°. The arrangement is such that the scattered α-particles always strike the screen P normally. The angle of scattering is read from the graduations of the platform A.

With a particular group of α-particles and a fixed distance between the scatterer and the foil, it appears from (1.16.8) that $N_\Phi \sin^4(\Phi/2)$ should be a constant for a given foil, N_Φ being the scintillations per sec at an angle Φ. In the table that follows, the results obtained by Geiger and Marsden are given. It bears testimony to the validity of Φ-dependence as given in (1.16.7) or (1.16.8).

Table 1.2 : Results of Geiger and Marsden

Scattering angle Φ (In degrees)	Scintillations NΦ per sec	$N\Phi \sin^4(\Phi/2)$
15	132,000	38.3
30	7,800	35.0
45	1,435	30.8
60	477	29.8
75	211	29.1
105	69.5	27.5
120	51.9	29.0
135	43.0	31.2
150	33.1	28.8

To study the dependence of scattering on the thickness of the foil, nature of the foil and also the speed of the α-particles, the apparatus as shown in Fig.1.11 is used. It is a cylindrical brass chamber B, each side being closed by glass plates P. The chamber is highly evacuated. Foils F of known thickness are mounted on a disc G which can be rotated from outside. The foils could thus be presented one after another to the normally incident beam of homogeneous α-rays from the source at O, placed just outside the chamber wall. The α-particles enter the chamber via a thin mica window W and are collimated by a diaphragm D. The scattered particles are viewed by the microscope M at the fluorescent screen S so mounted that the angle of scattering is fixed at 25°.

Results – The experimental results could be summarised as follows. The amount of scattering:

α the thickness of the foil

α (atomic number)^2,

α (velocity of particle)$^{-4}$

α $1/\sin^4(\Phi/2)$.

Rutherford’s scattering formula thus stands well verified experimentally.

---

3. 8 Liquid Drop Model

The constant density of the nuclear matter and the constant binding energy per nucleon are very similar to those found in a liquid drop. The very strong short range interaction between the nucleon permits us to consider their collective behavior in determining the properties of the nucleus. There are reasons to believe that each individual molecule within a liquid drop exerts an attractive force upon a group of molecules in its immediate neighborhood. The force of interaction does not extend to all the molecules within the drop. This is known as the saturation of the force. In order to calculate the potential of the interaction, it is necessary to know the number of interacting pairs of molecules within the drop. The binding energy BE of a nucleus is proportional linearly to the number of nucleus within it, so that the binding fraction fB is linearly constant for most nuclei. This fact shows a close resemblance of the nucleus with a liquid drop. Thus we come to the conclusion that the inter nucleon force within the nucleus attains a saturation value, so that each nucleon can interact only with a limited number of nucleon in its close vicinity. Apart from this, there are certain other points of resemblance between the nucleus of an atom and a liquid drop:

1. The nuclear force is similar to the force of surface tension on the surface of the liquid drop.

2. As in the case of a liquid drop, the density of the nuclear matter is independent of its volume. The nuclear radius R0 A1/3 where A is the mass number. Hence the nuclear volume VA. Since the nuclear mass M~A, the density of the nuclear matter ρm = M/V is independent of A.

3. The different types of particles, e.g., neutrons, protons, deuterons, α-particles etc. are emitted during nuclear reactions. These processes are analogous to the emission of the molecules from the liquid drop during evaporation.

4. The internal energy of the nucleus is analogous to the heat energy within the liquid drop.

5. The formation of the short lived compound nucleus by the absorption of a nuclear particle in a nucleus during a nuclear reaction is analogous to the process of condensation from the vapour to the liquid phase in the case of the liquid drop. The liquid drop model is not very successful in describing the low lying excited states of the nucleus. Because of the collective motions of the large number of nucleons involved, the model gives rise to closely spaced energy levels. Actually however, these are found to be quite widely spaced at low excitation energies.

Deformation of liquid drop

The fission process can be explained with the help of liquid drop model. The incident neutron combines with the nucleus to form highly energetic compound nucleus. Its extra energy is partly the kinetic energy of the neutron but largely the added binding energy of the incident neutron. This energy appears to initiate a series of rapid oscillations in the drop, which tend to distort the spherical shape so that the drop may become ellipsoidal in shape. The surface tension forces tend to make the drop return to its original spherical shape, while the excitation energy tends to distort the shape still further. If the excitation energy is sufficiently large, the drop may attain the shape of a dumb-bell. If the oscillations become so violent that the critical state, stage fourth of Figure, is reached then the final fission into stage fifth is inevitable. Thus there is a *threshold energy* or a *critical energy* required to produce stage fourth after which the nucleus cannot return to stage first. When the distortion produced is not pronounced enough to get the nucleus beyond the critical point, the ellipsoid will return to the spherical shape with the excitation energy being liberated in the form of γ-rays and we have a radiative capture rather than fission.

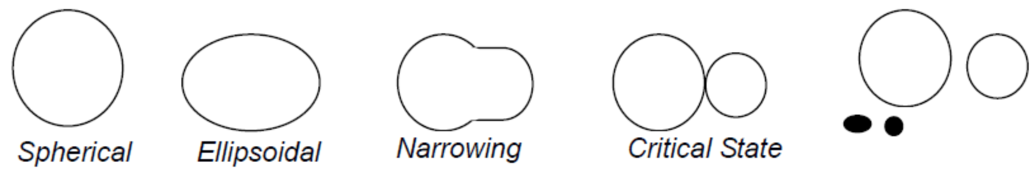


Figure : Schematic representation of nuclear fission

3.9 Shell Model of Nucleus

It is believed that protons and neutrons in a nucleus to be in a continuous process of collision with each other. With the enormous strong force acting between them and with so many nucleons to collide with, how can nucleons possibly complete whole orbits without interacting. According to Pauli's exclusion principle, no two electrons cannot occupy the same quantum state. The evidence for a kind of shell structure and a limited number of allowed energy states suggests that a nucleon moves in some kind of effective potential well created by the forces of all the other nucleons. This leads to energy quantization in a manner similar to the square well Potential . The labels on the levels are somewhat different from the corresponding symbols for atomic energy levels. The energy levels increase with orbital angular momentum quantum number l, and the s,p,d,f... symbols are used for l =0,1,2,3... just like the atomic case. But there is really no physical analog to the principal quantum number n, so the numbers associated with the level just start at n=1 for the lowest level associated with a given orbital quantum number. In addition to the dependence on the details of the potential well and the orbital quantum number, there is a sizable spin-orbit interaction which splits the levels by an amount which increases with orbital quantum number. This leads to the overlapping levels as shown in the illustration.

The subscript indicates the value of the total angular momentum j, and the multiplicity of the state is 2j + 1. The contribution of a proton to the energy is somewhat different from that of a neutron because of the coulomb repulsion, but it makes little difference in the appearance of the set of energy levels. It is found that nuclei with even numbers of protons and neutrons are more stable than those with odd numbers. In particular, there are "magic numbers" of neutrons and protons which seem to be particularly favored in terms of nuclear stability, they are : 2,8,20,28,50,82,126. Nuclei which have both neutron number and proton number equal to one of the magic numbers can be called "doubly magic", and are found to be particularly stable.

Illustrsted Examples

► Example 1. Calculate the binding energy in MeV of ^{4}He from the following data: Mass of ^{1}H = 1.008145 u and the mass of a neutron is 1.008986 u.

Solution. ^{4}He nucleus contains 2 protons and 2 neutrons. So, the mass of the constituents = 2(1.008145 + 1.008986) = 4.034262 u. But the mass of ^{4}He nucleus = 4.003875 u.

∴ Mass different (loss) = (4.034262 – 4.003875) = 0.30387 u.

∴ Binding energy, E_B = 0.030387 x 931 MeV = 28.29 MeV

► Example 2. The masses of the hydrogen atom and the neutron are 1.008142 u and 1.008982 u respectively. Calculate the packing fraction and the binding energy per nucleon of ^{16}O nucleus.

Solution. ^{16}O nucleus consists of 8 protons and 8 neutrons.

∴ Mass of constituents = 8(1.008142 + 1.008982) = 16.136992 u.

∴ Mass of ^{16}O nucleus = 15.994915 u

∴ Mass difference (loss) = 0.142077 u

∴ Binding energy, E_B = 0.142077 x 931 = 132.27 MeV

∴ Mass defect, $\Delta M' = M(A,Z) – A = 15.994915 \sim 16 = 0.05085$ u

∴ Packing fraction = $\Delta M'/A = 0.005085/16 = 3.178 \times 10^{-4}$

► Example 3. Since $^{27}_{14}\mathrm{Si}$ and $^{27}_{13}\mathrm{A1}$ are mirror nucli, their ground sates are identical except for charge. If their mass different is 6 MeV, estimate their radius (neglect the proton-neutron mass difference).

Solution. The mass-difference between mirror nuclei can be attributed to the difference in electrostatic energy. Now, the electrostatic energy of a charge Q distributed uniformly throughout a sphere of radius R is $W = 3Q^2/5R$.

$$\therefore \Delta_W = \frac{3e^2}{5R}(Z_1^2 - Z_2^2) \Rightarrow R = \frac{3e^2}{5\Delta W}(14^2 - 13^2) = \frac{3hc}{5\Delta W}\left(\frac{e^2}{hc}\right) \text{ x } 27$$

Where e^2/he = the fine structure constant = 1/137.

$$\therefore R = \frac{3 \text{ x } 1.07 \text{ x } 10^{-11}}{5 \text{ x } 6} \text{ x } \frac{1}{137} \text{ x } 27$$

$$= 0.0388 \text{ x } 10^{-11} \text{ cm} = 388 \text{ fm}$$

► Example 4. Find the value of the impact parameter and the cross section for an α-particle of energy 7.68 MeV, scattered by a thin gold foil ($Z = 79$) of thickness $6 \text{ x } 10^{-10}$ cm for a scattering angle of 90°. Also, calculate the number of α-particles incident on it. (At wt. of gold = 197.2, density of gold = 19.39 g/c.c., Avogadro number = $6.02 \text{ x } 10^{23}$).

Solution. Impact parameter, b is given by the relation

$$b = k \cot \frac{\Phi}{2} = \frac{Ze^2}{2\pi\epsilon 0 M v^2} \cot \frac{\Phi}{2}$$

$$= \frac{79 x (1.6 x 10^{-10})^2}{2 x 3.14 x 8.85 x 10^{-12} x (2.7.68 x 1.6 x 10^{-13})} \text{ x } \cot(90°/2)$$

$$= 1.48 \text{ x } 10^{-14} \text{ m, on simplification.}$$

Cross section, $\sigma = \pi b^2 = 3.14 \text{ x } (1.48 \text{ x } 10^{-14})^2 = 6.80 \text{ x } 10^{-28} \text{ m}^2$

Thickness of the foil, t = 6 x 10^{-7} m; at. wt. of gold, W = 197.2; density of gold p = 19.39 x 10^3 kg/m^3 and Avogadro number N_A = 6.02 x 10^{26} kg.mol.

Now, number of atoms per unit volume, n = $\frac{\rho N_A}{W} = \frac{19.39 \, X \, 10^3 \, X \, 6.02 \, X \, 10^{26}}{197.2}$

= 5.9 x 10^{28}

From the relation: $N_s / N_i = \sigma$nt, we have, on substitution of the values,

N_s / N_i = 6.89 x 10^{-28} x 5.9 x 10^{28} x 6 x 10^{-7}

= 2.43 X 10^{-5}.

► Example 5. An α-particles of energy 5 MeV is scattered through 180° by a fixed uranium nucleus. Calculate the distance of closet approach.

Solution. The distance of closet approach is

$$d = \frac{Ze^2}{\pi\epsilon 0 M v^2} = \frac{92 \, x \left(1.6 \, x \, 10^{-19}\right)^2}{3.14 \, x \, 8.85 \, x \, 10^{-12} \, x \, 2 \, x \, 5 \, x \, 1.6 \, x \, 10^{-13}}$$

= 5.3 x 10^{-14} m

► Example 6. In Rutherford's scattering experiment, the number of α-particles observed at an angle of 10° is 10^6 per minute. What will be the number observed at an angle of 90° and 180° respectively?

Solution. Number of particles scattered and striking unit area of the screen at an angle Φ is given by

$$N\Phi = N_i nt \left(\frac{ZeE}{8\pi\epsilon 0 M v^2}\right)^2 \frac{1}{\sin^4(\Phi/2)} = \frac{K}{\sin^4(\Phi/2)} \text{, say.}$$

When Φ = 10°, $N_{10} = \frac{K}{\sin^4(5^\circ)} \Rightarrow 10^6 = \frac{K}{(0.0871)^4}$

$\therefore$ K = 10^6 x $(0.0871)^4$ = 57.6

∴ When Φ = 90°, $N_{90} = \frac{K}{\sin^4(45^\circ)} = \frac{57.6}{(0.7071)^4}$

$= \frac{57.6}{0.25} = 230.4$

When Φ = 180°, $N_{180} = \frac{K}{\sin^4(90^\circ)} = K = 57.6$

► **Example 7.** All odd-A nuclides have a nuclear spin, I = $\frac{1}{2}$ (2n + 1) h/2$^{\pi}$, where n = 0,1,2, … and the electron, proton, neutron all have a spin angular momentum $\frac{1}{2}$ h/2$^{\pi}$. Prove that electron cannot exist inside the nucleus.

Solution. If there are electrons in the nucleus, instead of neutrons, then in the nucleus we have A protons and (A-Z) electrons. This ensures that the mass is A and the charge of the nucleus is Z (= A- $\overline{A-Z}$).

