MSc. Previous Year

Physics

MP – 02

CLASSICAL AND
STATISTICAL MECHANICS

मध्यप्रदेश भोज (मुक्त) विश्वविद्यालय – भोपाल
MADHYA PRADESH BHOJ (OPEN) UNIVERSITY – BHOPAL

Reviewer Committee

1. Dr. P.K. Agrawal
   Director
   IEHE College, Bhopal (M.P.)

2. Dr. S.K.Vijay
   Professor
   Govt. MLB College, Bhopal (MP)

3. Dr. Seema Hardikar
   Professor
   Sarojini Naidu Govt. Girls PG (Autonomous)
   College, Bhopal (MP)

• •

Advisory Committee

1. Dr. Jayant Sonwalkar
   Hon'ble Vice Chancellor
   Madhya Pradesh Bhoj (Open) University,
   Bhopal

2. Dr. L.S. Solanki
   Registrar
   Madhya Pradesh Bhoj (Open) University, Bhopal

3. Dr. Neelam Wasnik
   Asst. Director
   Madhya Pradesh Bhoj (Open) University, Bhopal

4. Dr. P.K. Agrawal
   Director
   IEHE College, Bhopal (M.P.)

5. Dr. S.K. Vijay
   Professor
   Govt. MLB College, Bhopal (MP)

6. Dr. Seema Hardikar
   Professor
   Sarojini Naidu Govt Girls PG (Autonomous)
   College, Bhopal (MP)

• •

COURSE WRITERS

Dr. Nikhil Ranjan Roy, Former Professor and Head, Department of Physics, Ranchi University
Units: (1.0-1.2, 1.3, 1.4, 1.6, 1.7, 1.9, 1.11, 1.12-1.16, 2.0-2.6, 2.8-2.13, 3.0-3.3, 3.5-3.9, 4.1, 4.3, 4.7-4.11, 5.0-5.2, 5.9, 5.15-5.19)

Dr Shikha Bansal, Assistant Professor, Centre for Nanoscience and Nanotechnology, Jamia Millia Islamia, New Delhi
Dr Manika Khanuja, Assistant Professor, Centre for Nanoscience and Nanotechnology, Jamia Millia Islamia, New Delhi
Units: (2.7, 3.4, 4.2)

Dr Partha Pratim Das, Assistant Professor, School of Applied Sciences & Humanities, Haldia Institute of Technology Haldia, Purba Medinipur (W.B.)
Abhisek Chakraborty, Scientist, Space Application Centre (ISRO), Ahmedabad (Gujrat)
Units: (4.0, 4.5, 5.3-5.4, 5.8, 5.14)

Dr Samiksha Sirakwar, Assistant Professor Integrated Basic Science School for Physical and Decision Science, Babasaheb Bhimrao Ambedkar University (A Central University), Lucknow.
Dr Bhupendra Pratap Singh, Former Faculty Deptt of Physics University of Lucknow
Units: (1.2.1, 1.4.1, 1.5, 1.8, 1.10, 1.11.1, 4.4, 4.6, 5.5-5.7, 5.10-5.13)

SYLLABI-BOOK MAPPING TABLE
Classical and Statistical Mechanics

Syllabi	Mapping in Book
UNIT I: LARGRANGIAN MECHANICS • Constraints, Generalised coordinates, D'Alembert Principle and derivation of Lagrangian equation, velocity dependent potentials and Rayleigh's dissipation function. • Variational Principle, Euler-Lagrange equation, Derivation Lagrange's equation from Hamilton principle • Two-body central force problem, Kepler's problem, inverse square law of force. Scattering in a central force field	**Unit-1:** Lagrangian Mechanics **(Pages 3-48)**
UNIT II: HAMILTON MECHANICS AND RIGID BODY • Derivation of Hamilton's equation from variational principle, principles of least action, Equations of canonical transformation, Lagrangian and Poisson Motion in Poisson bracket notation. • Euler's equations of motion for a rigid body and its applications to torsion free symmetric rigid body.	**Unit-2:** Hamilton Mechanics and Rigid Body **(Pages 49-99)**
UNIT III: HAMILTON-JACOBI THEORY • Hamilton-Jacobi equation for Hamilton's principle function. Harmonic oscillator problem using Hamilton Jacobi method. Hamilton-Jacobi equation for Hamilton's charactristics functions. Separation of variables in the Hamilton-Jacobi equation. Action-angle variables. Kepler's problem in action angle variables.	**Unit-3:** Hamilton-Jacobi Theory **(Pages 101-130)**
UNIT IV: CLASSICAL STATISTICAL MECHANICS • A priori probability, phase, space, liouville's theorem; Statistical equilibrium, Maxwell Boltzman distribution law of velocity, Equation of Energy Micro Canonical and Grand Canonical ensembles Canonical Partition function thermodynamic functions in different ensembles	**Unit-4:** Classical Statistical Mechanics **(Pages 131-147)**
UNIT 5: Quantum Statistical Mechanics • Bose-Einstein Statistics, Blackbody radiation energy, and pressure of an ideal Bose Gas, Einstein condensation, theory of liquid helium, Fermi-Dirac Statistics, Energy and Pressure of ideal Fermi Gas, Free Electron theory of solids, Landau Theory of Phase transition, critical indices, scale transformation, dimensional analysis, Density and Energy function with electron spin in hydrogen like atom.	**Unit-5:** Quantum Statistical Mechanics **(Pages 149-227)**

CONTENTS

INTRODUCTION

Classical mechanics describes the motion of macroscopic objects, from projectiles to parts of machinery, and astronomical objects, such as spacecraft, planets, stars and galaxies. If the present state of an object is known it is possible to predict by the laws of classical mechanics how it will move in the future (determinism) and how it has moved in the past (reversibility). The earliest development of classical mechanics is often referred to as Newtonian mechanics. It consists of the physical concepts employed by and the mathematical methods invented by Isaac Newton and Gottfried Wilhelm Leibniz and others in the 17th century to describe the motion of bodies under the influence of a system of forces. Later, more abstract methods were developed, leading to the reformulations of classical mechanics known as Lagrangian mechanics and Hamiltonian mechanics. These advances, made predominantly in the 18th and 19th centuries, extend substantially beyond Newton's work, particularly through their use of analytical mechanics. They are, with some modification, also used in all areas of modern physics. Classical mechanics provides extremely accurate results when studying large objects that are not extremely massive and speeds not approaching the speed of light. When the objects being examined have about the size of an atom diameter, it becomes necessary to introduce the other major subfield of mechanics: quantum mechanics. To describe velocities that are not small compared to the speed of light, special relativity is needed. In case that objects become extremely massive, general relativity becomes applicable. However, a number of modern sources do include relativistic mechanics into classical physics, which in their view represents classical mechanics in its most developed and accurate form.

This book, *Classical and Statistical Mechanics* is divided into five units that follow the self-instruction mode with each unit beginning with an Introduction to the unit, followed by an outline of the Objectives. The detailed content is then presented in a simple but structured manner interspersed with Check Your Progress Questions to test the student's understanding of the topic. A Summary along with a list of Key Terms and a set of Self-Assessment Questions and Exercises is also provided at the end of each unit for recapitulation.

UNIT 1   LAGRANGIAN MECHANICS

Structure

1.0   INTRODUCTION

Swiss mathematician Leonhard Euler and Italian-French mathematician Joseph Louis Lagrange evolved Lagrange's equation in relationship with their discussion of the tautochrone problem. They used Lagrange's method to mechanics, which led to the formation of Lagrangian mechanics. Lagrange's equation is a second-order partial differential equation whose solutions are the functions for which a given functional is stationary. In this unit you will study Lagrange's equations for simple systems and important properties of the Lagrangian function. You will also learn principle of virtual work and D'Alembert's principle along with the derivation of Lagrange's equation for general and conservative system.

1.1   OBJECTIVES

After going through this unit, you will be able to:

- Explain Lagrange's equations for simple systems
- Describe gauge function, law of inertia, central force and virtual displacement
- Understand D' Alembert's principle and principle of virtual work
- Derive Lagrange's equations for general and conservative systems
- Discuss applications of Lagrangian formulation

Lagrangian Mechanics

- Describe important properties of the Lagrangian function
- Explain Kepler's laws of planetary motion

1.2 CONSTRAINTS IN LAGRANGIAN MECHANICS

The work of a force on a particle along a virtual displacement is known as the virtual work. The principle of virtual work explains that in equilibrium the virtual work of the forces applied to a system is zero.

Virtual Displacement

Consider a system of N particles 1, 2, ….., N having s degrees of freedom. Let q_1, q_2, ….., q_s be the generalized coordinates that describe the system. The configuration space of the system is s-dimensional. At any instant of time t, the configuration of the system is specified by a point in the configuration space, the point being defined by a particular set of values for the generalized coordinates.

Let the system be subjected to arbitrary displacement in the configuration space consistent with the constraints imposed on the system at the instant. The corresponding change in the configuration of the system is independent of time, i.e., no actual displacement of the system occurs with respect to time. Such displacements in the configuration space are called **virtual displacements**. It is usual to denote virtual displacement of the generalized coordinates, say q_k, as δq_k.