∴ Total number of particles in the nucleus = A + (A – Z) = 2A – Z.

For odd-A nuclides, we must have Z = odd, or Z = even.

When Z is odd, 2A – Z is also odd. So, the spin will be an odd multiple of $\frac{1}{2}$ h/2$^{\pi}$, i.e. I = 0,1,2, … which is not found for odd A nuclides. Hence, electrons and protons cannot together be particles inside the nucleus.

► Example 8. Find the parties of the following two functions:

$$\text{(i) } \psi(x) = \sin\left(^{\pi x/a}\right), \quad \text{(ii) } \psi(x) \cos\left(^{\pi x/a}\right)$$

Solution. (i) Given: $\psi(x) = \sin(^{\pi x/a})$, Changing x to $-x$, we have

$$\psi(-x) = \sin(-^{\pi x/a}) = -\sin(^{\pi x/a}) = -\psi(x)$$

Hence the given function has odd parity.

(ii) Given $\psi(x) = \cos\left(\pi x/a\right)$. When x is change to –x, we obtain

$$\psi(-x) = \cos\left(-\pi x/a\right) = \cos\left(\pi x/a\right) = \psi(x)$$

Hence the given function has even parity.

► Example 9. Find the mean square radius of a spherical nucleus of radius R with sharp edge.

Solution. The mean square radius is defined as

$$\langle r^2 \rangle = \int r^2 \rho(r) d^3r$$

For a spherical nucleus of radius R, with sharp edge,

$$\rho(r) = 3/4\pi R^3, \quad \text{for } r < R$$

$$= 0, \quad \text{for } r > R$$

$$\therefore \langle r^2 \rangle = \frac{3}{4\pi R^3} \int_0^R r^3 .4\pi r^2 dr = \frac{3}{5} R^2$$

✍**LEARNING ACTIVITY**

1. **Discuss in detail about Liquid drop model.**
2. **Explain the term Binding energy.**
3. **Explain the shell model in detail.**

Note:

a) Write your answer in the space given below.

b) Check the answer with your academic counsellor.

SUMMARY

In this chapter, we studied the Genreal properties of nucleus also detail study about binding energy and packing fraction. We detail studied about Nuclear stability and nuclear models.

Block IV

Radioactivity

STRUCTURE

OVERVIEW

In this chapter on `Nuclear Physics', we shall study the Radio activity and laws of disintegration . We shall also study the term Half life and mean life period of Nucleas and Particle accelerator.

LEARNING OBJECTIVES

After completing this Unit, you should be able to:

- Understand the Natural Radioactivity and properties of alpha, Beta and gamma rays.
- Derivation about laws of disintegration with examples.
- various types of particle accelerators

4.1 Introduction

Nuclear physics may be said to begin with the discovery of radioactivity. During his experiments on fluorescence of Uranium, Becqurerel in 1896 was amazed to notice that a photographic plate, wrapped in black paper for protection against radiation, was affected by the salts of Uranium kept outside it. In tracking the origin and nature of this invisible penetrating radiation, radioactivity was discovered.

4.2 Radioactivity

Radioactivity is the phenomenon of spontaneous disintegration, with emission of corpuscular or e.m. radiations of heavy atomic nuclei like U, Ra etc. at a constant rate, unaffected by any physical or chemical change or influences such as temperature, pressure etc. to which the atom (nucleus) may be subjected. It is a nuclear property of the active element and in all radioactive processes, a transmutation of the element occurs and a new nucleus (new element) is formed. Radiations from different radioactive substances were classified as α-rays and β-rays by Rutherford from a study on their penetrating power. Later, a third energetic radiation, γ-rays, was discovered by Villard. In α and β-emission processes, either Z or A or both Z and A of the nucleus change leading to the creation of a new nucleus (transmutation).

In γ-emission process, however, no transmutation i.e., no change of Z occurs, the nucleus makes only a transition from a quantum state of higher energy (excited state) to another of lower energy. Any nuclide that undergoes a change in its structure by shedding nuclear particles such as α and β, and giving off γ-rays of its own is called a radioactive nucleus. It has been found that there are 272 stable nucli of naturally occurring elements (they are non-radioactive); the rest are all unstable and hence radioactive, and are known as radio-isotopes.

● In spite of the strong nuclear forces holding the constituent nucleons, heavy elements like U, Th, Ra etc. disintegrate spontaneously giving off α, β and γ-radiations. What then is the reason behind this spontaneous decay? The cause of this disintegration is that an unstable nucleus decays to attain a new configuration which is stable.

4.3 Properties of Alpha, Beta and Gamma Rays

Properties of* α–*rays

(i) An α - particle is a helium nucleus consisting of two protons and two neutrons. It

carries two units of positive charge.

(ii) They move along straight lines with high velocities.

(iii) They are deflected by electric and magnetic fields.

(iv) They produce intense ionisation in the gas through which they pass. The ionising power

is 100 times greater than that of β-rays and 10,000 times greater than that of γ–rays.

(v) They affect photographic plates.

(vi) They are scattered by heavy elements like gold.

(vii) They produce fluorescence when they fall on substances like zinc sulphide or barium

platinocyanide.

Properties of* β *– rays

(i) β–particles carry one unit of negative charge and mass equal to that of electron.Therefore, they are nothing but electrons.

(ii) The β–particles emitted from a source have velocities over the range of 0.3 c to 0.99 c, where c is the velocity of light.

(iii) They are deflected by electric and magnetic fields.

(iv) The ionisation power is comparatively low

(v) They affect photographic plates.

(vi) They penetrate through thin metal foils and their penetrating power is greater than that of α−rays

(vii) They produce fluorescence when they fall on substances like barium platinocyanide.

Properties of* γ *– rays

(i) They are electromagnetic waves of very short wavelength.

(ii) They are not deflected by electric and magnetic fields.

(iii) They travel with the velocity of light.

(iv) They produce very less ionisation.

(v) They affect photographic plates.

(vi) They have a very high penetrating power, greater than that of β-rays.

(vii) They produce fluorescence.

(viii) They are diffracted by crystals in the same way like X−rays are diffracted.

4.4 Geiger-Nuttal Law

Geiger-Nuttal law.

The range R of an a-particle and the disintegration constant λ of the radioactive element that emits it are related as follows:

$\log \lambda = A + B \log R$.

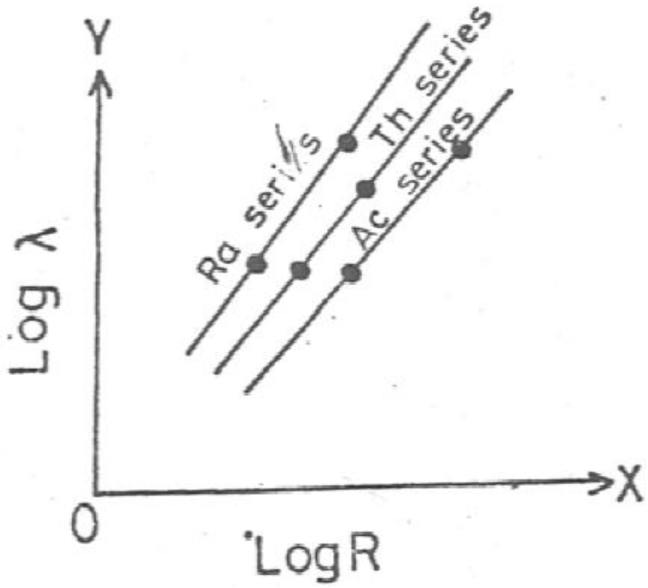


Fig.12

This relation is called Geiger-Nuttal law. If log λ are plotted against log R for the different α-emitters in the three series, three nearly parallel straight lines are obtained, one for each series (Fig. 12).

In the relation log $\lambda = A + B \log R$, the constant B is the same for all the series while A is different for the different series. According to this relation, when the disintegration constant is high, the range is also high. Since the range also depends on the energy, we conclude that radioactive substances of large decay constants emit high energy α-particles.

This law is helpful in determining roughly the decay constants of radioactive substances of very short or very long lives. Experimentally measuring the ranges of α -particles of such radioelement, the respective decay constants can be obtained by extrapolation from the curves representing the Geiger-Nuttal relation.

4.5 *Soddy Fajan's displacement law*

During a radioactive disintegration, the nucleus which undergoes disintegration is called a parent nucleus and that which remains after the disintegration is called a daughter nucleus. In 1913, Soddy and Fajan framed the displacement laws governing radioactivity.

α-decay

When a radioactive nucleus disintegrates by emitting an α-particle, the atomic number decreases by two and mass number decreases by four. The α-decay can be expressed as

$_{z}X^{A} \rightarrow {}_{z-2}YA^{-4} + {}_{2}He^{4}$

Example : Radium ($_{88}Ra^{226}$) is converted to radon ($_{86}Rn^{222}$) due to α−decay

$_{88}Ra^{226} \longrightarrow {}_{86}Rn^{222} + {}_{2}He^{4}$

β−decay

When a radioactive nucleus disintegrates by emitting a β− particle, the atomic number increases by one and the mass number remains the same. β−decay can be expressed as

$_{z}X^{A} \rightarrow {}_{Z+1}Y^{A} + {}_{-1}e^{0}$

Example : Thorium ($_{90}Th^{234}$) is converted to protoactinium ($_{91}Pa^{234}$) due to β−decay

$_{90}Th^{234} \longrightarrow {}_{91}Pa^{234} + {}_{-1}e^{0}$

At a time, either α or β−particle is emitted. Both α and β particles are not emitted during a single decay.

γ−decay

When a radioactive nucleus emits γ−rays, only the energy level of the nucleus changes and the atomic number and mass number remain the same. During α or β− decay, the daughter

nucleus is mostly in the excited state. It comes to ground state with the emission of γ−rays.

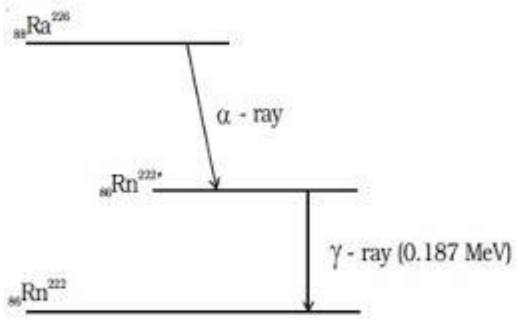


Fig Gamma decay

Example : During the radioactive disintegration of radium ($_{88}Ra^{226}$) into radon ($_{86}Rn^{222}$), gamma ray of energy 0.187 MeV is emitted, when radon returns from the excited state to the ground state

4.6 Radioactive disintegration Law

Experimental studies on radioactivity, as conducted by Rutherford and Soddy, show convincingly that

1. On emission of α or β-rays which is usually, but not invariably, accompanied by γ-emission, the emitting parent nuclide transforms into a new daughter element; the daughter element again may be radioactive so that the process of successive of disintegration continues till the original active parent nuclide gets transformed into a stable one, usually lead (Pb).
2. The rate of radioactive disintegration, that is, the number of atoms (nuclides) that disintegrate at any instant t is directly proportional to the number N_t of the active nuclides present in the sample under study at that instant.

Delay law or decay equation – Let N_i be the number of active nuclides present in the sample at any instant t. Then, we have, experimentally

$$-\frac{dNt}{dt}\ \alpha\ N_t \Rightarrow \frac{dNt}{dt} = -\lambda N_t$$

where λ, the constant of proportionality, is known as the decay constant – a characteristic constant of the element (nuclide). The negative sign indicates that N_i decreases (disintegration) with t.

Re-arranging equation, we obtain

$$\frac{dNt}{N_t} = -\lambda dt$$

Integrating equation , we have

$$\text{In}\ N_t = -\lambda t + A$$

Where A is the constant of integration.

At t = 0, $N_t = N_0$, the initial number of nuclides. So,

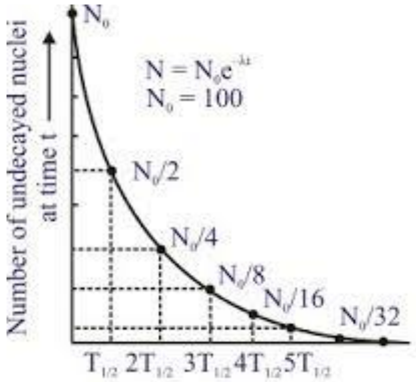
$$A = \text{In } N_0.$$

∴ we obtain finally,

$$\text{In } (N_t / N_o) = - \lambda t$$

or, $$\boxed{N_t = N_0 e^{-\lambda t}}$$

The above relation is the radioactive decay law or decay equation. It shows that the number of active nuclides decreases exponentially with time



Decay curve for a radioactive element

According to the decay law, therefore, an infinite time is theoretically needed for the complete disintegration of a radio-element and in this respect, all radio-element are the same. So, to distinguish one radio-element from another, a quantity called half-life T is more often used and we shall soon return to it.

- The significance of λ should be properly understood. For this, we write, as

$$\lambda = \frac{dN/dt}{N}$$

The r.h.s of is the number of disintegrations per nucleus per second. So, λ may be interpreted as the decay probability per nucleus per second.

Statistical nature of radioactivity:

Decay law from probability

So, long we treated N_t, the number of nuclei present at time t as a continuous variable. As long as N_t is very large compared to dN, the number of nuclei decaying during t and t + dt, this is justified and is usually the case. But, in fact, N_t varies discontinuously and the smallest value of $dN_t = 1$, corresponding to the decay of a single nucleus. Thus the decay law is valid only for a sample sufficiently large to treat dN_t as a differential. The concepts of half-life and mean life (see later) are meaningless when the sample consists of just a few nuclei. As the number of nuclei in the sample is increased, we can say that half of them will decay after one half-life, although we cannot say which particular nuclei will decay at a given instant.

The decay law is thus statistical in nature and was deduced by Schweidler assuming that the disintegration of a nuclide of a radio-element is subject to the laws of probability or chance only.

The probability p for a nuclide to disintegrate in an interval Δt depends only on the length of the interval and for sufficiently short intervals is proportional to Δt.

$$\therefore\ p = \lambda \Delta t$$

Where the constant proportionality λ is the distintegration constant.

Then the probability that the given nuclide will not disintegrate during the interval Δt is would be given by

$$1 - p = 1 - \lambda \Delta t$$

The probability that this atom will not disintegrate in a second interval Δt is also $1 - \lambda \Delta t$. The probability that this atom will survive both the intervals is therefore $(1- \lambda \Delta t)^2$. So, for n such intervals, the probability of survival of the atom is

$$(1 - \lambda \Delta t)^n = \left[1 - \lambda \left(\frac{t}{n}\right)\right]^n$$

if the total time $t = \Delta t + \Delta t + \ldots = n.\Delta t$

∴ The probability that the atom will remain unchanged after a time t is given by

$$\lim_{n\to\infty} \left(1 - \lambda \frac{t}{n}\right)^n = e^{-\lambda t}$$

One may interpret this satistically as follows:

If N_o be the initial number of nuclei, the fraction remaining unchanged after times t is $N_t/N_0 = e^{-\lambda t}$, where N_t is the number of unchanged nuclei after time t.

$$\therefore \quad \boxed{N_t = N_o e^{-\lambda t}}$$

- The decay constant is also called disintegration constant or the radioactive constant or transformation constant. It depends on the energy that is available for the nuclear transformation and on the characteristics of the parent and the daughter nuclei. It is independent of the external conditions and the age of the sample.

- Exponential growth we find that out of N_0 stoms at t = 0, the number left after time t is N_t implying that $N_0 - N_t$ atoms of parent A are converted into daughter (product) B in time t. If the daughter is non-radionactive, the rate of growth of the daughter is

$$N_0 - N_t - N_o - N_o e^{-\lambda t} = N_0 (1 - e^{-\lambda t})$$

Thus daughter product thus grows on expomentially (Fig 2.2)

Statistical fluctutation: Poisson distribution

The time t required for observing a finite number of counts N of radioactive particles (say, in a GM counter) is subject to statistical flucturations in count rate, *n* = *N/t* and (say, in a GM counter) is subject to statistical fluctuations in count rate, *n* = *N/t* and leads to an error. We shall now show that the number of counts in a given time t obeys the Poisson's distribution.

Proof. Let P_N = probability of counting N particles in time t. Divide t into n equal intervals, each equal to Δt = t/n. Now, Δt is so small (i.e., n so very large) that we may neglect the probability of emission of two particles within Δt.

∴ Probability of emission of N particles in first N intervals (n ≫ N) and none in remaining (n – N) intervals is

$$\left(\frac{N}{n}\right)^N \quad \left(1 - \frac{N}{n}\right)^{n-N}$$

Which is just one possible way to having N particles in time t.

The first particle may appear in any of n invervals, the second one in any of remaining (n – 1) intervals and so on. So finally the Nth particle may appear in any of the remaining (n – $\overline{N-1}$) or (n – N + 1) intervals. The number of ways of distributing N particles in n intervals in thus

$$N\ (n-1)\ (n-2)\ \ldots\ (n - N + 1)$$

But all of these are not mutually independent and hence, without in any way influencing the results, the position can be interchanged. The number of possible ways of interchanging the particles – N.

∴ P_N = Probability of obtaining N counts

$$= \frac{n(n-1)(n-2)\ldots(n-N+1)}{N!}$$

$$\times \left(\frac{N}{n}\right)^{N} \left(1 - \frac{N}{n}\right)^{n-N}$$

which is the binomial distribution law.

Let now n → ∞. Then reduces to

$$P_N = \frac{n^N}{N!}\left(\frac{N}{n}\right)^{N} e^{-N} = \frac{N^N e^{-N}}{N!}$$

The well-known Poisson distribution formula,

A continuous curve of P_N vs. N plotted through points gives a curve having an asymmetric maximum for small N. As N → large values, the curve turns symmetric about the maximum and approaches the Gaussian error curve.

4.7 Half Life, Mean Life periods

Half life

Half –life - The half-life or the half-value period of a radioactive nuclide is defined as the time T in which the original amount of radioactive nuclide is reduced by way of disintegration to half its value.

Substituting N_t by $N_0/2$ in (2.2.4). T is given for

$$\frac{N_0}{2} = N_0 e^{-\lambda T}$$

$$\therefore \quad \boxed{T = \frac{In2}{\lambda} = \frac{0.693}{\lambda}}$$

or, $\lambda T = 0.693 = \text{constant}$

So, T is independent of the instant from which it is measured. Depending on the active nuclide, its value may range from 10^{10} years to 10^{-7}s. The unit of T is second, the unit of λ is s^{-1}.

Mean or Average life – The fundamental law of radioactive decay, as already indicated, is a statistical law implying that the probability of decay of a given nuclide in a short time interval *dt* at time *t* is $|dN_t/N_t| = \lambda\, dt$ which is independent of the age of the nuclide. It does not state anything about the decay of an individual atom. Of the N_t nuclides, which one would disintegrate at any instant t is at the mercy of change. The decay law only states that the decay rate is proportional to N_t. Among the nuclides, some may disintegrate almost immediately, while others may exist for an infinitely long time. We do not know why it is so. The actual life a nuclide may this vary from 0 to ∞ and all radio-elements are the same in this respect. We can however speak of an average or mean life T of a radio-element.

The average or mean life τ of a radioelement is the average life time of all the atoms in the given sample and is defined as the ratio of the total life time of all the atoms or nuclei to the total number of atoms or nuclei.

$$\therefore \quad r = \frac{Total\ life\ time\ of\ all\ nuclei}{Total\ number\ of\ nuclei}$$

$$= \frac{t_1 dN_1 + t_2 dN_2 + \dots}{dN_1 + dN_2 + \dots} = \frac{\sum t dN}{\sum dN}$$

$$= \frac{\int t dN}{\int dN} = \frac{\int t dN}{-No}$$

Where dN_1 atoms have a life time t_1, dN_2 atoms a life-time t_2 and so on.

But, we have, $dN - d\,(N_0 e^{-\lambda T}) = -\lambda\, N_0 e^{-\lambda T} dt$

$\therefore\ \tau = \lambda \int_0^{\infty} t e^{-\lambda T}$ dt, substituting for *dN* in (2.3.5.)

$$= \lambda \left[-\frac{t}{\lambda} e^{-\lambda T} + \frac{1}{\lambda} \int_0^{\infty} e^{-\lambda T} dt \right]$$

$$= \lambda \left[\frac{t}{\lambda} e^{-\lambda T} - \frac{1}{\lambda 2} e^{-\lambda T} \right]_0^{\infty} = \frac{1}{\lambda}$$

$$\boxed{r = \frac{1}{\lambda} = 1.443T}$$

∴ Mean or average life,

The mean or average life of a radioactive element is thus not the same on its half life. The mean life is the reciprocal of the decay constant, that is, the decay probability per second and is greater than the half-life.

- To get the equation (2.3.5), the curve of Fig.2.4 is illuminating. The curve shows that each of dN number of active nuclei has lived a life time of t seconds, i.e., the total life span of dN nuclei is dn.t seconds.

- If t = r, then from the decay law,

$$N_t = N_o e^{-\lambda t} - N_o e^{-1} = N_0/e$$

Eq. (2.3.7) implies that the mean life is the time for the nuclei in an isotope to decay to 1/e of their original number.

- Since the actual life of an atom extends from 0 to infinity, the mean life τ is not of much physical importance. The decay constant λ and the half-life T are more significant than r.

- The decay constant, the half – life and the mean life are related to
 each other as

$$\boxed{T = \frac{0.693}{\lambda} = 0.693\,\tau}$$

Caution The decay operation N_t - $N_o e^{-1}$ applies to the behaviour of single, pure radioactive element. Mixtures of radioactive substances and radio-samples having products of initial disintegrations that are themselves radioactive do not obey the above decay law.

- As λ increase, T decreases and conversely. Also, from the decay law, as λ increase (i.e., T decreases), the number of active atoms decreases more rapidly. Conversely, if λ is small (T large), the number of active atoms decreases very slowly, e.g., uranium and theoreium. The half-lives of U and Th are very long so that their λ-values are very small.

- The λ and T are characteristic constants of a radioactive substances (Fig 2.1). The number of radioactive atoms is reduced by a factor ½ after a time 2 x T i.e., after two half-lives. In general, the number would be reduced by $\left(\frac{1}{2}\right)^n$ or $\frac{1}{2^n}$ after a time n x T , that is, after n half-lives.

Activity or strength of a radio-sample

Differentiating the decay equation $N_t = N^{oe^{-\lambda t}}$, with respect to t, we obtain

$$\frac{dN_t}{dt} = \lambda N^{oe^{-\lambda t}}$$

When t = 0, $\left(\frac{dN_t}{dt}\right)_o = \lambda N_0$. Hence, from (2.4.1) above, we have

$$\frac{dN_t}{dt} = \left(\frac{dN_t}{dt}\right)_o e^{-\lambda T}$$

$$\boxed{A_t = A_0 e^{-\lambda T}}$$

where $A_t = dN_t/dT$ and $A_0 = (dN_T/dt)_0$ = original activity = λN_0

A_t is called the activity or the strength of the sample and is proportional to the rate of disintegration.

Definiction: The acitivity or strength A_t of a radioactive sample at any instant t is defined as the number disintegrations occurring in the sample in unit time at t, that is

$$\text{Activity, } A_t = \left|\frac{dN_t}{dt}\right| = \lambda N^{oe^{-\lambda t}} = \lambda N_t = \frac{0.693}{T} N_t$$

The acitivity per unit mass of a sample is called its specific acitivity.

- A_0 [$=(dN_t/dt)_0$] is the initial activity of the same. The half life may as well be defined in terms of activity. It is the time in which the activity drops to one half of the initial activity ,

- Note that a very short lived substance gives rise to a large activity, even if it is present in minute quentites.

Units of activity – The usual unit of radioactivity is called the Curie (C_i) and is defined (since 1950) as the activity of any radioactive substance that distegrates at the rate of 3.70 x 10^{10} disintegrations per second.

A thousandth part of a Curie is called millicurie (mCi). Still smaller unit is the micro-curie (μCi). So, by definition.

1 Ci = 1 curie = 3.7 x 10^{10} disint /sec

1 mCi = 10^{-3} Curie

1 μCi = 10^{-6} Curie

The SI unit of activity is the becquerel (Bq). 1 Bq – 1 disintegration/second

Another unit of activity is the rutherford (rd), defined as the activity of a radioactive substance disintegrating at the rate of 10^6 disint /s.

- We shall now find the quantity of U-238 having 1 Curie of activity, for U-238

$$\lambda_U N_U = 1Ci = 3.70 \times 10^{10} \text{ distint/s, and}$$

$$\lambda_U = \frac{0.693}{T} = \frac{0.693}{4.5 \times 10^9 \times 365 \times 24 \times 3600} = 4.88 \times 10^{-18}$$

$$\therefore N_U = \frac{3.70 X 10^{10}}{\lambda U} = \frac{3.70 X 10^{10}}{4.88 X 10^{-18}} = 7.58 X 10^{27}$$

$\therefore$ Amount of U-238, having 1 Curie of activity is given by

$$mU = \frac{7.58 \times 10^{27} \times 238}{6.02 \times 10^{26}} = 2995 \text{ kg}$$

4.8 Law of Successive disintegration

The α-particles having been identified as helium nuclei ($_2^4$He), an α-emission by a parent melide of atomic number Z and mass number A transforms it into a daughter nuclide of atomic number (Z-2) and mass number (A-4). Hence, it becomes a new element whose position in the periodie table is two places lower. The α-emission process may be represented as

$$ {}^{A}_{Z}X \xrightarrow{\alpha} {}^{A-4}_{Z-2}Y $$

Parent element → daughter product + $\alpha-$ particle.

Few examples of alpha decay are:

$$ {}^{238}_{92}\mathrm{U} \quad \underline{\alpha} \quad {}^{234}_{90}\mathrm{Th} + {}^{4}_{2}\mathrm{He} $$

$$ {}^{226}_{88}\mathrm{Ra} \quad \underline{\alpha} \quad {}^{222}_{86}\mathrm{Rn} + {}^{4}_{2}\mathrm{He} $$

Similarly, β-particles having been identified as electrons a β-emission changes the parent nuclide (Z,A) into a daughter (Z + 1,A). The mass number in this case remains unaltered as the mass of an electron is negligible. According to the present day knowledge electron does not exist in the nucleus $\beta \rightarrow$ The neutron in the nucleus spontaneously transforms into a proton during β-emission process may be represented as

$$ {}^{A}_{Z}X \qquad {}_{Z+1}^{A}Y $$

Few examples of beta decay are:

$$ {}^{234}_{90}\mathrm{Th} \qquad {}^{234}_{91}\mathrm{Pa} + {}^{0}_{-1}\mathrm{e} $$

$$ {}^{226}_{88}\mathrm{Bi} \qquad {}^{210}_{84}\mathrm{Po} + {}^{0}_{-1}\mathrm{e} $$

The new element formed by β-emission is an isobar to the parent nucleus.

- The fact that an element comes down by two places in the periodic table by an α-emission and shoots up by one place by a β-emission constitutes the displacement law of Society and Fajan, enunciated empirically in 1913, when our present day knowledge regarding neutron proton structure of nuclei was absent.