The concept of virtual displacement has been found useful for mathematical analysis of the properties of mechanical systems.

Virtual Work

Consider the system of N particles described above. Let $\vec{F_1}, \vec{F_2}, ..., \vec{F_N}$ be the forces acting on the particles. If we consider the system to be in equilibrium, we have

$$\vec{F_k} = 0 \quad \left(k = 1, 2, ..., N \right) \tag{1.1}$$

Let $\vec{r_1}, \vec{r_2}, ..., \vec{r_N}$ be the equilibrium position vectors of the particles.

Let $\delta\vec{r_1}, \delta\vec{r_2}, ..., \delta\vec{r_N}$ be the infinitesimal virtual displacements of the particles from their equilibrium positions.

We then have according to Equation (1.1)

$$\sum_{k=1}^{N} \vec{F_k} \cdot \delta \vec{r_k} = 0 \qquad (1.2)$$

However, if the forces $\vec{F_k}$ are continuous functions of positions then the left hand side of Equation (1.2) can be interpreted as the net work done in the virtual displacements of the particles. If the system changes from its equilibrium configuration, we may write Equation (1.2) as

$$\delta W = 0 \qquad (1.3)$$

The result given by Equation (1.3) is referred to as the **principle of virtual work.**

Let us consider the presence of constraints in the system. We then have the force on any particle as a vector sum of the applied force $\vec{F}^{(a)}$ and the force of constraint $\vec{F}^{(c)}$. Thus, we get

$$\vec{F_k} = \vec{F_k}^{(a)} + \vec{F_k}^{(c)} \qquad (1.4)$$

Equation (1.2) then becomes

$$\sum_{k=1}^{N} \left(\vec{F_k}^{(a)} + \vec{F_k}^{(c)} \right) \cdot \delta \vec{r_k} = 0$$

or $\qquad \sum_{k=1}^{N} \vec{F_k}^{(a)} \cdot \delta \vec{r_k} + \sum \vec{F_k}^{(c)} \cdot \delta \vec{r_k} = 0 \qquad (1.5)$

Unlike in Equation (1.2), the left hand side of Equation (1.5) can not be interpreted as the net work done in the virtual displacements of the particles of the system. This is because forces of constraints $\vec{F_k}^{(c)}$'s are not continuous functions of positions.

Let us restrict our considerations to only such systems for which

$$\vec{F_k}^{(c)} \cdot \delta \vec{r_k} \geq 0 \qquad (1.6)$$

For all δr_k which are consistent with the constraints, from Equations (1.5) and (1.6), we get

$$\sum_k \vec{F_k}^{(a)} \cdot \delta \vec{r_k} \leq 0 \qquad (1.7)$$

The only forces involved in Equation (1.7) are the applied forces which may be considered as continuous functions of positions, in general. We are then in a position to interpret the left hand side of Equation (1.7) as the net work done by the applied forces during the virtual displacements of the particles consistent with the constraints and express Equation (1.7) as

$$\delta W = \sum \vec{F}_k^{(a)} \cdot \delta \vec{r}_k \leq 0 \qquad (1.8)$$

Let the virtual displacements under consideration be restricted to displacements which are reversible in the geometrical sense. Denoting reversible displacements by $\delta' \vec{r}_k$ we get from Equation (1.7)

$$\sum \vec{F}_k^{(a)} \cdot \delta' \vec{r}_k \leq 0$$

Also, $$\sum \vec{F}_k^{(a)} \cdot \left(-\delta' \vec{r}_k \right) \leq 0$$

The above two results hold only if

$$\sum \vec{F}_k^{(a)} \cdot \delta' \vec{r}_k = 0 \qquad (1.9)$$

Equation (1.9) is the generalized form of the principle of virtual work. The principle can be stated as follows:

The work done in infinitesimal reversible virtual displacements, consistent with the constraints, from the equilibrium configuration of a system is zero.

It is important to note that the equilibrium of the system we have referred to in the above discussion is static equilibrium. If we extend this argument to systems in motion, we obtain another important principle called the D'Alembert's principle.

1.2.1 Generalized Coordinates

In analytical mechanics, the term generalized coordinates refers to the parameters that describe the configuration of the system relative to some reference configuration. These parameters must uniquely define the configuration of the system relative to the reference configuration. This is done assuming that this can be done with a single chart. The generalized velocities are the time derivatives of the generalized coordinates of the system. An example of a generalized coordinate is the angle that locates a point moving on a circle. The adjective "generalized" distinguishes these parameters from the traditional use of the term coordinate to refer to Cartesian coordinates. For example, describing the location of the point on the circle using x and y coordinates.

1.3   D'ALEMBERT'S PRINCIPLE

If $\vec{r}_k$ is the radius vector of the kth particle in the system of particles considered

above we have the equation of motion of the particle

$$\vec{F}_k = \frac{d}{dt}\left(m_k \, \dot{\vec{r}}_k \right) \tag{1.10}$$

or

$$\vec{F}_k - \frac{d}{dt}\left(m_k \, \dot{\vec{r}}_k \right) = 0 \tag{1.11}$$

If $\delta\vec{r}_k$ is an infinitesimal virtual displacement of the particle, we obtain

from Equation (1.11)

$$\left[\vec{F}_k - \frac{d}{dt}\left(m_k \, \dot{\vec{r}}_k \right) \right] \cdot \delta\vec{r}_k = 0$$

Considering all the particles in the system, the above equation gives

$$\sum_k \left[\vec{F}_k - \frac{d}{dt}\left(m_k \, \dot{\vec{r}}_k \right) \right] \cdot \delta\vec{r}_k = 0 \tag{1.12}$$

In the presence of constraints in the system, the above equation can be
written as

$$\sum_k \left[\left(\vec{F}_k^{(a)} + \vec{F}_k^{(c)} \right) - \frac{d}{dt}\left(m_k \, \vec{r}_k \right) \right] \cdot \delta\vec{r}_k = 0$$

or,

$$\sum_k \vec{F}_k^{(c)} \cdot \delta\vec{r}_k + \sum_k \left[\vec{F}_k^{(a)} - \frac{d}{dt}\left(m_k \, \dot{\vec{r}}_k \right) \right] \cdot \delta\vec{r}_k = 0 \tag{1.13}$$

Once again restricting our consideration to only such systems for which

$$\vec{F}_k^{(c)} \cdot \delta \vec{r}_k \geq 0$$

for all $\delta \vec{r}_k$ which are compatible with the constraints, we get from Equation (1.13)

$$\sum_k \left[\vec{F}_k^{(a)} - \frac{d}{dt}\left(\vec{p}_k \right) \right] \cdot \delta \vec{r}_k \leq 0 \qquad (1.14)$$

where $\vec{p}_k = m_k \dfrac{d}{dt}\left(\vec{r}_k \right)$ is the momentum of the kth particle.

Further, considering only such virtual displacements which are reversible and denoted as $\delta' \vec{r}_k$ we obtain

$$\sum_k \left[\vec{F}_k^{(a)} - \frac{d}{dt}\left(\vec{p}_k \right) \right] \cdot \delta' \vec{r}_k \leq 0 \qquad (1.15)$$

and also

$$\sum_k \left[\vec{F}_k^{(a)} - \frac{d}{dt}\left(\vec{p}_k \right) \right] \cdot \left(-\delta \vec{r}_k \right) \leq 0 \qquad (1.16)$$

Simultaneous validity of Equations (1.15) and (1.16) gives

$$\sum_k \left[\vec{F}_k^{(a)} - \frac{d}{dt}\left(\vec{p}_k \right) \right] \cdot \delta' \vec{r}_k = 0 \qquad (1.17)$$

The term $\sum_k F_k^{(a)} \cdot \delta' \vec{r}_k$ is the net work done by the applied forces in the course of virtual displacements of the particles of the system. It is usual to call $-\dfrac{d}{dt}\left(\vec{p}_k \right)$ as the **force of inertia** which may be denoted as $\vec{F}_k^{(I)}$. In view of this we may write Equation (1.17) as

$$\sum_k \left[\vec{F}_k^{(a)} + \vec{F}_k^{(I)} \right] \cdot \delta' \vec{r}_k = 0 \qquad (1.18)$$

$\left(\vec{F}_k^{(a)} + \vec{F}_k^{(I)} \right)$ can be called the *effective force* on the kth particle and denoted as $\vec{F}_k^{\text{eff}}$.

Equations (1.17) and (1.18) are different mathematical forms of D'Alembert's principle. The principle may be stated as follows:

For any dynamical system, the total work done by the effective force is zero in the course of reversible infinitesimal virtual displacement compatible with the constraints imposed on the system.

It is important to note that the coefficients of $\delta' \vec{r_k}$ in Equation (1.17) and (1.18) cannot be put equal to zero because $\delta' \vec{r_k}$ are not independent of each other.

Here is another mathematical form of D'Alembert's principle,

$$\sum \vec{F_k}^{eff} \cdot \delta^1 \vec{r_k} = 0 \qquad (1.19)$$

1.4 DERIVATION OF LAGRANGE'S EQUATION

Consider a mechanical system of N particles. At some instant of time t, let $\vec{r_1}, \vec{r_2} \ldots, \vec{r_N}$ be the position vectors of the particles with respect to some fixed origin. If the system is described by s generalized coordinates $q_1, \ldots, q_s$, then we have the transformation equations

$$\vec{r_i} = \vec{r_i}(q_1, \ldots, q_s, t) \cdot (i = 1, \ldots, N) \qquad (1.20)$$

Velocity vectors for the particles are then given by

$$\vec{v_i} = \frac{d\vec{r_i}}{dt} = \sum_{k=1}^{s} \frac{\partial \vec{r_i}}{\partial q_k} \dot{q_k} + \frac{\partial \vec{r_i}}{\partial t} \qquad (1.21)$$

If $\delta \vec{r_i}$ is an infinitesimal virtual displacement of the ith particle, we get

$$\delta \vec{r_i} = \sum_{j} \frac{\partial \vec{r_i}}{\partial q_j} \delta q_j + \frac{\partial \vec{r_i}}{\partial t} \delta t$$

However, since virtual displacement does not refer to displacement with respect to time, we have $\dfrac{\partial \vec{r_i}}{\partial t} = 0$ and hence we obtain

$$\delta \vec{r_i} = \sum_{j} \frac{\partial \vec{r_i}}{\partial q_j} \delta q_j \qquad (1.22)$$

According to D'Alembert's principle we have

$$\sum_{i} \left(\vec{F_i} - \dot{\vec{p_i}} \right) \cdot \delta \vec{r_i} = 0 \qquad (1.23)$$

where $\vec{F}_i$ is the actual force acting on the ith particle and $\dot{\vec{p}}_i$ is the reverse

effective force. Using Equation (1.22) in Equation (1.23) we get

$$\sum_i \left(\vec{F}_i - \dot{\vec{p}}_i \right) \cdot \sum_j \frac{\partial \vec{r}_i}{\partial q_j} \delta q_j \; = \; 0$$

or

$$\sum_{i,j} \vec{F}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j} \delta q_j - \sum_{i,j} \dot{\vec{p}}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j} \delta q_j \; = \; 0$$

or

$$\sum_j Q_j \delta q_j - \sum_{i,j} \dot{\vec{p}}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j} \delta q_j \; = \; 0 \tag{1.24}$$

where

$$Q_j \; = \; \sum_{i=1}^{N} \vec{F}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j} \tag{1.25}$$

are the components of the generalized forces.

We further have

$$\sum_{i,j} \dot{\vec{p}}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j} \delta q_j \; = \; \sum_{i,j} m_i \ddot{\vec{r}}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j} \delta q_j$$

$$= \sum_{i,j} \left[\frac{d}{dt} \left(m_i \vec{r}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j} \right) - m_i \vec{r}_i \cdot \frac{d}{dt} \left(\frac{\partial \vec{r}_i}{\partial q_j} \right) \right] \delta q_j \tag{1.26}$$

Now

$$\frac{d}{dt} \left(\frac{\partial \vec{r}_i}{\partial q_j} \right) \; = \; \sum_k \frac{\partial}{\partial q_k} \left(\frac{\partial \vec{r}_i}{\partial q_j} \right) \dot{q}_k + \frac{\partial}{\partial t} \left(\frac{\partial \vec{r}_i}{\partial q_j} \right)$$

$$= \; \sum_k \frac{\partial^2 \vec{r}_i}{\partial q_k \partial q_j} \dot{q}_k + \frac{\partial^2 \vec{r}_i}{\partial t \partial q_j}$$

$$= \; \sum_k \frac{\partial^2 \vec{r}_i}{\partial q_j \partial q_k} \dot{q}_k + \frac{\partial^2 \vec{r}_i}{\partial q_j \partial t}$$

$$= \frac{\partial}{\partial q_j} \left[\sum \frac{\partial \vec{r_i}}{\partial q_k} \dot{q}_k + \frac{\partial \vec{r_i}}{\partial t} \right]$$

or $$\frac{d}{dt}\left(\frac{\partial \vec{r_i}}{\partial q_j} \right) = \frac{\partial}{\partial q_j} \frac{d\vec{r_i}}{dt} = \frac{\partial \vec{v_i}}{\partial q_j} \qquad (1.27)$$

We have from Equation (1.21)

$$\frac{\partial \vec{v_i}}{\partial q_j} = \frac{\partial}{\partial q_j}\left[\sum_k \frac{\partial \vec{r_i}}{\partial q_j} \dot{q}_k + \frac{\partial \vec{r_i}}{\partial t} \right]$$

or $$\frac{\partial \vec{v_i}}{\partial \dot{q}_j} = \frac{\partial \dot{\vec{r_i}}}{\partial \dot{q}_j} \qquad (1.28)$$

Substituting Equations (1.27) and (1.28) in Equation (1.26) we obtain

$$\sum_{i,j} \vec{p_i} \cdot \frac{\partial \vec{r_i}}{\partial q_j} \delta q_j = \sum_{i,j} \left[\frac{d}{dt}\left(m_i \vec{v_i} \cdot \frac{\partial \vec{v_i}}{\partial q_j} \right) - m_i \vec{v_i} \cdot \frac{\partial \vec{v_i}}{\partial q_j} \right] \delta q_j$$

$$= \sum_j \left[\frac{d}{dt}\left\{ \frac{\partial}{\partial \dot{q}_j}\left(\sum_i \frac{1}{2} m_i v_i^2 \right) \right\} - \frac{\partial}{\partial q_j}\left(\sum \frac{1}{2} m_i v_i^2 \right) \right] \delta q_j \qquad (1.29)$$

$$= \sum_j \left[\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} \right] \delta q_j \qquad (1.30)$$

where $T = \sum \frac{1}{2} m_i v_i^2$ = Kinetic energy of the system of particles.

Using Equation (1.29) in Equation (1.24) we then obtain

$$\sum_j Q_j \delta q_j - \sum_j \left[\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} \right] \delta q_j = 0$$

The above can be re-written as

$$\sum_j \left[\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} - Q_j \right] \delta q_j = 0 \qquad (1.31)$$

For holonomic constraints, q_j's are independent of each other and hence the coefficient of each δq_j in Equation (1.31) separately vanishes giving

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = Q_j; \ (j = 1, \ ....., \ S) \tag{1.32}$$

Considering the system under consideration to be conservative, the potential energy U of the system is a function of only the position vectors, i.e., $U = U(\vec{r}_1, \ ....., \vec{r}_N)$ and force on each particle can be derived from the potential energy function U according to

$$\vec{F}_i = -\frac{\partial U}{\partial \vec{r}_i} \tag{1.33}$$

The generalized force Q_j, given by Equation (1.25) can thus be written as

$$Q_j = \sum_{i=1}^{N} -\frac{\partial U}{\partial r_i} \frac{\partial \vec{r}_i}{\partial q_j} = -\frac{\partial \vec{U}}{\partial q_j} \tag{1.34}$$

In view of Equation (1.34), we get from Equation (1.32)

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = -\frac{\partial U}{\partial q_j}$$

or
$$\frac{d}{dt}\frac{\partial (T)}{\partial \dot{q}_j} - \frac{\partial (T-U)}{\partial q_j} = 0$$

Since U does not depend on the generalized velocities, the above equation can also be written as

$$\frac{d}{dt}\frac{\partial}{\partial \dot{q}_j}(T-U) - \frac{\partial}{\partial q_j}(T-U) = 0 \tag{1.35}$$

If we replace $T-U$ by L, Equation (1.35) becomes

$$\frac{d}{dt}\left(\frac{\partial L}{\partial q_j}\right) - \frac{\partial L}{\partial \dot{q}_j} = 0 \tag{1.36}$$

Since T is a function of the generalized velocities $\dot{q}_k$'s and U is a function of generalized coordinate q_k's, we find L to be a function of coordinates, velocities and time in general, and in view of Equation (1.36) we can identify L as the Lagrangian function of the system. Thus, for a conservative system we obtain

L = Kinetic energy of the system – Potential energy of the system

Applications of Lagrangian Formulation

1: Motion of a Simple Pendulum Placed in a Uniform Gravitational Field

A simple pendulum consists of a point mass m at one end of a weightless, inelastic string of length l, the other end being rigidly clamped. The mass m

can swing back and forth in a vertical plane about the position of rest once it is displaced from the position of rest and released. As the mass is constrained to move on a circular arc in the vertical plane, the pendulum has only one degree of freedom. Thus, the pendulum is described, at any time, by only one generalized coordinate which can be conveniently taken as θ, the angle the string makes with the vertical as shown in Figure 1.1.
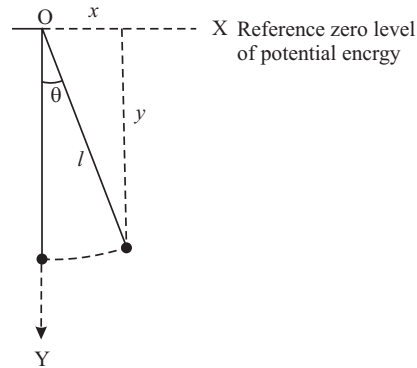
Fig. 1.1 *Movement of Pendulum*