- Note that during a radioactive transformation, the mass number and the total charge is conserved. The examples of α-decay (2.7.1) and β-decay (2.7.2) are examples of nuclear reactions. Radioactivity is entirely a nuclear phenomenon; the radioactive radiations come out of the atomic nucleus.

Radioactive series – Now, the daughter of a parent radionuclide may itself be radioactive and decay in its turn into another radionuclide and the process may be repeated till the product is stable, that is, non-radioactive. These successive transformations of a radionuclide on being studied are generally found to lie in the range of atomic number Z – 81 to Z – 92 and form what is called a radioactive series. A radioactive series is named after the longest lived member in it.

There are three naturally occurring, radioactive series showing successive transformations. These are (i) uranium series, (ii) thorium series and (iii) actinium series.

Uranium series – The uranium series starts with $^{238}_{92}\mathrm{U}$ and ends at $^{206}_{82}\mathrm{Pb}$, a stable isotope of lead. The series is also called (4n + 2) series, for the mass number of any nuclide in the series is given by A – 4n + 2, n being an integer.

Thorium series – The thorium series begins with $^{238}_{92}\mathrm{Th}$ and ends at $^{208}_{82}\mathrm{Pb}$, another stable isotope of lead. Since the mass number of any nuclide in the series is given by A – 4n, where n is an integer, it is sometimes known as 4n series and the (N, Z)-chart is given in Fig.2.7

Actinium series – The actinium series starts with actino-uraniu, that is $^{235}_{92}\mathrm{U}$ and ends at $^{207}_{82}\mathrm{Pb}$, a third stable isotope of lead. The mass number of any nuclide in the series is given by A = 4n + 3, n being an integer. The series is therefore sometimes called (4n + 3) series.

The discovery of these radioactive families is largely due to the pioneering and monumental work conducted by Soddy.

Table 2.1 : The four radioactive series

No.	Name of the series	Mass number A	Parent
1.	Thorium Series	A – 4n	$^{232}_{90}\mathrm{Th}$
2.	Uranium Series	A – 4n + 2	$^{238}_{92}\mathrm{U}$
3.	Actinium Series	A – 4n + 3	$^{235}_{92}\mathrm{U}$
4.	Neptanium Series	A – 4n + 1	$^{237}_{93}\mathrm{Np}$

This series is man-made. Other three are naturally occurring.

Neptunium series – After the discovery of nuclear fission and the production of transuranic elements (i.e., man-made elements beyond uranium, Z > 92), a fourth radioactive series called the neptunium (Z – 93) series has been observed. The series starts with plutonium, $^{241}_{94}\mathrm{Pu}$ and finishes at $^{209}_{83}\mathrm{Bi}$, a stable isotope of bismuth. It is also called (4n + 1) series as the mass number of any nuclide in the family is given by A – 4n + 1. Its (N, Z) - chart is shown in Fig.2.9.

- The several branching disintegrations in the above different series are worth noting. For example, the $^{219}_{82}\mathrm{Pb}$ according to the following two alternative transformation:

(i) $^{219}_{82}\mathrm{Pb}$

- The reason of having exactly four radioactive series is due to the fact that the alpha decay reduces the mass number of the nucleus by 4.

Successive transformation

Consider a radioactive nuclide P. symbolized by script 1, to decay into another radioactive nuclide Q (script 2); the latter again decays into a stable end-product R (script 3). For instance,

$$P(1)\ ^{\lambda 1}\ Q(2)\ \lambda 2\ R(3)\ [\text{stable}]$$

Let λ_1, λ_2 be the decay constants of nuclides 1 and 2 respectively, and N_1, N_2, N_3 be the number of atoms of the three kinds at any instant t. To determine N_1, N_2, N_3.

The second nuclide would be formed at the rate $\lambda_1 N_1$ by the decay of the parent atom and disappear by its own decay at the rate $\lambda_2 N_2$; the atoms of the of the end-product i.e., R appear at the rate $\lambda_2 N_2$ by the decay of nuclide 2, but being stable do not disappear, So, we have

$$\frac{dNt}{dt} = -\lambda_1 N_1$$

$$\frac{dN_2}{dt} = \lambda_1 N_1 - \lambda_2 N_2$$

$$\frac{dN_2}{dt} = \lambda_2 N_2$$

Solving: $(N_1)_t = N\ ^{ioe^{-\lambda 1t}}$

Where N_{10} is the number of atoms of nuclide 1 at time t = 0.

we obtain

$$\frac{dN_2}{dt} = \lambda_1 N\ ^{ioe^{-\lambda 1t}} - \lambda_2 N_2$$

$$\Longrightarrow \frac{dN_2}{dt} + \lambda_2 N_2 = \lambda_1 N\ ^{ioe^{-\lambda 1t}}$$

Multiplying both side by $e^{-\lambda 2t}$, the integrating factor,

$$e^{-\lambda 2t}\left(\frac{dN_2}{dt} + \lambda_2 N_2\right) = \lambda_1 N\ ^{ioe^{(\lambda_2 - \lambda_1)t}}$$

$$\text{or, } \frac{d}{dt}\left(N_{2e}{}^{\lambda 2t}\right) = \lambda_1\ N_{10e}{}^{(\lambda_2 - \lambda_1)t}$$

Integrating, $$N_{2e}{}^{\lambda 2t} = \frac{\lambda_1}{\lambda_2 - \lambda_1} N_{10e}{}^{(\lambda_2 - \lambda_1)t} + A$$

Where A is the constant of integration.

Now at t = 0, $(N_2)_t = N_{20} = 0$, since only the first nuclide is present. we have

$$A = -\frac{\lambda_1}{\lambda_2-\lambda_1} N_{10}$$

$$\therefore (N_2)t = \frac{\lambda_1}{\lambda_2-\lambda_1} N_{10} \left(e^{-\lambda 1t} - e^{-\lambda 2t}\right)$$

This equation is known as the Batemann equation and it gives the number of daughter atoms at time t.

From (2.8.3), using (2.8.7), we obtain on integration

$$(N_3)_t = \left(\frac{\lambda_1}{\lambda_2-\lambda_1} N_{10}\right) e^{-\lambda 2t} - \frac{\lambda_1}{\lambda_2-\lambda_1} N_{10e}{}^{-\lambda 1t} + B\lambda$$

At $t = 0$, $(N_3)_t = N_{30}\text{-}0$. So from equation (2.8.8), we get β - N_{10}

$$\therefore (N_3)_t - N_{10} \left(1 + \frac{\lambda 1}{\lambda 2-\lambda 2} e^{-\lambda 2t} - \frac{\lambda 2}{\lambda 2-\lambda 1} e^{-\lambda 1t}\right)$$

The decay of the first nuclide P, the growth and decay of the second nuclide Q and the growth of the third nuclide R

Time for Q(2) to attain a maximum – Eq. (2.8.7) gives the number of daughter atoms N_2 at time t. It shows that $N_2 = 0$ at $t = 0$. It increases with increasing t and attains a maximum at $t - t_m$, say. This t_m can be determined by imposing the condition $dN_2/dt - 0$, at $t - t_m$.

$$\therefore \frac{dN_2}{dt} = 0 = \frac{\lambda_1}{\lambda_2-\lambda_1} N_{10} \left(-\lambda i e^{-\lambda i t} + \lambda 2 e^{-\lambda 2t}\right)$$

$$\Longrightarrow \lambda 2 e^{-\lambda 2t} = \lambda 1 e$$

$$\therefore e^{(\lambda 2-\lambda r)tm} = \frac{\lambda_2}{\lambda_1} \Longrightarrow (\lambda_2 - \lambda_1)t_m = \text{In} \frac{\lambda_2}{\lambda_1}$$

$$\therefore \boxed{t_m = \frac{In\ (\lambda 2/\lambda 1)}{\lambda 2 - \lambda 1}}$$

That this t_m corresponds to maximum N_2 can be readily verified by obtaining the second time-derivative of N_2 which becomes negative at $t - t_m$.

Radioactive chain: General case

We can extend the above case of successive transformations to a radioactive chain of any number of members the so-called general case.

$$P(1) \xrightarrow{\lambda 1} Q(2)$$

Where P(1) is the parent nucleus, and Q(2), R(3), S(4) etc., are the successive products of disintegration.

If $\lambda_1, \lambda_2, \lambda_3, \ldots$ be the decay constants and $N_1, N_2, N_3, \ldots$ be the number of stone at any instant *t* of nuclides 1,2,3, etc. we have the following set of equations.

$$\left.\begin{aligned} \frac{dN_1}{dt} &= -\lambda_1 N_1 \\ \frac{dN_2}{dt} &= \lambda_1 N_1 - \lambda_2 N_2 \\ \frac{dN_3}{dt} &= \lambda_2 N_2 - \lambda_3 N_3 \\ \ldots &= \ldots\ldots\ldots\ldots\ldots \\ \frac{dN_n}{dt} &= \lambda_{n-1} N_{n-1} - \lambda_n N_n \end{aligned}\right\}$$

Bateman was the first to solve the above set of equations assuming that only the parent nuclide is present at t = 0, that is,

$$N_{20} = N_{30} = \ldots = N_{no} = 0.$$

Let us first consider the simple case of three successive disintegrations, the last (fourth) nuclide being stable.

$$\therefore \frac{dN_3}{dt} = \lambda_2 N_2 - \lambda_3 N_3$$

$$= \frac{\lambda_1 N_2}{\lambda_1 - \lambda_2} N_{10} \left(e^{-\lambda 2t} - e^{-\lambda 1t}\right) - \lambda_3 N_3, \text{ using (2.8.7)}$$

Multiplying both sides by $e^{\lambda 3t}$, we obtain

$$\frac{dN_3}{dt} e^{\lambda 3t} + \lambda_3 N_3\, e^{\lambda 3t} = \frac{\lambda_1 N_2}{\lambda_1 - \lambda_2} N_{10} \left(e^{-\lambda 2t} - e^{-\lambda 1t}\right) e^{\lambda 3t}$$

$$\text{or, } \frac{d}{dt}\left(N_3\, e^{\lambda 3t}\right) = \frac{\lambda_1 N_2}{\lambda_1 - \lambda_2} N_{10} \left(e^{-\lambda 2t} - e^{-\lambda 1t}\right) e^{\lambda 3t}$$

$$\text{Integrating } N_3\, e^{\lambda 3t} = \frac{\lambda_1 N_2}{\lambda_1 - \lambda_2} N_{10} \left\{\frac{e^{(\lambda_3 - \lambda_2)t}}{\lambda 3 - \lambda 2} - \frac{e^{(\lambda_3 - \lambda_1)}}{\lambda 3 - \lambda 1}\right\} + A$$

Where A is the constant of integration

Since $N_{30} = 0$, this at once gives

$$\frac{\lambda_1 N_2}{\lambda_1 - \lambda_2} N_{10} \frac{\lambda_2 - \lambda 1}{(\lambda_3 - \lambda_2)(\lambda_3 - \lambda_1)} + A = 0$$

whence $A = \frac{\lambda_2\, \lambda 1}{(\lambda_3 - \lambda_1)(\lambda_3 - \lambda_2)} N_{10}$

$$\therefore (N_3)_t = \lambda_1 \lambda_2 N_{10} \left\{A_{1e}^{\;-\lambda 1t} + A_{2e}^{\;-\lambda 2t} + A_{3e}^{\;-\lambda 3t}\right\}$$

Where, we have

$$\left.\begin{aligned} A_1 &= \frac{1}{(\lambda_2-\lambda_1)(\lambda_3-\lambda_1)} \\ A_2 &= \frac{1}{(\lambda_1-\lambda_2)(\lambda_3-\lambda_2)} \\ A_3 &= \frac{1}{(\lambda_1-\lambda_3)(\lambda_2-\lambda_3)} \end{aligned}\right\}$$

In the general case of n successive dintegrations, the number of atoms of nth nuclide, i.e., in the nth step, is

$$(N_n)_t = N_{10}\,(A_{1e}{}^{-\lambda 1t} + A_{2e}{}^{-\lambda 2t} + A_{3e}{}^{-\lambda 3t} + \cdots + A_{ne}{}^{-\lambda nt})$$

Where

$$\left.\begin{aligned} A_1 &= \frac{\lambda_1\lambda_2\lambda_3 \ldots \lambda_n-1}{(\lambda_2-\lambda_1)(\lambda_3-\lambda_1)\ldots(\lambda_n-\lambda_1)} \\ A_2 &= \frac{\lambda_1\lambda_2\lambda_3 \ldots \lambda_n-1}{(\lambda_1-\lambda_2)(\lambda_3-\lambda_2)\ldots(\lambda_n-\lambda_2)} \\ \ldots &= \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ A_3 &= \frac{\lambda_1\lambda_2\lambda_3 \ldots \lambda_n-1}{(\lambda_1-\lambda_n)(\lambda_2-\lambda_n)\ldots(\lambda_{n-1}-\lambda_n)} \end{aligned}\right\}$$

Radioactive equilibrium

An equilibrium, by definition, refers to the condition in which the time-derivative of a function vanishes.

Applying the above condition to different radioactive nuclides of a chain, it would mean that the number of atoms of a nuclide in the chain does not change with time. Mathematically, from the set of equation (2.9.1), the conditions of equilibrium are:

$$\left.\begin{aligned} &\frac{dN_1}{dt} = -\lambda_1 N_1 - 0 \\ &\lambda_1 N_1 = \lambda_2 N_2 \\ &\lambda_2 N_2 = \lambda_3 N_3 \\ &\ldots = \ldots\ldots\ldots\ldots\ldots \\ &\lambda_{n-1} N_{n-1} = \lambda_n N_n \end{aligned}\right\}$$

Secular equilibrium – Plainly, the first equation of the above set (2.10.1) is a contradiction, for it implies $\lambda_1 = 0$. So the condition of equilibration cannot be satisfied rigorously.