If x and y are the coordinates of the mass point with respect to the origin at the point of suspension O, then we have

$$x = l \sin \theta$$
$$y = l \cos \theta$$

The kinetic energy of the point mass is

$$T = \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2\right) = \frac{1}{2}m\left[\left(l\cos\dot{\theta}\right)^2 + \left(-l\sin\dot{\theta}\right)^2\right]$$

or

$$T = \frac{1}{2}ml^2\left[\cos^2\theta\dot{\theta}^2 + \sin^2\theta\dot{\theta}^2\right]$$

or

$$T = \frac{1}{2}ml^2\dot{\theta}^2 \qquad (1.37)$$

The potential energy of the point mass in the gravitational field is

$$V = -mgy = -mgl\cos\theta \qquad (1.38)$$

The Lagrangian of the pendulum is thus given by

$$L = T - V = \frac{1}{2}ml^2\dot{\theta}^2 + mgl\cos\theta \qquad (1.39)$$

The Lagrange's equation for the coordinate θ is

$$\frac{d}{dl}\left(\frac{\partial L}{\partial \dot{\theta}}\right) = \frac{\partial L}{\partial \theta}$$

Using L given by Equation (1.39), the above gives

$$\frac{d}{dt}\left[\frac{1}{2}ml^2 2\dot{\theta}\right] = -mgl\sin\theta$$

or $$\frac{d}{dt} ml\dot{\theta} = mgl \sin \theta$$

or $$l\ddot{\theta} = -g \sin \theta$$

or $$\ddot{\theta} = -\frac{g}{l} \sin \theta \qquad (1.40)$$

Considering θ to be small, we get $\sin \theta \approx \theta$ and hence the above equation reduces to

$$\ddot{\theta} = -\frac{g}{l} \theta \qquad (1.41)$$

Equation (1.41) shows that under the condition that the angular amplitude is very small the motion of the pendulum is simple harmonic of time period

$$T = 2\pi \sqrt{\frac{l}{g}}$$

2: Motion of a Compound Pendulum in a Uniform Gravitational Field

Any rigid body capable of oscillating in a vertical plane about a horizontal axis passing through any point (excepting the centre of gravity) of the body is called a compound pendulum.

Let the vertical plane of oscillation of the compound pendulum be the XY plane.

Let us choose the origin of the coordinate system as the point O through which the horizontal axis (the X axis) passes.

Let *G* be the position of the centre of gravity of the body when at rest.

OG = *l* (say)

On displacing the pendulum slightly from the position of rest and releasing, the pendulum begins to oscillate about the horizontal axis through O.

At any instant of time *t*, let *G'* be the new position of the centre of gravity and $G\hat{O}G'$ be equal to θ as shown in Figure 1.2.

The kinetic energy of the pendulum at the instant *t* is

$$T = \frac{1}{2} I \dot{\theta}^2 \qquad (1.42)$$

where *I* is the moment of inertia of the pendulum about the axis of oscillation.

Reference zero level
of potential energy
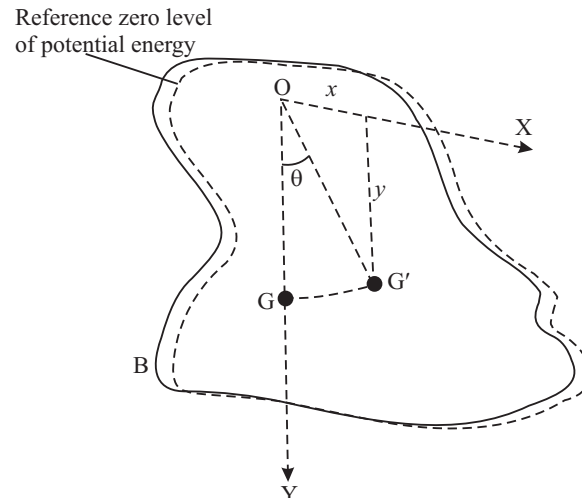
Fig. 1.2 *Positions of Centre of Gravity*

Taking the horizontal axis OX as the reference zero of potential energy, we get the potential energy of the pendulum at the instant *t* as

$$V = -mgy = -mgl \cos \theta \qquad (1.43)$$

The Lagrangian of the pendulum is thus

$$L = T - V = \frac{1}{2}I\dot{\theta}^2 + mgl\cos\theta \qquad (1.44)$$

From Equation (1.44) we find that the only generalized coordinate for the pendulum is θ. We thus have the Lagrange's equation for the compound pendulum

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) = \frac{\partial L}{\partial \theta} \qquad (1.45)$$

Using *L* given by Equation (1.44) the above equation gives

$$\frac{d}{dt}\left(\frac{1}{2}I2\dot{\theta}\right) + mgl\sin\theta = 0$$

or $\qquad\qquad I\ddot{\theta} + mgl\sin\theta = 0$

or $\qquad\qquad \ddot{\theta} + \dfrac{mgl}{I}\sin\theta = 0 \qquad (1.46)$

Considering θ small Equation (1.46) reduces to

$$\ddot{\theta} = -\frac{mgl}{I}\theta \qquad (1.47)$$

Clearly, the motion of the pendulum is simple harmonic of time period

$$T = 2\pi\sqrt{\frac{I}{mgl}} \qquad (1.48)$$

3: Motion of a Spherical Pendulum

A spherical pendulum consists of a point mass *m* constrained to move on the surface of a sphere. The position of the point mass at any instant is located by the Cartesian coordinates *x, y, z*, or more conveniently by the spherical polar coordinates *r*, θ, φ (*r* = radius of the sphere is constant) with respect to a coordinate frame XYZ having the origin at the centre of the sphere as shown in Figure 1.3.
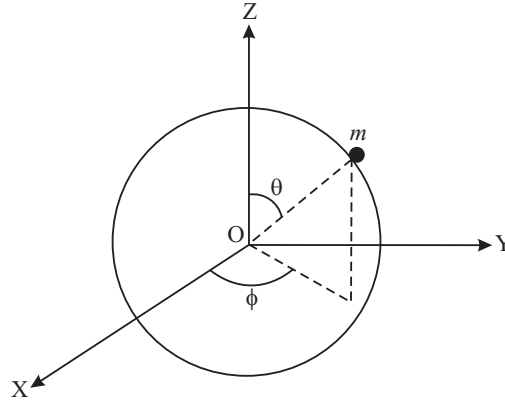
Fig. 1.3 *Motion of Spherical Pendulum*

We have the transformation equations

$$x = r \sin \theta \cos \phi$$
$$y = r \sin \theta \sin \phi \qquad (1.49)$$
$$z = r \cos \theta$$

The kinetic energy of the body at the instant of time under consideration is given by

$$T = \frac{1}{2} m \left[\dot{x}^2 + \dot{y}^2 + \dot{z}^2 \right] \qquad (1.50)$$

$\dot{x}, \dot{y}$ and $\dot{z}$ found by differentiating Equation (*i*) with respect to time,

when substituted in Equation (1.50) gives

$$T = \frac{1}{2} m \dot{r}^2 \left[\dot{\theta}^2 + \sin^2 \theta \dot{\phi}^2 \right] \qquad (1.51)$$

Considering the horizontal plane XOY as the plane of zero potential energy, we get the potential energy of the body as

$$V = mgz = mgr \cos \theta \qquad (1.52)$$

The Lagrangian of the spherical pendulum is then given by

$$L = T - V = \frac{1}{2} m \dot{r}^2 \left[\dot{\theta}^2 + \sin^2 \theta \dot{\phi}^2 \right] - mgr \cos \theta \ (1.53)$$

From Equation (1.53) we find the generalized coordinates for the pendulum to be θ and φ (since *r* is constant), so that the Lagrange's equations are

(*i*)
$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) = \frac{\partial L}{\partial \theta}$$

or
$$\frac{d}{dt}\left[\frac{1}{2}mr^2 2\dot{\theta}\right] = \frac{1}{2}mr^2\dot{\phi}^2 2\sin\cos\theta + mgr\sin\theta$$

or
$$\frac{d}{dt}\left(mr^2\dot{\theta}\right) - mr^2\sin\cos\theta\dot{\phi}^2 - mgr\sin\theta = 0$$

or $\qquad r\ddot{\theta} - r\sin\cos\theta\dot{\phi}^2 - g\sin\theta = 0 \qquad (1.54)$

(*ii*)
$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\phi}}\right) = \frac{\partial L}{\partial \phi}$$

or
$$\frac{d}{dt}\left[\frac{1}{2}mr^2\sin^2\theta 2\dot{\phi}\right] = 0$$

or $\qquad mr^2\dfrac{d}{dt}\left[\sin^2\theta\dot{\phi}\right] = 0 \qquad (1.55)$

4: Motion of a Particle Under a Central Force

Central force is that force which acts either towards or away from a fixed point (called the centre of the force) and depends only on the distance r from the fixed point.