A state very close to equilibrium however can be attained if the parent substance is very long-lived. i.e., decays much more slowly compared to the other nuclides in the chain. N_1 then can be taken as nearly constant and $\lambda_1 \ll$ other λ's in the chain. Under this condition, the first equation of set (2.10.1) is a good approximation but the other conditions hold good rigorously. This type of equilibrium is called secular or permanent equilibrium and is given by the condition

$$\lambda_1 N_1 = \lambda_2 N_2 = \dots \lambda_{n-1} N_{n-1} = \lambda_n N_n$$

or,

$$\boxed{\frac{N_1}{(T)_1} = \frac{N_2}{(T)_2} = \cdots = \frac{N_n}{(T)_n}}$$

where $(T)_i$ is the half-life of the ith member of the given chain.

The relation shows that at secular equilibrium, the rate of decay of any radioactive product is just equal to its rate of production from the previous member in the chain.

4.9 Particle Accelerator

A particle accelerator is a device used to accelerate the charged particles, which are required in the study of artificial transmutation of elements. Hence the accelerator is the basic device in high energy particle physics.

Accelerators can be divided broadly into two types.

(i) The first type belongs to electrostatic accelerators in which charged particles are accelerated by applying a constant potential difference. The final energy is determined by the amount of the potential difference. The Cockcroft – Walton and Van de Graaff generators belong to this class. These accelerators can accelerate particles only upto a few million electron-volts.

(ii) The second type is the cyclic or synchronous accelerator, which accelerate particles in multiple steps imparting a small energy in each successive step. Example : Linear accelerator, cyclotron, betatron, synchrocyclotron and synchrotron. They can accelerate particles to energy of the order of 10^9 eV (GeV), of course, with each machine having its own energy limit.

Accelerators

Accelerators are some of the most remarkable tools of modern science. They are precision instruments constructed on a gargantuan scale. They have to track and accelerate particles that traverse millions of km in just matters of seconds, and maneuver and constrain particle motion to accuracies of the order of 1 /zm. They can provide sufficient numbers of energetic particles to vaporize macroscopic targets with single pulses of beam. Because of their immensity and their challenging complexity, and because of their symbolic reflection of the intellectual aspirations and creativity of mankind, modern accelerators have been likened by Robert R. Wilson to the great Gothic cathedrals of medieval Europe. Their impact on nuclear and particle physics, the fields for which they were initially developed, has, of course, been pivotal. They have served as the microscopes for probing nuclear and particle structure, and, in fact, were it not for the development of accelerators, the fields of nuclear and particle physics would still be in their infancies.

After the pioneering experiments of Rutherford and his colleagues revealed the presence of a nucleus within the atom, it became clear that higher-energy scattering experiments could provide an invaluable probe of the nucleus. For example, it was understood that with sufficient energy to penetrate the Coulomb barrier, projectiles could break apart nuclei and reveal their constituents. It was also recognized that the more energy a particle had, the more deeply it could probe within the nucleus. This is simply a consequence of the fact that, through the uncertainty principle, large momentum transfers correspond to small distances, and vice versa. The study of the short-distance behavior of nuclei and of elementary par- tides requires therefore the availability of high energy beams that can be used to impart large momentum transfers either to target particles or to other beams of particles.

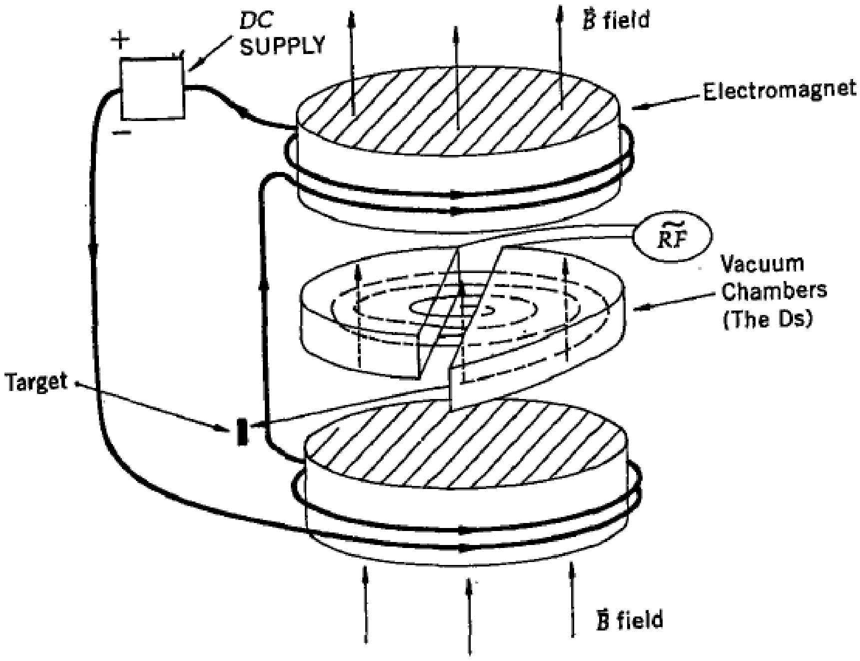
Although high energy particles are available in the cosmic rays, their fluxes are quite low, and their energies, clearly, cannot be controlled. In fact, the excitement brought about by discoveries of new phenomena in experiments with cosmic rays only added impetus to the development of techniques for accelerating charged particles.

The increase in accelerator energies achieved over the past 75 years has been astounding. The first accelerators that were constructed (around 1930) provided beams of particles with energies of hundreds of keV, while the largest modern accelerators will soon have beam energies of almost 10^8 greater than that. And because of the advent of colliding-beam techniques, the effective increase in beam energy (that is, considering the energy available in the center-of-mass) has gone up by an even more spectacular factor of about 10^{12}! Such changes correspond to differences of about 10^6 in the sensitivity to distance scales that can be studied with the next-generation accelerators, which expect to provide sensitivity in the range of $10^{\sim 18}$ cm. Nowadays, besides being used in nuclear and particle physics, accelerators are in demand in a variety of applications, ranging from experiments in condensed matter physics, the electronics industry, biomedical and geophysical areas, to food processing and sewage treatment. Accelerator science is therefore no longer just an appendage of nuclear and particle physics, but is a separate intellectual discipline in its own right.

There is a variety of ways of accelerating charged particles, and the methods used for any specific application depend upon the kinds of probes that are required, their energies, the desired beam intensities, and, of course, any economic constraints. We will now sketch several of the key historical developments in particle acceleration during the past 75 years.

Cyclotron

Fixed-voltage machines have an inherent limitation to their energy because of voltage breakdown and discharge. An alternative method, which uses the resonance principle, is more important for accelerating particles to higher energy. The cyclotron (or cyclic accelerator), first built by Ernest Lawrence, is the simplest of the machines that use this principle (see Fig.). The accelerator is constructed out of two hollow evacuated D-shaped metal chambers (referred to as Ds), which are connected to an alternating highvoltage source. The entire system is placed inside a strong magnetic field perpendicular to the Ds. The principle of operation of the cyclotron is as follows.



Although the hollow Ds are connected to the source of high voltage, because of the shielding effect of the metallic chamber walls, there is no electric field within the Ds. Consequently, a strong alternating electric field exists only in the gap between the Ds. A source for producing ions is placed in the gap between the Ds, and, depending on the sign of the voltage at that moment, any ion in the gap is attracted towards one of the Ds. However, the trajectory of the ions is circular because of the bending effect of the magnetic field. Once an ion is inside the D, it stops sensing the electric force, but continues in its circular motion because of the presence of the static magnetic field. But after a half circle, when the ion is about to emerge from the D, the direction of the voltage can be changed and the ion can be accelerated again before it enters the other D. Similarly, when it is about to exit from the second D, the applied voltage can again be reversed and the particle accelerated further. If the frequency of the alternating voltage source is just right, then the charged particle can be accelerated continuously and move in ever increasing radial orbits, until it is extracted to strike a target (for example, by suddenly turning off the *B* field).

For non-relativistic motion, the frequency appropriate for the alternating voltage can be calculated from the fact that the magnetic force provides the centripetal acceleration for a circular orbit. That is,

$mv^2r = qvB/C$

$v/r = qB/mc$

Now, for circular motion at constant speed, the angular frequency *u* is related to the radius and circular velocity of the orbit as

$$\omega = \frac{v}{r}.$$

We can therefore express the frequency of the motion as

$$\nu = \frac{\omega}{2\pi} = \frac{qB}{2\pi mc} = \frac{1}{2\pi}\left(\frac{q}{m}\right)\frac{B}{c}.$$

Clearly, to keep the acceleration in phase with the particle motion requires that the frequency of the electric field be the same as v. This frequency is referred to as the *cyclotron resonance frequency*, and is the origin of the label "resonance accelerator" for this kind of machine. Equation (8.3) therefore provides a means for determining the frequency of the accelerating fields as a function of other parameters. The maximum energy that a charged particle has when it is extracted at a radius $r = R$ is given by

$$T_{\max} = \frac{1}{2}mv_{\max}^2 = \frac{1}{2}m\omega^2R^2$$

$$= \frac{1}{2}m\left(\frac{qB}{mc}\right)^2R^2 = \frac{1}{2}\frac{(qBR)^2}{mc^2}$$

Equation (8.4) relates the magnitude of the magnetic field and the size of the magnet that is needed to accelerate a particle to any given energy. In a typical cyclotron, $B < 2$ T , the alternating voltage applied to the Ds is « 200 kV, at a frequency of « 10-20 MHz. The maximum proton energy that can be attained in such cyclotrons is about 20 MeV (for Ds of R « 30 cm), as will be shown shortly in an example. As we increase the energy of charged particles, they become relativistic, and the frequency relation in Eq. (8.3) starts failing. Consequently, a fixed-frequency cyclotron cannot accelerate ions to relativistic energies. For electrons, relativistic effects set in at even lower energies, and consequently, such simple cyclotrons are not useful for accelerating electrons.

Linac or Linear Accelerator

Linear accelerators, as the name implies, accelerate particles along linear trajectories rather than in circular orbits. These accelerators are also based on the resonance principle, and operate as follows. A series of metal tubes, called drift tubes, are located in a vacuum vessel and connected successively to alternate terminals of a radio frequency oscillator, as shown in Fig.

Let us suppose that at some time the fields are as shown in the figure. Positive ions from the source will then be accelerated by the electric field towards the first drift tube. If the alternator can change its direction before the ions pass through that tube, then they will be accelerated again on their way between the exit of the first and entry into the second tube, and so on. However, as the particles accelerate, their velocities increase, and consequently, if the drift tubes are all of the same length, the phase between the particle positions and the potentials at the next tube may not keep in step (that is, the next gap may not accelerate). To avoid this, the drift tubes are made longer along the path so that one radio-frequency (RF) alternator can accelerate the particles all the way to the end.

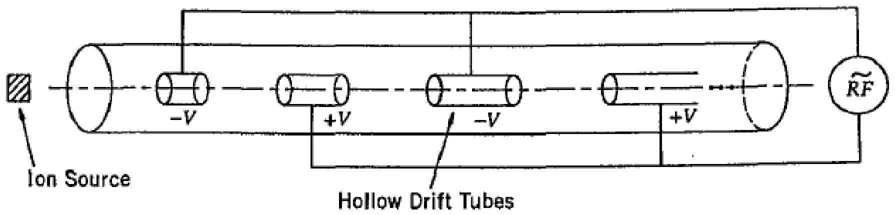


Fig. Linear ion accelerator

Because electrons become relativistic at relatively low energies, electron linear accelerators act on a slight variation of the principle just described. The electron source is usually a hot wire filament which, effectively, boils off electrons. These are accelerated through a positive potential grid and rapidly become relativistic. Bunches of these electrons are then passed through accelerating tubes that are fed with microwave power delivered by klystron amplifiers. Electrons radiate easily as they get accelerated (this is referred to as *synchrotron radiation)*, and therefore much power is needed to increase their energy. This power is supplied *by* the microwave fields that travel in step with electrons in specially shaped iris-loaded waveguides. The longest linear accelerator is the two-mile Stanford Linac (SLAC), and it accelerates electrons to energies of 50 GeV.

Betatron

Electrons may be accelerated to high energies by having them move in a circular path of constant radius and at the same time increasing the magnetic flux through the circular orbit in such a way that the electrons acquire additional energy during each revolution. Such an accelerator is known as a *betatron.*

The first successful betatron was designed and built by D. W. Kerst (1940) from the theory worked out by Kerst and R. Serber. The original betatron,as shown in Figure 12-5, which accelerated electrons up to 2.3 MeV, was operated as an x-ray tube; the x-rays were produced in the conventional manner by allowing the high-energy electrons to strike a target. Most betatrons built since then have also been used as x-ray sources, but some have been used as sources of high-energy electrons for nuclear experiments. Betatrons that accelerate electrons up to about 300 MeV have been built.