We may thus express the magnitude of central force as $F = F(r)$.

Further, any central force can be derived from a potential function V according to

$$F = -\frac{dV}{dr}$$

which gives $\qquad dV = -Fdr$ and hence

$$V = -\int Fdr$$

Since F depends only on the distance r, we find from the above, the potential V to depend only on r, i.e., on the distance from the force centre. Thus, $V = V(r)$.

The most important characteristic of motion of a particle under central force is that the motion is restricted to take place in a plane. The number of degrees of freedom of the particle is thus two and the convenient generalized coordinates are polar coordinates r and θ as indicated in Figure 1.4.
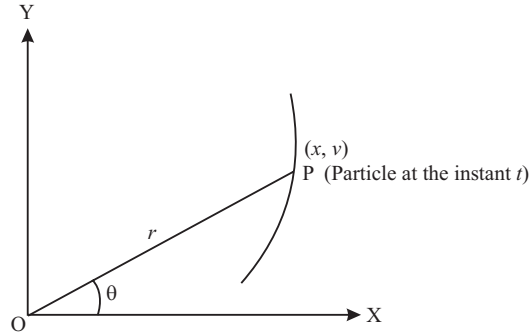
Fig. 1.4 Polar Coordinates r and α

The kinetic energy of the particle at the instant *t* is

$$T = \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2\right) \tag{1.56}$$

Coordinate transformation equations are

$$x = r\cos\theta \tag{1.57}$$

and

$$y = r\sin\theta$$

Equations (1.57) give

$$\dot{x} = \dot{r}\cos\theta - r\sin\theta\dot{\theta} \tag{1.58}$$

$$\dot{y} = \dot{r}\sin\theta + r\cos\theta\dot{\theta} \tag{1.59}$$

Using Equations (1.58) and (1.59) in Equation (1.56) we obtain

$$T = \frac{1}{2}m\left[r^2\cos^2\theta + r^2\sin^2\theta\dot{\theta}^2 - 2r\dot{r}\cos\theta\sin\theta\dot{\theta} + r^2\sin^2\theta + r^2\cos^2\theta\dot{\theta}^2 + 2r\dot{r}\sin\theta\cos\theta\dot{\theta}\right]$$

or

$$T = \frac{1}{2}m\left[\dot{r}^2 + r^2\dot{\theta}^2\right] \tag{1.60}$$

The potential energy of the particle is

$$V = V(r) \tag{1.61}$$

Thus, the Lagrangian of the particle is given by

$$L = T - V = \frac{1}{2}m\left[\dot{r}^2 + r^2\dot{\theta}^2\right] - V(r) \tag{1.62}$$

Lagrange's equations are

(a) $\dfrac{d}{dt}\left(\dfrac{\partial L}{\partial \dot{r}}\right) = \dfrac{\partial L}{\partial r}$

Using *L* given by Equation (1.62), the above gives

$$\frac{d}{dt}\left[\frac{1}{2}m2\dot{r}\right] = \frac{1}{2}m\dot{\theta}^2 2r - \frac{dV(r)}{dr}$$

or $\qquad m\ddot{r} = mr\dot{\theta}^2 - \dfrac{dV(r)}{dr}$

or $\quad m\ddot{r} - mr\dot{\theta}^2 = -\dfrac{dV(r)}{dr}$ $\qquad$ (1.63)

(b) $\qquad \dfrac{d}{dt}\left(\dfrac{\partial L}{\partial \dot{\theta}}\right) = \dfrac{\partial L}{\partial \dot{\theta}}$

Using L given by Equation (1.62), the above gives

$$\dfrac{d}{dt}\left[\dfrac{1}{2}mr^2 2\dot{\theta}\right] = 0 \qquad (1.64)$$

$$\dfrac{d}{dt}\left(\dfrac{1}{2}mr^2\dot{\theta}\right) = 0$$

or $\qquad mr^2\dot{\theta} = \text{constant} \qquad (1.65)$

Note: By definition, the generalized momentum (p_q) conjugate to the generalized coordinate θ is given by

$$p_q = \dfrac{\partial L}{\partial \dot{\theta}}$$

Using Equation (1.62), the above becomes

$$p_q = mr^2\dot{\theta} \qquad (1.66)$$

Equations (1.65) and (1.66) show that for a particle moving under a central force, the angular momentum is a constant quantity.

From Equation (1.64) we get

$$m2r\dot{r}\dot{\theta} + mr^2\ddot{\theta} = 0$$

or $\qquad r\ddot{\theta} + 2\dot{r}\dot{\theta} = 0 \qquad (1.67)$

5: Motion of a Linear Harmonic Oscillator

Let a particle mass m undergo simple harmonic motion along the X-axis. Let us measure displacement of the particle from the mean position O which is taken as the origin of the X-axis as shown in Figure 1.5.

Fig. 1.5

If x is the displacement of the particle at any instant of time t, the kinetic energy of the particle at that instant is

$$T = \dfrac{1}{2}m\dot{x}^2 \qquad (1.68)$$

If k is the restoring force per unit displacement acting on the particle then the potential energy of the particle when the displacement is x is given by

$$V = \frac{1}{2}kx^2 \qquad (1.69)$$

The Lagrangian of the oscillator is thus

$$L = T - V = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2 \qquad (1.70)$$

The Lagrange's equation for the oscillator is given by

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = \frac{\partial L}{\partial x} \qquad (1.71)$$

Using Equation (1.70) we get

$$\frac{\partial L}{\partial \dot{x}} = \frac{1}{2}m2\dot{x} = m\dot{x} \qquad (1.72)$$

and

$$\frac{\partial L}{\partial \dot{x}} = -\frac{1}{2}k2x = -kx \qquad (1.73)$$

Using Equations (1.72) and (1.73) in Equation (1.71) we obtain

$$\frac{d}{dt}(m\dot{x}) = -kx$$

or

$$m\ddot{x} + kx = 0$$

which is the familiar equation of motion for a linear harmonic oscillator.

6: Motion of a Coplanar Double Pendulum Placed in a Uniform Gravitational Field

A double pendulum consists of two mass points m and m_1 at the ends of two weightless rods of lengths l and l_1.

The rod of length l is suspended from a rigid support at O while the rod of length l_1 is suspended from a hinge at the mass point m as shown in Figure 1.6.
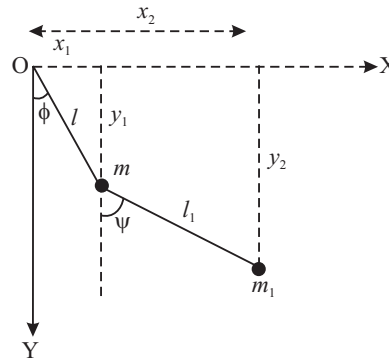
Fig. 1.6 *Pendulum Motion: In Uniform Gravitational Field*

Both the pendulums are constrained to move in the same vertical plane, say the XY plane and hence the number of degrees of freedom for the pendulum is two.

The two generalized coordinates are conveniently chosen as angles ϕ and ψ which the two rods make with the Y-axis which is assumed to be vertical.

The Cartesian coordinates of the mass point m are

$$x_1 = l \sin \phi; \, y_1 = l \cos \phi \qquad (1.74)$$

The Cartesian coordinates of the mass point m_1 are

$$x_2 = l \sin \phi + l_1 \sin \psi; \, y_2 = l_1 \cos \phi + l_2 \cos \psi \qquad (1.75)$$

For the mass point m, the kinetic energy is

$$T_1 = \frac{1}{2} m \left(\dot{x}_1^2 + \dot{y}_1^2 \right) = \frac{1}{2} m \left[\left(l \cos \phi \dot{\phi} \right)^2 + \left(-l \sin \phi \dot{\phi} \right)^2 \right] = \frac{1}{2} m l^2 \dot{\phi}^2 \qquad (1.76)$$

The potential energy of the mass point m is

$$V_1 = -mgy_1 = -mgl \cos \phi \qquad (1.77)$$

For the mass point m_1, the kinetic energy is

$$T_2 = \frac{1}{2} m_1 \left(\dot{x}_2^2 + \dot{y}_2^2 \right) = \frac{1}{2} m_1 \left[\left(l \cos \phi \dot{\phi} + l_1 \sin \psi \dot{\psi} \right)^2 + \left(-l \sin \phi \dot{\phi} - l_1 \sin \psi \dot{\psi} \right)^2 \right]$$

$$= \frac{1}{2} m_1 \left[l^2 \dot{\phi}^2 + l^2 \dot{\psi}^2 + 2 l l_1 \cos (\phi - \psi) \dot{\phi} \dot{\psi} \right] \qquad (1.78)$$