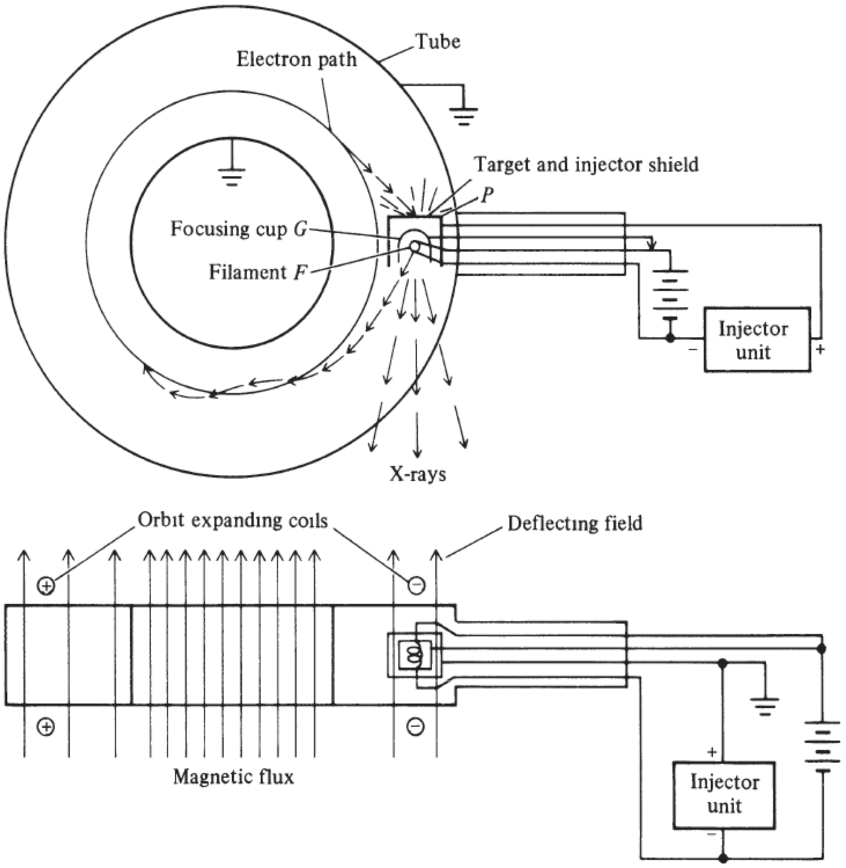


Fig. Path of an electron in betatron

In the operation of the betatron electrons from the heated filament Fare injected into the circular or doughnut-shaped tube by applying a difference of potential between the filament and the plate *P*, as shown in Figure. The electrons are focused with the aid of the grid G. When an alternating magnetic field is applied parallel to the axis of the tube, two effects are produced: an electromotive force is produced in the electron orbit by the changing magnetic flux that gives the electrons additional energy; a radial force is produced by the action of the magnetic field whose direction is perpendicular to the electron velocity which keeps the electron moving in a circular path. The magnetic flux through the orbit has to be chosen in such a way that the electrons will move in a stable orbit of fixed radius *R*. The electrons make several hundredthousand revolutions in this circular path while the alternating magnetic field is increasing in intensity from zero to a maximum-that is, during a quarter of a cycle. With each revolution they gain additional energy. When the electrons have acquired the desired amount of energy, a capacitor is discharged through two coils of wire, one directly above and the other directly below the stable orbit, thus producing a sudden addition to the magnetic flux. This destroys the condition for the stability of this orbit and the electron beam moves out to larger radii until it strikes the back of the irtiector *P* which acts as the x-ray target.

We can think of the circular electron path of fixed radius *R* as a circuit; the emf *V* induced in this circuit by the changing magnetic flux is, according to Faraday's law,

$V= d¢/dt$

where ¢ is the instantaneous value of the magnetic flux which is perpendicular to the plane of the circuit. The work done on an electron of charge *e* in one revolution is therefore

$Ve=e\ d¢/dt$

This work can also be expressed in terms of the tangential force *F* which, acting on the electron over a distance *ds*, does an amount of work *dW* given by

$dW=F\cdot ds$

from which $F=dW/ds$

Thus the tangential force acting on the electron is equal to the work done per unit length of path. Evaluating this force for one revolution for which the path length is *2nR*, we get

$$F = \frac{Ve}{2\pi R} = \frac{e}{2\pi R}\frac{d\Phi}{dt}$$

Now, from Newton's second law,

$F=dp/dt$

hence

$$\frac{dp}{dt} = \frac{e}{2\pi R}\frac{d\Phi}{dt}$$

Or

$$dp = \frac{e}{2\pi R}\, d\Phi$$

Because of the presence of the magnetic flux perpendicular to the plane of the electron orbit, the electron will experience a radial force inward given by

$$Bev = \frac{m\gamma v^2}{R}$$

where "y = $(1 - V^2/C^2)^{-1/2}$ and *B* is the value of the magnetic induction at the electron orbit of constant radius *R.* From the above equation

$myv = BeR$

$p=BeR$

If *R* is kept constant, then differentiation of this equation yields

$dp = eR\, dB$

Comparing Equations, we see that

$$\frac{e}{2\pi R}\, d\Phi = eR\, dB$$

from which

$$d\Phi = 2\pi R^2\, dB$$

Integrating this equation between the limits of zero and *B*, respectively, we get

$$\Phi = 2\pi R^2 B$$

for the instantaneous relationship between the total magnetic flux cP and the magnetic induction *B* at a distance *R* from the center. This equation shows that the magnetic flux within the orbit of radius *R* is always proportional to the intensity of the magnetic field at the orbit and, furthermore, that the magnetic flux through the orbit is twice what it would have been if the magnetic induction were uniform throughout the orbit at the value *B.* This distribution of magnetic flux is obtained in an air gap between specially shaped pole faces of an electromagnet.

Most modern betatrons are operated from a 60-Hz ac source. Since the magnet and its coils constitute a large inductance, a very large capacitance is introduced into the circuit to bring the power factor closer to unity for efficient operation. In a 1 OO-Me V betatron the electrons are accelerated during a quarter of a cycle-that is, during 1/240 sec. The energy acquired by an electron per revolution is 400 e V; hence it has to make 250,000 revolutions to acquire the maximum energy. In practice the energy of the electrons, hence that of the x-ray photons, can be varied from about 10 to 100 MeV by applying the orbit-shifting magnetic field at different times during the quarter-cycle that the field is increasing.

✍LEARNING ACTIVITY

1. **What is natural radioactivity.**
2. **State Laws of disintegration .**
3. **Derive an expression for Half life and Mean life of Nuclease**
4. **Describe the construction and working of Linear accelerator.**

Note:

a) Write your answer in the space given below.

b) Check the answer with your academic counsellor.

SUMMARY

In this chapter, we studied the radioactivity and properties of alpha, beta and gamma rays. Also detail studied about laws of disintegration and its types. We detail studied about half life and mean life and different types of particle acceletaors.

Block V

NUCLEAR REACTION

STRUCTURE

OVERVIEW

In this chapter you will be learning about types of nuclear reactions and conservation always connected to nuclear reaction. Also it will be discussed in this lesson about the theories of nuclear reactions. Cross section of nuclear reaction, compound nuclear formation and statistical theory connect to compound values.

LEARNING OBJECTIVES

After completing this Unit, you should be able to:

- Classification of Nuclear reactions
- Energy of nuclear reaction – Q value and artificial radioactivity.
- Distinguish Between Nuclear fission and Nuclear fusion

5.1 Introduction

A nuclear reaction is a process that takes place when a nuclear particle such as proton, neutron, deuteron, a-particle etc comes in close contact (with in 10-15m) with another, and energy and momentum exchanges occur. The final products of the reactions are also some nuclear particle or particles that leave the point of contact in different directions. The process results in the *transmutation* of target nucleus. The changes that occur in a nuclear reaction usually involve strong nuclear force. Process that involve nuclear interactions (e.g. b-decay) are purely electromagnetic (e.g. coulomb scattering) are not included under nuclear reactions but changes of nuclear states due to electromagnetic interactions are included. A general equation representing a nuclear reaction is of the form

$$ {}^{A}_{Z}X + {}^{A}_{Z}x \rightarrow {}^{A'}_{Z'}Y + {}^{A'}_{Z'}y $$

$$ or \quad {}^{A}X(x, y)^{A'}Y $$

Where X is the target nucleus, x the bombarding particle, Y is residual product nucleus and y the ejected particle. In nuclear reaction charge number, mass number, total energy etc are conserved.

5.2 Types of Nuclear Reaction

A) Elastic scattering: The incident particle strikes the target nucleus and leaves without energy loss, but scattered in different directions. The residual nucleus is the same as the target nucleus and is left in the same state (ground state) and is represented as X(x, y) X. Scattering of a particles in gold is an example of this process.

Example:

$${}_2He^4 + {}_{79}Au^{197} \rightarrow {}_{79}Au^{197} + {}_2He^4$$

B) Inelastic Scattering: Here same particles are scattered in different directions with different energy, as there is loss of energy due to collision. The residual nucleus, which is the same as target nucleus, is left in an excited state and the process is represented as X(x, y) X*.

Example:

$${}_3Li^7 + {}_1H^1 \rightarrow \left({}_3Li^7\right)^* + {}_1H^1$$

The star is used to indicate that, after scattering nucleus is left in an excited state. In the above example the excess energy is radiated away in the form of a g-quantum.

C) Disintegration: In scattering the target nucleus the incident particle is absorbed and a different particle is ejected. The product nucleus differs from target nucleus. The incident particle may be an a-particle, proton, neutron etc. The product particle may be a charged particle or a neutron.

Example:

$${}_7N^{14} + {}_2He^4 \rightarrow {}_8O^{17} + {}_1H^1$$

D) Photo- Disintegration: Here the g-rays are absorbed by the target nucleus, exciting to a higher quantum state. If the energy is high enough, one or more particles may be liberated.

Example:

$${}^1_2H + \gamma \rightarrow {}^1_1H + {}^1_0n$$

E) Radioactive Capture: The incident particle is absorbed or captured by the target nucleus to form the excited compound nucleus which disintegrates to produce one or more g-photons and goes down to the ground state.

Example:

$$^{26}_{12}Mg + ^{1}_{1}H \rightarrow \left(^{27}_{13}Al\right)^{*} \rightarrow ^{27}_{13}Al + \gamma$$

The process may be represented as X(x, r) Y*

F) Direct Reaction: A collision of an incident particle with the nucleus may immediately pull one of the nucleons out of the target nucleus and is called *pick up reaction.* In the reverse process, a bombarding particle composed of more than one nucleon may loss one of them to the target by *stripping reaction.*

Example:

$$^{63}_{29}Cu + ^{2}_{1}d \rightarrow ^{64}_{29}Cu + ^{1}_{1}H$$

G) Nuclear Fission: In heavy nuclei, neutrons produce fission reaction in which the nuclei disintegrate into two nuclei of comparable size along with the emission of neutrons and release of energy.

There are many other types of nuclear reactions such as elementary particle reaction involving production of elementary particles many body reaction in which two or more particles are emitted by the compound nucleus, nuclear fusion in which light nuclei combine to form a heavier nucleus etc.,

5.3 Energy of nuclear Reactions: Q value

It was found that in conservation of mass-energy, neither the kinetic energy nor the rest mass is conserved by itself but their sum is always conserved. In the derivations of energy relations in low-energy nuclear physics the non-relativistic approximation is adequate.

Consider the nuclear reaction X(x, y) Y represented by the equation

X + x = Y + y

Where X is the target nucleus of mass Mx, Y the product or recoil nucleus of mass My, x the incident particle of mass mx and y the product particle of mass my.

Also let K_X, K_Y, Kx, Ky are the K.E. of X, Y, x , y. Assume that X is initially at rest so that $K_X = 0$. The total energy of a particle is the sum of its kinetic energy and rest energy.

As energy is conserved, we have

$M_X c^2 + (Kx + mx\, c^2) = (K_Y + M_Y c^2) + (Ky + my\, c^2)$

Or

$K_Y + Ky – Kx = (M_X + Mx – M_Y – my)\, c^2$

(1)

Where, c is the velocity of light in free space. The quantity on the L.H.S. of equation (1) is the difference between the kinetic energies of the products of the reaction and that of the interacting particles. This is denoted by Q which is equal to the difference of the rest mass energies of the interacting particles and that of the product particles. Here

$Q = K_Y + Ky - Kx$

$= (M_X + mx – M_Y – my)\, c^2$

(2)

$= M_X + mx – M_Y – my$, energy units

The quantity Q is known as the energy balance of the nuclear reaction, or Q-value. It is nothing but the net surplus or deficit of energies of the reaction products $K_Y + Ky$ over the energy supplied Kx.

Three possibilities:

1) Exoergic reaction: From eqn (2)

if $(M_X + Mx) > (M_Y + My)$ then $(K_Y + Ky) > Kx$ and $Q > 0$.

Such a nuclear reaction is said to be exoergic or exothermic reaction because it release energy. In this case there is a net decrease in the mass after the nuclear reaction and the lost mass gets converted into surplus energy according to the relation $E = mc^2$. This energy is imparted as K.E. to the product particles.

2) Endoergic reaction:

If $(M_X + Mx) < (M_Y + My)$, then $(K_Y + Ky) < Kx$ and $Q < 0$.

Such a nuclear reaction is said to be endoergic or endothermic reaction because some energy is to be supplied from outside for the reaction to occur. In exoergic reaction, there is a net increase in the mass after the nuclear reaction and the increase in mass is provided by supplying energy to the incident particle so that the reaction proceeds in the forward direction. In endoergic reaction, a minimum amount of energy equal to the negative Q-value of the reaction must be given to the incident particle so that the reaction occurs. Such an energy is called the threshold energy of the endoergic reaction.

3) Q-Value:

The third possibility is that Q = 0 which occurs if $(M_X + Mx) = (M_Y + My)$ that is, the sum of he masses of the interacting particles is equal to the sum of the rest masses of the product particle. It gives rise to the elastic collision where there is no loss of energy but only change in the direction of the particles.