The potential energy of the mass m_1 is

$$V_2 = m_1 g y_2 = m_1 g (l \cos \phi + l_1 \cos \psi) \qquad (1.79)$$

The Lagrangian of the double pendulum is thus

$$L = T_1 - V_1 + T_2 - V_2$$

Using Equations (1.76), (1.77), (1.78) and (1.79) we obtain

$$L = \frac{1}{2} m l^2 \dot{\phi}^2 + mgl \cos \phi + \frac{1}{2} m_1 \left[l^2 \dot{\phi}^2 + l_1^2 \dot{\psi}^2 + 2 l l_1 \cos (\phi - \psi) \dot{\phi} \dot{\psi} \right] + m_1 g \left(l \cos \phi + l_1 \cos \psi \right)$$

or

$$L = \frac{1}{2} (m + m_1) l^2 \dot{\phi}^2 + \frac{1}{2} m_1 l_1^2 \dot{\psi}^2 + m_1 l l_1 \dot{\phi} \dot{\psi} \cos (\phi - \psi) + (m + m_1) gl \cos \phi + m_1 g l_1 \cos \psi$$

$$(1.80)$$

The Lagrange's equations for the pendulum are

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\phi}} \right) = \frac{\partial L}{\partial \phi} \, ; \, \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\psi}} \right) = \frac{\partial L}{\partial \psi} \qquad (1.81)$$

We obtain using Equation (1.80)

$$@(m + m_1) l^2 \ddot{\phi} + m_1 l l_1 \ddot{\psi} \cos (\phi - \psi) + m_1 l l_1 \dot{\psi}^2 \sin (\phi - \psi) + (m + m_1) gl \sin \phi = 0 \quad (1.82)$$

and

(*b*) $m_1 l^2 \ddot{\phi} + m_1 l l_1 \ddot{\phi} \cos(\phi - \psi) - m_1 l l_1 \dot{\phi}^2 \sin(\phi - \psi) + m_1 g l_1 \sin \psi = 0$ (1.83)

In the special case when $m_1 = m = m_0$, and $l_1 = l = l_0$, the above equations assume the simple forms

$$2\ddot{\phi} + \ddot{\psi} \cos(\phi - \psi) + \dot{\psi}^2 \sin(\phi - \psi) + 2\frac{g}{l_o} \sin \phi = 0 \qquad (1.84)$$

$$\ddot{\psi} + \ddot{\phi} \cos(\phi - \psi) - \dot{\phi}^2 \sin(\phi - \psi) + \frac{g}{l_o} \sin \psi = 0 \qquad (1.85)$$

Further, if we consider both ϕ and ψ small, the above equations further reduce to

$$2\ddot{\phi} + \ddot{\psi} + 2\frac{g}{l_o}\phi = 0 \qquad (1.86)$$

$$\ddot{\psi} + \ddot{\phi} + \frac{g}{l_o}\psi = 0 \qquad (1.87)$$

The above are coupled differential equations for the double pendulum.

7: Lagrangian of a Hoop Rolling Down an Inclined Plane without Slipping

Consider a hoop (circular ring) of radius r and mass m rolling down an inclined plane without slipping (velocity of the instantaneous point of contact of the hoop along the plane is zero) as shown in Figure 1.7.
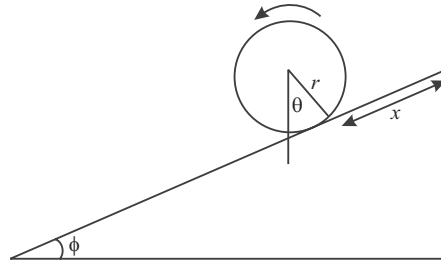
***Fig. 1.7** Motion in an Inclined Plane*

Let us measure the displacement of the centre of mass of the hoop from the top of the incline.

Let at some instant of time *t*, the centre of mass be at a distance *x* from the top.

The velocity of the centre of mass is then

$$v = \dot{x} \qquad (1.88)$$

Since there is no slipping, the angular velocity of the hoop about the axis of rotation through the centre of mass is

$$\dot{\theta} = \frac{v}{r} = \frac{\dot{x}}{r} \tag{1.89}$$

Now, the kinetic energy of the hoop at the instant t is

$$T = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}mr^2\dot{\theta}^2 \tag{1.90}$$

The potential energy of the hoop at the instant t is

$$V = mg(l - x)\sin\phi \tag{1.91}$$

where l is the length of the inclined plane.

The Lagrangian of the hoop is thus given by

$$L = T - V = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}mr^2\dot{\theta}^2 - mg(l - x)\sin\phi \tag{1.92}$$

Using Equation (1.89) we may write the Lagrangian as

$$L = m\dot{x}^2 + mgx\sin\phi - mgl\sin\phi \tag{1.93}$$

8: Lagrangian of a Charged Particle Moving in an Electromagnetic Field

The electric field vector $\vec{E}$ and the magnetic field vector $\vec{B}$ which describe an electromagnetic field satisfy the Maxwell's equations

$$\text{curl}\,\vec{E} + \frac{1}{c}\frac{\partial \vec{B}}{\partial t} = 0 \;;\; \text{div}\,\vec{D} = 4\pi\vec{P}$$

$$\text{curl}\,\vec{H} - \frac{1}{c}\frac{\partial \vec{D}}{\partial t} = \frac{4\pi}{c}\vec{j} \;;\; \text{div}\,\vec{B} = 0 \tag{1.94}$$

The force $\vec{F}$ experienced by a particle of mass m having charge q moving with a velocity $\vec{v}$ in the electromagnetic field is given by

$$\vec{F} = q\vec{E} + \frac{q}{c}\left(\vec{v} \times \vec{B}\right) \tag{1.95}$$

The electromagnetic field can alternatively be described by a scalar potential ϕ and vector potential $\vec{A}$ defined according to

$$\vec{B} = \text{curl}\,\vec{A}$$

and

$$\vec{E} = -\vec{\nabla}\phi = \frac{1}{c}\frac{\partial \vec{A}}{\partial t} \tag{1.96}$$

In terms of ϕ and $\vec{A}$ the force $\vec{F}$ becomes

$$\vec{F} = q\left[-\vec{\nabla}\phi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t} + \frac{1}{c}\left(\vec{v}\times\vec{\nabla}\times\vec{A}\right)\right] \qquad (1.97)$$

The x-component of $\vec{F}$ is

$$F_x = q\left[-\left(\vec{\nabla}\phi\right)_x - \frac{1}{c}\left(\frac{\partial \vec{A}}{\partial t}\right)_x + \frac{1}{c}\left(\vec{v}\times\vec{\nabla}\times\vec{A}\right)_x\right] \qquad (1.98)$$

Now,

$$\left(\vec{\nabla}\phi\right)_x = \frac{\partial \phi}{\partial x} \qquad (1.99)$$

$$\frac{d}{dt}A_x = \frac{\partial}{\partial t}A_x + \frac{\partial A_x}{\partial x}\frac{dx}{dt} + \frac{\partial A_y}{\partial y}\frac{dy}{dt} + +\frac{\partial A_z}{\partial z}\frac{dz}{dt} \qquad (1.100)$$

$$= \frac{\partial A_x}{\partial t} + \frac{\partial A_x}{\partial x}v_x + \frac{\partial A_y}{\partial y}v_y + \frac{\partial A_z}{\partial z}v_z$$

Also,

$$\left(\vec{v}\times\vec{\nabla}\times\vec{A}\right)_x = v_y\left(\vec{\nabla}\times\vec{A}\right)_z - v_z\left(\vec{\nabla}\times\vec{A}\right)_y$$

$$= v_y\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) - v_z\left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right)$$

$$= v_x\frac{\partial A_x}{\partial x} + v_y\frac{\partial A_y}{\partial y} + v_z\frac{\partial A_z}{\partial z} - v_y\frac{\partial A_x}{\partial y} - v_z\frac{\partial A_x}{\partial z} - v_x\frac{\partial A_x}{\partial x}$$

$$= \frac{\partial}{\partial x}\left(\vec{v}.\vec{A}\right) - \frac{dA_x}{dt} + \frac{\partial}{\partial t}A_x \qquad (1.101)$$

Using Equations (1.99) and (1.101) in Equation (1.98) we obtain

$$F = q\left[-\frac{\partial \phi}{\partial x} - \frac{1}{c}\frac{\partial A_x}{\partial t} + \frac{1}{c}\frac{\partial}{\partial x}\left(\vec{v}.\vec{A}\right) - \frac{1}{c}\frac{dA_x}{dt} + \frac{1}{c}\frac{\partial A_x}{\partial t}\right]$$

$$= \left[-\frac{\partial}{\partial x}\left(\phi - \frac{1}{c}\vec{v}.\vec{A}\right) - \frac{1}{c}\frac{d}{dt}\left\{\frac{\partial}{\partial v_x}\left(\vec{v}.\vec{A}\right)\right\}\right]$$

or

$$F = -\frac{\partial U}{\partial x} + \frac{d}{dt}\frac{\partial U}{\partial v_x} \qquad (1.102)$$

where $$U = q\phi - \frac{q}{c}\vec{v}.\vec{A} \qquad (1.103)$$

We note that $\dfrac{\partial \phi}{\partial v_x} = 0$ since ϕ is independent of velocities. Thus, U is

a kind of generalized velocity-dependent potential.

The Lagrangian of the charged particle is thus given by

$$L = T - U \ (T \text{ is the kinetic energy})$$

or $$L = T - q\phi + \frac{q}{c}\vec{v}.\vec{A} \qquad (1.104)$$

1.4.1 Velocity Dependent Potentials

Velocity-dependent potential functions can sometimes be used to determine the field of force that can be applied in order that particles may move in specified paths. In particular, the electromagnetic field vectors E and B can be determined from such a potential function if the paths on which charged particles move are specified.

The velocity-dependent potential U is related to the kinetic energy T and the Lagrangian function L by the equation U = T - L, where L is an arbitrary solution of the Lagrange equation (d/dt)(δL/δp') - (δL/δp) = 0, where p = constant represents the orthogonal trajectories of the curves which describe the paths the particles are to follow. From the velocity-dependent potential function U, the field of force can be calculated by the definition Qp = -(δU/δp) + (d/dt)(δU/δp').

Check Your Progress

4. Define D'Alembert's principle.
5. What would be the Lagrangian function of the system for a conservative system?

1.5 RAYLEIGH'S DISSIPATION FUNCTION

In physics, the Rayleigh dissipation function, named for Lord Rayleigh, is a function used to handle the effects of velocity-proportional frictional forces in Lagrangian mechanics. If the frictional force on a particle with velocity $\vec{v}$ can be written as $\vec{F}_f = -\vec{k}\cdot\vec{v}$, the Rayleigh dissipation function can be defined for a system of N particles as

$$G(v) = \frac{1}{2}\sum_{i=1}^{N}(k_x v_{i,x}^2 + k_y v_{i,y}^2 + k_z v_{i,z}^2).$$

The force of friction is negative the velocity gradient of the dissipation function, $\vec{F}_f = -\nabla_v G(v)$. The function is half the rate at which energy is being dissipated by the system through friction.

1.6 VARIATIONAL EQUATION AND EULER-LAGRANGE'S EQUATION

The D'Alembert's principle, which is widely used together with Newton's laws of motion for dealing with mechanical systems, is a differential principle. This is because, in using this principle, we need to consider the instantaneous state of a system (defined by positions and velocities in the configuration space of the system) along with some infinitesimal virtual displacements from the instantaneous position.

The variational principle finds immense usefulness in treating mechanical system on the one hand, while on the other hand, it considers the motion of the system as a whole between the given time limits along some small variation in the motion of the system between the same time limits from the actual motion. In this sense, the variational principle is essentially an integral principle. In the following, we will discuss some important aspects of the calculus of variation that happens to be useful for future development of different formulations of mechanics.

Consider a curve given by

$$y = y(x) \tag{1.105}$$

defined between two points (x_1, y_1) and (x_2, y_2) as shown in Figure 1.8. We may conveniently call the two points as end points.

Let a function $f = f\left(y, \dfrac{dy}{dx}, x\right) = f(y, y'x)$ be defined on the above curve.

Our basic problem is to obtain the curve for which the line integral of the function f between the end points (x_1, y_1) and (x_2, y_2) is stationary, i.e.,

$$I = \int_{x_1}^{x_2} f(y, y', x)\, dx = \text{Extremum (either maximum or minimum)}$$
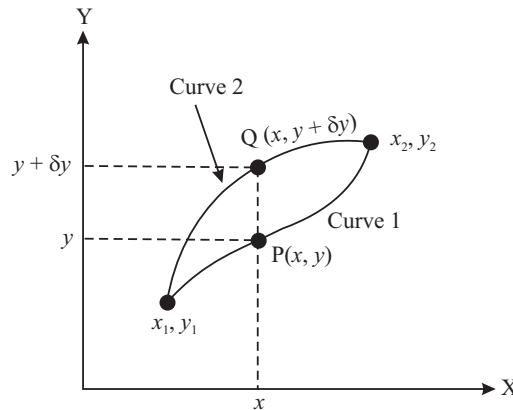
Fig. 1.8 *Neighbouring Curves*

In Figure 1.8, two neighboring curves governed by Equation (1.105) are shown between the end points $y(x_1) = y_1$ and $y(x_2) = y_2$, the curve-1 corresponding to the stationary value of the integral I. Consider the point $P(x, y)$ on Curve 1 for $x = x$.

The point on Curve 2 for the same value of x is the point $Q(x, y + \delta y)$

. Here, δy defines the variation in y as we go over from Curve 1 to Curve 2, keeping x the same.

It is useful to associate some parameter, say α, with all the possible curves determined by the constraints between the end points indicated. The α should be such that for some given value, for simplicity, say, for $\alpha = 0$ the Curve 2 coincides with Curve 1.

Corresponding to the extremum value of the integral, y is then a function of both the independent variable x and the parameter α. We may express

$$y(\alpha, x) = y(x) + \alpha \eta(x) \tag{1.106}$$

where $\eta(x)$ is a function of x which has continuous first derivative and vanishes at the two end points. Clearly, $y(\alpha, x)$ reduces to $y(x)$ at the two end points.

In view of our considerations above, the integral I becomes a function of the parameter α and we get

$$I(\alpha) = \int_{x_1}^{x_2} f\left[y(\alpha, x), y'(\alpha, x), x\right] dx \tag{1.107}$$

Condition that $I(\alpha)$ has an extremum value is thus

$$\left|\frac{\partial I(\alpha)}{\partial \alpha}\right|_{\alpha=0} = 0 \tag{1.108}$$

Differentiating Equation (1.107) with respect to α we obtain

$$\frac{\partial I(\alpha)}{\partial \alpha} = \frac{\partial}{\partial \alpha}\left[\int_{x_1}^{x_2} f\left[y(\alpha, x), y'(\alpha, x), x\right]\right] dx$$

$$= \int_{x_1}^{x_2}\left[\frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha} + \frac{\partial f}{\partial y'}\frac{\partial y'}{\partial \alpha}\right] dx \quad \text{(we may note that } \frac{\partial x}{\partial \alpha} = 0)$$

or $$\frac{\partial I(\alpha)}{\partial \alpha} = \int_{x_1}^{x_2}\left[\frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha} + \frac{\partial f}{\partial y'}\frac{\partial^2 y}{\partial \alpha \partial x}\right] dx \tag{1.109}$$

Integrating by parts, we have

$$\int_{x_1}^{x_2}\frac{\partial f}{\partial y'}\frac{d}{dx}\left(\frac{\partial y}{\partial x}\right) dx = \left\{\frac{\partial f}{\partial y'}\frac{\partial y}{\partial x}\right\}_{x_1}^{x_2} - \int_{x_1}^{x_2}\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\frac{\partial y}{\partial x} dx \tag{1.110}$$

Clearly, $$\frac{\partial y}{\partial \alpha}\bigg|_{x_1}^{x_2} = \eta(x)\big|_{x_1}^{x_2} = \eta(x_2) - \eta(x_1) = 0 \tag{1.111}$$

We thus get

$$\int_{x_1}^{x_2}\frac{\partial f}{\partial y'}\frac{d}{dx}\left(\frac{\partial y}{\partial x}\right) dx = \int_{x_1}^{x_2}\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\frac{\partial y}{\partial \alpha} dx \tag{1.112}$$

Using Equation (1.112) in Equation (1.109) we obtain

$$\frac{\partial I(\alpha)}{\partial \alpha} = \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} \frac{\partial y}{\partial \alpha} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\frac{\partial y}{\partial \alpha} \right] dx$$

$$= \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) \right] \eta(x)\, dx \qquad (1.113)$$

We may note that the functions y and y' with respect to which the derivatives of the function f appear on the right hand side of Equation (1.113) are functions of α. However, for $\alpha = 0$ we get $y(\alpha, x) = y(x)$, $y'(\alpha, x) = y'(x)$ and Equation (1.113) becomes independent of α. Since $\eta(x)$ is an arbitrary function, for $\left|\dfrac{\partial I(\alpha)}{\partial \alpha}\right|_{\alpha=0}$ to vanish so that $I(\alpha)$ has an extremum value, we find from Equation (1.113)

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0$$

or $$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = \frac{\partial f}{\partial y} \qquad (1.114)$$

Equation (1.114) is the Euler-Lagrange equation as obtained earlier.

The Euler-Lagrange equation can be generalized to the case when

$$f = f\left(y_1, \ldots, y_s, y'_1, \ldots, y'_s, x\right) \qquad (1.115)$$

In this case, the equation reads

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'_k}\right) = \frac{\partial f}{\partial y_k}; \qquad k = 1, \ldots, s \qquad (1.116)$$

The results of the calculus of variation can be expressed in terms of δ-notation as

$$\delta I = \delta \int_{x_1}^{x_2} f\left(y_1, \ldots, y_s, y'_1, \ldots, y'_s, x\right) dx = 0 \qquad (1.117)$$

1.7 DERIVATION OF LAGRANGE'S EQUATION FROM HAMILTON'S PRINCIPLES

Using Hamilton's principle of least action, it is possible to derive the dynamical equations for the system under consideration.