Hence nuclear reactions involve information about the nuclear masses and the particle energies and can be employed to derive information about the mass of nuclei, particle energies or Q-values. The quantity KY, the recoil energy of the product nucleus is usually small and difficult to measure. It can be eliminated by applying the conservation of momentum to the reaction.

5.4 Artificial radioactivity

Artificial transmutation

Artificial transmutation is the conversion of one element into another by artificial methods. The first successful artificial transmutation was carried out by Rutherford. When nitrogen was bombarded with α-particles of sufficient energy, a rare isotope of oxygen ($_8O^{17}$) and a proton were formed.

$$_7N^{14} + {}_2He^4 \rightarrow {}_8O^{17} + {}_1H^1$$

This process is called nuclear reaction.

In a nuclear reaction, the sum of initial atomic numbers (charges) is equal to the sum of the final atomic numbers, i.e conservation of charges is satisfied. In the same way, conservation of nucleons is also satisfied. But the initial rest mass is not equal to the final rest mass. The energy equivalence of the difference between the rest masses is equal to the nuclear reaction energy Q. The nuclear reaction can be written as

$$A + a \rightarrow B + b + Q$$

where A is the target nucleus and a, the bombarding particle. B and b are the product nucleus and particle.

Artificial radioactivity

Artificial radioactivity or induced radioactivity was discovered by Irene Curie and F. Joliot in 1934. This is also known as man-made radioactivity. The phenomenon by which even light elements are made radioactive by artificial or induced methods is called artificial radioactivity.

Curie and Joliot showed that when lighter elements such as boron and aluminium were bombarded with α-particles, there was a continuous emission of radioactive radiations, even after the α–source had been removed. They showed that the radiation was due to the emission of a particle carrying one unit positive charge with mass equal to that of an electron. This particle is known as positron and the reaction is

$$_5B^{10} + {}_2He^4 \rightarrow {}_7N^{13*} + {}_0n^1$$

The nitrogen atom is radioactive and decays with a half life of about 10.1 minutes into a stable isotope of carbon with the emission of a positron

$$_7N^{13*} \longrightarrow {}_6C^{13} + {}_1e^0$$

In the case of Aluminium,

$$_{13}A^{l27} + {}_2He^4 \longrightarrow {}_{15}P^{30*} + {}_0n^1$$

Radio-phosphorous decays into a stable isotope of silicon with the emission of a positron. The half life of radioactive phosphorous is about 3 minutes.

$_{15}P^{30*} \longrightarrow {}_{14}Si^{30} + {}_1e^0$

Artificial radioactivity follows the same laws of natural radioactivity. Artificial radioactive elements emit electrons, positrons, neutrons and γ rays.

Production of artificial radio-isotopes

(i) Artificial radio-isotopes are produced by placing the target element in the nuclear reactor, where plenty of neutrons are available.

(1)$_{15}P^{31} + {}_0n^1 \longrightarrow {}_{15}P^{32*} + \gamma$, and (2) $_{11}Na^{23} + {}_0n^1 \longrightarrow {}_{11}Na^{24*} + \gamma$

(ii) Another method of production of artificial radio-isotope is to bombard the target element with particles from particle accelerators like cyclotron.

${}_{11}Na^{23} + {}_{1}H^{2} \rightarrow {}_{11}Na^{24*} + {}_{1}H^{1}$

Applications of radio-isotopes

The radio-isotopes have wide applications in medicine, agriculture, industry and research. A radio-isotope is added to a particular system and the course of the isotope is studied to understand the system.

(i) Medical applications

In medical field, radio-isotopes are used both in diagnosis and therapy. Radio cobalt (Co^{60}) emitting γ–rays is used in the treatment of cancer. Gamma rays destroy cancer cells to a greater extent.

Radio-sodium (Na^{24}) is used to detect the presence of blocks in blood vessels, to check the effective functioning of heart in pumping blood and maintaining circulation. Radio-iodine (I^{131}) is used in the detection of the nature of thyroid gland and also for treatment. Radioiodine is also used to locate brain tumours.

Radio-iron (Fe59) is used to diagnose anaemia. An anaemic patient retains iron in the blood longer than normal patient. Radio-phosphorous (P^{32}) is used in the treatment of skin diseases.

(ii) Agriculture

In agriculture, radio-isotopes help to increase the crop yields. Radio-phosphorous (P^{32}) incorporated with phosphate fertilizer is added to the soil. The plant and soil are tested from time to time. Phosphorous is taken by the plant for its growth and radio-phosphorous is found to increase the yield. Sprouting and spoilage of onions, potatoes, grams etc. Are prevented by exposure to a very small amount of radiation. Certain perishable cereals remain fresh beyond their normal life span when exposed to radiation.

(iii) Industry

In Industry, the lubricating oil containing radio-isotopes is used to study the wear and tear of the machinery.

(iv) Molecular biology

In molecular biology radio-isotopes are used in sterilizing pharmaceutical and surgical instruments.

(v) Radio-carbon dating

In the upper atmosphere, C^{14} is continually formed from N^{14} due to the bombardment by neutrons produced from cosmic rays.

${}_7N^{14} + {}_0n^1 \rightarrow {}_6C^{14*} + {}_1H^1$.

The C^{14} is radioactive with half life of 5570 years.

The production and the decay of C^{14} are in equilibrium in atmosphere. The ratio of C^{14} and C^{12} atoms in atmosphere is $1 : 10^6$. Hence, carbon dioxide present in the atmosphere contains a small portion of C^{14}.

Living things take C^{14} from food and air. However with death, the intake of C^{14} stops, and the C^{14} that is already present begins to decay. Hence the amount of C^{14} in the sample will enable the calculation of time of death i.e, the age of the specimen could be estimated. This is called radio-carbon dating. This method is employed in the dating of wooden implements, leather clothes, charcoal used in oil paintings, mummies and so on.

Biological hazards of nuclear radiations

When γ-ray or any high energy nuclear particle passes through human beings, it disrupts the entire normal functioning of the biological system and the effect may be either pathological or genetic. The biological effects of nuclear radiation can be divided into three groups

(i) Short term recoverable effects

(ii) long term irrecoverable effects and

(iii) genetic effect

The extent to which the human organism is damaged depends upon

(i) the dose and the rate at which the radiation is given and

(ii) the part of the body exposed to it.

Smaller doses of radiation exposure produce short term effects such as skin disorder and loss of hair. If the exposure is 100 R*, it may cause diseases like leukemia (death of red blood corpuscle in the blood) or cancer. When the body is exposed to about 600 R, ultimately it causes death. Safe limit of receiving the radiations is about 250 milli roentgen per week. The genetic damage is still worse. The radiations cause injury to genes in the reproductive cells. This gives rise to mutations which pass on from generation to generation.

The following precautions are to be taken for those, who are working in radiation laboratories.

(1) Radioactive materials are kept in thick−walled lead container.

(2) Lead aprons and lead gloves are used while working in hazardous area.

(3) All radioactive samples are handled by a remote control process.

(4) A small micro−film badge is always worn by the person and it is checked periodically for the safety limit of radiation.

5.5 Nuclear fission

In 1939, German scientists Otto Hahn and F. Strassman discovered that when uranium nucleus is bombarded with a neutron, it breaks up into two fragments of comparable masses with the release of energy.

The process of breaking up of the nucleus of a heavier atom into two fragments with the release of large amount of energy is called nuclear fission. The fission is accompanied of the release of neutrons. The fission reactions with 92U235 are represented as

$_{92}U^{235} + {}_0n^1 \rightarrow {}_{56}Ba^{141} + {}_{36}Kr^{92} + 3\ {}_0n^1 + Q$ ...(1)

$_{92}U^{235} + {}_0n^1 \rightarrow {}_{54}Xe^{140} + {}_{38}Sr^{94} + 2\ {}_0n^1 + Q$ ...(2)

In the above examples the fission reaction is taking place with the release of 3 neutrons and 2 neutrons. On an average 2.5 neutrons per fission are actually released. Of the many possible fission reactions of $_{92}U^{235}$, the reaction given in equation (1) is the most favourable.

Energy released in fission

Let us calculate the amount of energy released during the fission of $_{92}U^{235}$ with a neutron. The fission reaction is

$_{92}U^{235} + {}_0n^1 \rightarrow {}_{56}Ba^{141} + {}_{36}Kr^{92} + 3{}_0n^1 + Q$

Mass of $_{92}U^{235} = 235.045733$ amu

Mass of ${}_0n^1$ = 1.008665 amu

Total mass of the reactant = 236.054398 amu

Mass of ${}_{56}Ba^{141}$ = 140.9177 amu

Mass of ${}_{36}Kr^{92}$ = 91.8854 amu

Mass of 3 neutrons = 3.025995 amu (3 × 1.008665)

Total mass of the products = 235.829095 amu

∴ Mass defect = 236.054398 – 235.829095 = 0.225303 amu

As, 1 amu = 931 MeV, energy released in a fission = 0.225303 × 931 = 200 MeV

Niels Bohr and John A. Wheeler explained the nuclear fission process with the help of liquid drop model. A liquid drop has a spherical shape due to surface tension. On applying external force the sphere changes into ellipsoid, which may change into a dumb bell shape when the force is larger. This may break at the narrow end into two portions.

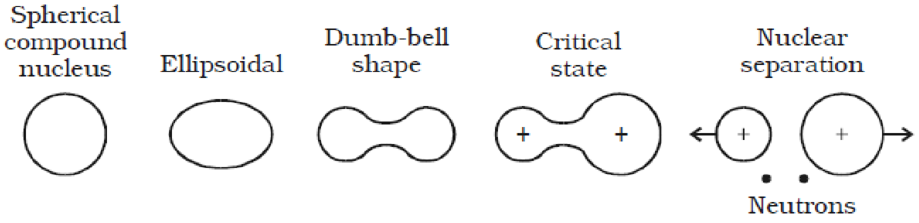


Fig 8.7 Bohr – Wheeler's explanation of Nuclear fission

In the same way, when the heavier nucleus absorbs a neutron, a compound nucleus is formed and is left in an excited state. The excitation energy sets up a series of rapid oscillations. The compound nucleus undergoes distortion from spherical to dumb bell shape. Each portion of the dumb bell has a positive charge and one repels the other. This results in fission and the formation of fission fragments (Fig)

5.6 *Chain reaction*

Consider a neutron causing fission in a uranium nucleus producing three neutrons. The three neutrons in turn may cause fission in three uranium nuclei producing nine neutrons. These nine neutrons in turn may produce twenty seven neutrons and so on. A chain reaction is a self propagating process in which the number of neutrons goes on multiplying rapidly almost in a geometrical progression.

Two types of chain reactions are possible. In the uncontrolled chain reaction, the number of neutrons multiply indefinitely and the entire amount of energy is released within a fraction of a second. This type of chain reaction takes place in atom bombs.

In the controlled chain reaction the number of fission producing neutron is kept constant and is always equal to one. The reaction is sustained in a controlled manner (Fig). Controlled chain reaction is taking place in a nuclear reactor.

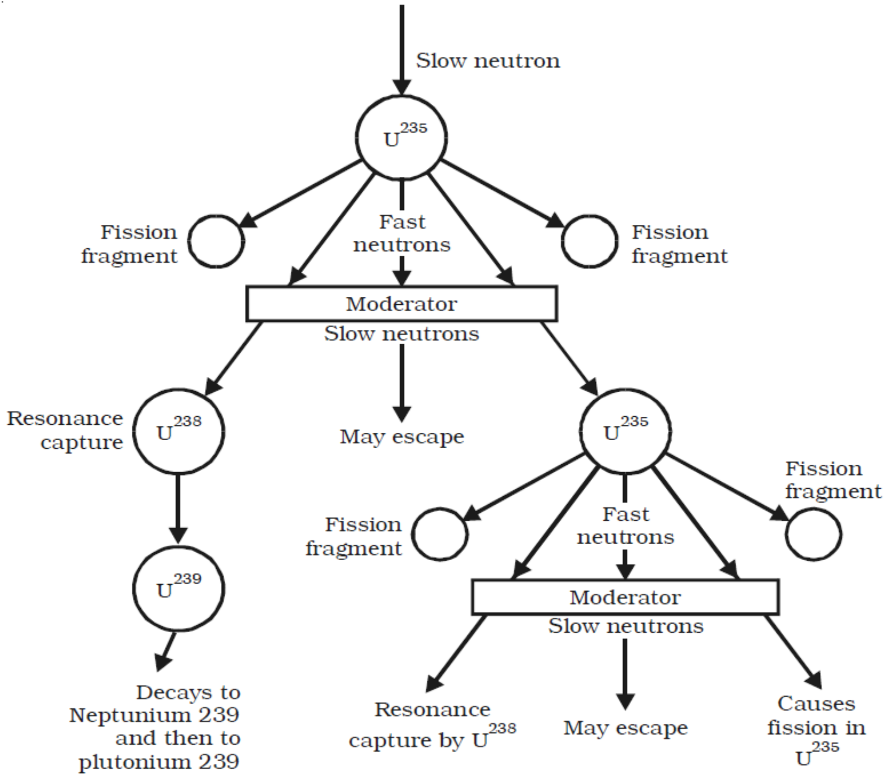


Fig Controlled chain reaction

When a thermal neutron bombards U^{235} nucleus, it breaks into two fission fragments and three fast neutrons (Fig 8.8). One neutron may escape and one neutron may be captured by U^{238} which decays to Np^{239} and then to Pu^{239}. One neutron is available for carrying out chain reaction. The chain reaction is possible, only when the loss of neutrons is less than the neutrons produced.

Critical size

Critical size of a system containing a fissile material is defined as the minimum size in which atleast one neutron is available for further fission reaction. The mass of the fissile material at the critical size is called critical mass. The chain reaction is not possible if the size is less than the critical size.

Atom bomb

Atom bomb is based on the principle of uncontrolled fission chain reaction. Natural uranium consists of 99.28% of U^{238} and 0.72% of U^{235}. U^{238} is fissionable only by fast neutrons. Hence, it is essential in a bomb that either U^{235} or Pu^{239} should be used, because they are fissionable by neutrons of all energies. An atom bomb consists of two hemispheres of U^{235} (or ${}_{94}Pu^{239}$), each smaller than the critical size and are kept apart by a separator aperture (Fig).

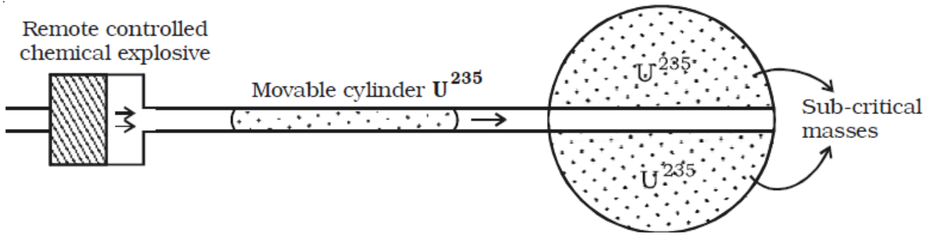


Fig Atom bomb

When the bomb has to be exploded, a third well fitting cylinder of U^{235} (or ${}_{94}Pu^{239}$) whose mass is also less than the critical mass, is propelled so that it fuses together with the other two pieces. Now the total quantity is greater than the critical mass and an uncontrolled chain reaction takes place resulting in a terrific explosion.

The explosion of an atom bomb releases tremendous amount of energy in the form of heat, light and radiation. Temperature of millions of degree celsius and pressure of millions of atmosphere are produced. Such explosions produce shock waves. The release of dangerously radioactive γ–rays, neutrons and radioactive materials produce a health hazards over the surroundings for a long time. These bombs were used in world war II and were exploded over Hiroshima and Nagasaki in Japan. We are quite familiar with the amount of destruction and hazard.

5.7 *Nuclear reactor*

A nuclear reactor is a device in which the nuclear fission reaction takes place in a self sustained and controlled manner. The first nuclear reactor was built in 1942 at Chicago USA.

Depending on the purpose for which the reactors are used, they may be calssified into research reactors, production reactors and power reactors. Research reactors are used primarily to supply neutrons for research purpose and for production of radio-isotopes. The purpose of production reactors is to convert fertile (non-fissile but abundant) material into fissile material. The power reactor converts nuclear fission energy into electric power. The power reactors can be further classified into boiling water reactor, pressurised water reactor, pressurised heavy water reactor and fast breeder reactor depending upon the choice of the moderator and the coolant used. Numerous reactors of different designs have been constructed all over the world for a variety of purposes, but there exists a number of general features common to all the reactors. The schematic diagram of a nuclear reactor is shown in Fig .

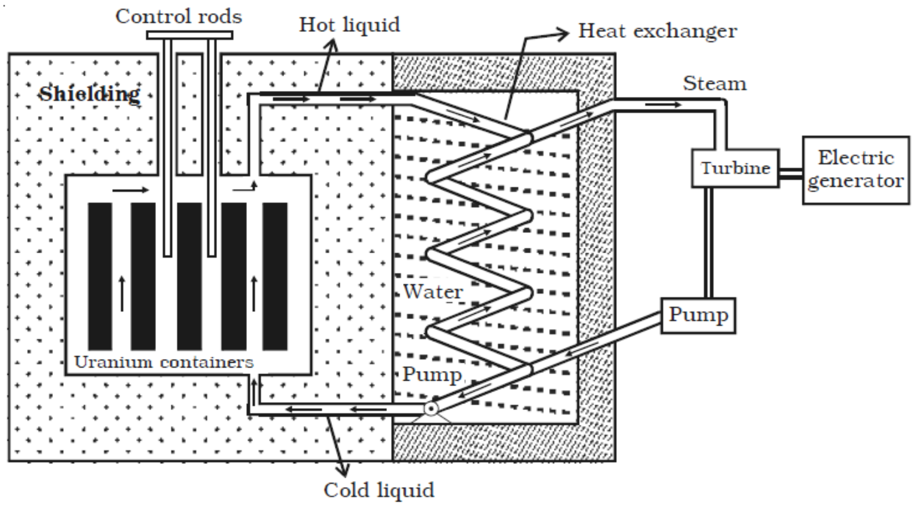


Fig Nuclear reactor

Fissile material or fuel

The fissile material or nuclear fuel generally used is $_{92}U^{235}$. But this exists only in a small amount (0.7%) in natural uranium. Natural uranium is enriched with more number of $_{92}U^{235}$ (2 – 4%) and this low enriched uranium is used as fuel in some reactors. Other than U^{235}, the fissile isotopes U^{233} and Pu^{239} are also used as fuel in some of the reactors.

In the pressurised heavy water reactors (PHWR) built in our country, natural uranium oxide is used as fuel. Tiny pellets of uranium oxide are packed in thin tubes of zirconium alloy and sealed to form a fuel rod. Nineteen such rods are tied together to form a fuel bundle. The reactor vessel which goes by the name 'Calandria' has about three hundred tubes passing through it. The fuel bundles are placed in these 10^8 tubes. The part of the reactor vessel which contains the fuel rods is known as reactor core. In the pressurised light water reactors (PWR), low enriched uranium is used. In the fast breeder test reactor (FBTR) at Kalpakkam, a mixture of the carbides of uranium and plutonium is used as fuel.

The fuel of the prototype fast breeder reactor (PFBR) that is being built at Kalpakkam, is a mixture of oxides of plutonium and uranium. In the fast breeder reactors, the fuel pellets are packed in special stainless steel tubes to withstand the high temperature.Kamini [Kalpakkam mini reactor] is the only operating reactor in the world which uses ${}_{92}U^{233}$ as fuel. In this reactor, the fuel is an alloy of uranium and aluminium and is in the form of plates.

(ii) Moderator

The function of a moderator is to slow down fast neutrons produced in the fission process having an average energy of about 2 *MeV* to thermal neutrons with an average energy of about 0.025 *eV*, which are in thermal equilibrium with the moderator. Ordinary water and heavy water are the commonly used moderators. A good moderator slows down neutrons by elastic collisions and it does not remove them by absorption. The moderator is present in the space between the fuel rods in a channel. Graphite is also used as a moderator in some countries. In fast breeder reactors, the fission chain reaction is sustained by fast neutrons and hence no moderator is required.

(iii) Neutron source

A source of neutron is required to initiate the fission chain reaction for the first time. A mixture of beryllium with plutonium or radium or polonium is commonly used as a source of neutron.

(iv) Control rods

The control rods are used to control the chain reaction. They are very good absorbers of neutrons. The commonly used control rods are made up of elements like boron or cadmium. The control rods are inserted into the core and they pass through the space in between the fuel tubes and through the moderator. By pushing them in or pulling out, the reaction rate can be controlled. In our country, all the power reactors use boron carbide (B_4C), a ceramic material as control rod.

(v) The cooling system

The cooling system removes the heat generated in the reactor core. Ordinary water, heavy water and liquid sodium are the commonly used coolants. A good coolant must possess large specific heat capacity and high boiling point. The coolant passes through the tubes containing the fuel bundle and carries the heat from the fuel rods to the steam generator through heat exchanger. The steam runs the turbines to produce electricity in power reactors.

The coolant and the moderator are the same in the PHWR and PWR. In fast breeder reactors, liquid sodium is used as the coolant. A high temperature is produced in the reactor core of the fast breeder reactors. Being a metal substance, liquid sodium is a very good conductor of heat and it remains in the liquid state for a very high temperature as its boiling point is about 1000^{o} C.

(vi) Neutron reflectors

Neutron reflectors prevent the leakage of neutrons to a large extent, by reflecting them back. In pressurised heavy water reactors the moderator itself acts as the reflector. In the fast breeder reactors, the reactor core is surrounded by depleted uranium (uranium which contains less than 0.7% of ${}_{92}U^{235}$) or thorium (${}_{90}Th^{232}$) which acts as neutron reflector. Neutrons escaping from the reactor core convert these materials into Pu^{239} or U^{233} respectively.

(vii) Shielding

As a protection against the harmful radiations, the reactor is surrounded by a concrete wall of thickness about 2 to 2.5 m.

Breeder reactor

${}_{92}U^{238}$ and ${}_{90}Th^{232}$ are not fissile materials but are abundant in nature. In the reactor, these can be converted into a fissile material ${}_{94}Pu^{239}$ and ${}_{92}U^{233}$ respectively by absorption of neutrons. The process of producing more fissile material in a reactor in this manner than consumed during the operation of the reactor is called breeding. A fast reactor can be designed to serve as a good breeder reactor.

Uses of reactors

(1) Nuclear reactors are mostly aimed at power production, because of the large amount of energy evolved with fission.

(2) Nuclear reactors are useful to produce radio-isotopes.

(3) Nuclear reactor acts as a source of neutrons, hence used in the scientific research.

5.8 Nuclear fusion

Nuclear fusion is a process in which two or more lighter nuclei combine to form a heavier nucleus. The mass of the product nucleus is always less than the sum of the masses of the individual lighter nuclei. The difference in mass is converted into energy. The fusion process can be carried out only at a extremely high temperature of the order of 10^7 K because, only at these very high temperatures the nuclei are able to overcome their mutual repulsion. Therefore before fusion, the lighter nuclei must have their temperature raised by several million degrees. The nuclear fusion reactions are known as thermo-nuclear reactions.

Hydrogen bomb

The principle of nuclear fusion is used in hydrogen bomb. It is an explosive device to release a very large amount of energy by the fusion of light nuclei. The temperature required for the purpose of fusion is produced by fission reactions. The explosion of an atom bomb produces temperature of the order of 50 million degree celcius. A suitable assembly of deuteron and triton is arranged at the sight of the explosion of the atom bomb. Favourable temperature initiates the fusion of light nuclei in an uncontrolled manner. This releases enormous amount of heat energy.

The fusion reaction in hydrogen bomb is

${}_1H^3 + {}_1H^2 \rightarrow 2He^4 + {}_0n^1 + \text{energy}$

5.9 Thermo Nuclear Reactions

Stellar energy

Fusion is the source of stellar energy. The temperature of the interior of the sun is about 1.4×10^7 K and the temperature of some stars is of the order 10^8 K. It has been estimated that the total energy radiated by sun is about 3.8×10^{26} joule per second. The origin of such a tremendous amount of energy is neither chemical nor gravitational. The fusion of protons into helium is supposed to release energy in sun and stars. All the elements like hydrogen and helium are in a highly ionised state called plasma at such a high temperature. The energy produced in fusion is responsible for the maintenance of high temperature of the sun and stars and also for the emission of energy in the form of heat and light.

Proton−proton cycle and carbon−nitrogen cycle are the two important types in which nuclear fusion takes place in sun and stars.

Proton – Proton cycle

${}_1H^1 + {}_1H^1 \rightarrow {}_1H^2 + {}_1e^0 + \nu$ (emission of positron and neutrino)

${}_1H^1 + {}_1H^2 \rightarrow {}_2He^3 + \gamma$ (emission of gamma rays)

$2\ {}_2He^3 \rightarrow {}_2He^4 + 2{}_1H^1$

The reaction cycle is written as

$4{}_1H^1 \rightarrow {}_2He^4 + 2{}_1e^0 + 2\nu + \text{energy (26.7 MeV)}$

Thus four protons fuse together to form an alpha particle and two positrons with a release of large amount of energy.

Carbon – Nitrogen Cycle

The following cycle of reactions take place in carbon – nitrogen cycle in which carbon acts as a catalyst.

${}_{1}H^{1} + {}_{6}C^{12} \longrightarrow {}_{7}N^{13*} + \gamma$ (emission of gamma rays)

${}_{7}N^{13*} \longrightarrow {}_{6}C^{13} + {}_{1}e^{0} + \nu$ (emission of positron and neutrino)

${}_{1}H^{1} + {}_{6}C^{13} \longrightarrow {}_{7}N^{14} + \gamma$ (emission of gamma rays)

${}_{1}H^{1} + {}_{7}N^{14} \longrightarrow {}_{8}O^{15*} + \gamma$ (emission of gamma rays)

${}_{8}O^{15*} \longrightarrow {}_{7}N^{15} + {}_{1}e^{0} + \nu$ (emission of positron and neutrino)

${}_{1}H^{1} + {}_{7}N^{15} \longrightarrow {}_{6}C^{12} + {}_{2}He^{4}$

The overall reaction of the above cycle is given as

$4\ {}_{1}H^{1} \longrightarrow {}_{2}He^{4} + 2\,{}_{1}e^{0} + 2\nu$ + energy (26.7 MeV)

✍**LEARNING ACTIVITY**

1. **What is meant by Nuclear fission.**
2. **What is Nuclear fusion.**
3. **Explain the construction and working of Nuclear reactor.**
4. **Explain the term thermo nuclear reaction with example..**

Note:

a) Write your answer in the space given below.

b) Check the answer with your academic counsellor.

SUMMARY

In this chapter, we studied the types of Nuclear reaction and its Q value. Also, detail studied about Artificial radioactivity, Radio Isotopes and its applications. We detail studied about nuclear fission and fusion with examples and also derive the thermo nuclear reactions.