We have the Lagrangian of the system given by

$$L = L\left(q_1, \ldots, q_s, \dot{q}_1, \ldots, \dot{q}_s, t\right)$$

For convenience, we may write L in a shorter form as

$$L = L(q_k, \dot{q}_k, t) \tag{1.118}$$

In the above, q_k stands for all the coordinates and $\dot{q}_k$ stands for all the velocities which describe the system.

q_k's in general depend on time explicitly so that we should write $q_k(t)$ instead of q_k. Let $q_k(t)$ be replaced by $q_k(t) + \delta q_k(t)$. where $\delta q_k(t)$ is a small variation in $q_k(t)$ in the interval of time from time t_1 to time t_2. The variation of action S for fixed t_1 and t_2 is then

$$\delta S = \int_{t_1}^{t_2} L(q_k + \delta q_k, \dot{q}_k + \delta \dot{q}_k, t) dt - \int_{t_1}^{t_2} L(q_k, \dot{q}_k, t) \tag{1.119}$$

The major contributions in the expansion of $L\ (q_k + \delta q_k, \dot{q}_k + \delta \dot{q}_k, t)$ in powers of δq_k and $\delta \dot{q}_k$ are of the first order. Hence, for S to be an extremum, these terms should be zero. Thus, Hamilton's principle given by the following equation:

$$\delta S = \delta \int_{t_1}^{t_2} L(q_1, \dots, q_s, \dot{q}_1, \dots, \dot{q}_s, t) = 0$$

takes the form,

$$\delta S = \delta \int_{t_1}^{t_2} L(q_k, \dot{q}_k, t) dt = \int_{t_1}^{t_2} \sum_{k=1}^{s} \left(\frac{\partial L}{\partial q_k} \delta q_k + \frac{\partial L}{\partial \dot{q}_k} \delta \dot{q}_k \right) dt = 0 \tag{1.120}$$

We may note that t is fixed in the δ-variation under consideration. The following identity holds

$$\delta \dot{q}_k = \frac{d}{dt}(\delta q_k) \tag{1.121}$$

Using Equation (1.121) in Equation (1.120) we get

$$\delta S \int_{t_1}^{t_2} \sum_{k=1}^{s} \left(\frac{\partial L}{\partial q_k} \delta q_k \right) dt + \int_{t_1}^{t_2} \sum_{k=1}^{s} \left(\frac{\partial L}{\partial \dot{q}_k} \frac{d}{dt}(\delta \dot{q}_k) \right) dt = 0$$

or $\quad \delta S = \int_{t_1}^{t_2} \sum_{k} \left(\frac{\partial L}{\partial q_k} \delta q_k \right) dt + \sum_{k=1}^{s} \left\{ \frac{\partial L}{\partial \dot{q}_k} \delta \dot{q}_k \right\}_{t_1}^{t_2} - \sum_{k=1}^{s} \left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k} \right) \delta q_k \right] dt = 0 \quad$ (1.122)

Since variations at the end points are zero, i.e.,

$$\delta q_k(t_1) = 0 = \delta q_k(t_2) \tag{1.123}$$

Equation (1.122) becomes

$$\delta S = \int_{t_1}^{t_2} \sum_{k=1}^{s} \left(\frac{\partial L}{\partial q_k} \delta q_k \right) dt - \int_{t_1}^{t_2} \sum_{k=1}^{s} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k} \right) \delta q_k dt = 0$$

or
$$\delta S = \sum_{k=1}^{s} \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q_k} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k} \right) \right] \delta q_k \, dt = 0 \tag{1.124}$$

The result given by Equation (1.124) holds for all arbitrary variations provided the coefficient of δq_k in the integrand on the right hand side vanishes for each k. We thus obtain

$$\frac{\partial L}{\partial q_k} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k} \right) = 0$$

or
$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k} \right) - \frac{\partial L}{\partial q_k} = 0; \quad (k = 1, 2, \ldots, s) \tag{1.125}$$

The above set of s number of second order differential equations satisfied by the Lagrangian of the system are called the Lagrange's equations of motion.

Lagrange's equations of motion given by Equation (1.125) can be seen to follow directly from Euler-Lagrange equation given by Equation

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'_k} \right) = \frac{\partial f}{\partial y_k}; \qquad k = 1, \ldots, s$$

. If in the function f given by Equation $(f = f(y_1, \ldots, y_s, y'_1, \ldots, y'_s, x)$ we replace $y_1, \ldots, y_s$ by the generalized coordinates $q_1, \ldots, q_s$, respectively, $y_1', \ldots, y_s'$ by the generalized velocities $\dot{q}_1, \ldots, \dot{q}_s$ respectively and x by t then the function f can be identified as the

Lagrangian $L(q_1, \ldots, q_s, \dot{q}_1, \ldots, \dot{q}_s, t)$ and the Euler-Lagrange equations given

by Equation

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'_k} \right) = \frac{\partial f}{\partial y_k}; \qquad k = 1, \ldots, s$$

become the Lagrange's equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k} \right) = \frac{\partial L}{\partial q_k}; \quad k = 1, \ldots, s.$$

Important Properties of the Lagrangian Function

In this section we discuss the important properties of the Lagrangian function and attempt to find its meaning, i.e., to find whether or not the Lagrangian function is representative of some physical quantity of the system under consideration.

(i) Lagrangian is Gauge Invariant

The Lagrangian function of a system having s degrees of freedom and described by the generalized coordinates $q_1, \ldots, q_s$ and the generalized velocities $\dot{q}_1, \ldots, \dot{q}_s$ is given by

$$L = L(q_1, ..., q_s, \dot{q}_1, ..., \dot{q}_s, t) = L(q_k, \dot{q}_k, t) \quad (1.126)$$

Consider an arbitrary function $F = F(q_1, ..., q_s, t) = F(q_k, t)$ and define a new function $L'(q_k, \dot{q}_k, t)$ as

$$L'(q_k, \dot{q}_k, t) = L(q_k, \dot{q}_k, t) + \frac{d}{dt} F(q_k, t) \quad (1..127)$$

The action of the system between the time limits t_1 and t_2 is the time integral

$$S = \int_{t_1}^{t_2} L(q_k, \dot{q}_k, t) dt \quad (1.128)$$

Let us consider the time integral of the new function L' between the same time limits

$$S' = \int_{t_1}^{t_2} L'(q_k, \dot{q}_k, t) dt \quad (1.129)$$

Using Equation (1.127) in Equation (1.129) we get

$$S' = \int_{t_1}^{t_2} \left[L(q_k, \dot{q}_k, t) + \frac{d}{dt} F(q_k, t) \right] dt$$

$$= \int_{t_1}^{t_2} L(q_k, \dot{q}_k, t) dt + \int_{t_1}^{t_2} \frac{d}{dt} F(q_k, t) dt$$

or $\qquad S' = S + \left\{ F(q_k, t) \right\}_{t_1}^{t_2} \quad (1.130)$

Taking δ-variation of Equation (1.130) we get

$$\delta S' = \delta S + \delta \left\{ F(q_k, t) \right\}_{t_1}^{t_2}$$

or, $\qquad \delta S' - \delta S = \delta \left[F(q_k, t) \right]_{t=t_2} - \delta \left[F(q_k, t) \right]_{t=t_1} = 0 \quad (1.131)$

because $\delta q_k = 0$ at $t = t_1$ and at $t = t_2$.

According to Hamilton's principle we have

$$\delta S = 0$$

In view of this and Equation (1.131) we find

$$\delta S' = 0 \quad (1.132)$$

The condition $\delta S' = 0$ leads to equations of motion which are the same as those given by the condition $\delta S = 0$. Hence, we may identify the function L' given by Equation (1.127) also as the Lagrangian of the system.

From the above we may conclude that the Lagrangian of a system cannot be defined uniquely, but can be defined only within an additive total time derivative of any function of coordinates relevant to the system and time.

The arbitrary function $F(q_1, ..., q_s, t) = F(q_k, t)$ is called **gauge function**.

Hence, the above result shows that Lagrangian of a system is **gauge invariant**.

(ii) Lagrangian is Additive

Let A and B be two non-interacting parts of a mechanical system. Let L_A and L_B be their Lagrangians, respectively. By additivity, we mean that the Lagrangian of the whole system is given by

$$L = L_A + L_B \tag{1.133}$$

A consequence of this property is that the equation of motion of the part A are completely independent of the quantities of the part *B* and vice-versa.

(iii) Lagrangian of a System is Arbitrary within an Overall Multiplicative Constant

This property means that if the Lagrangian is multiplied by any arbitrary constant then the equations of motion remain unaltered.

Lagrangian of a Particle Moving Freely in Space

Consider a particle of mass *m* moving freely in space with respect to the origin of an inertial frame of reference. Let $\vec{r}$ be the position vector of the particle at the instant of time *t*. Let $\vec{v}$ be the velocity of the particle at the instant *t* as observed from the frame under consideration. The Lagrangian function of the particle is given by

$$L = L(\vec{r}, \vec{v}, t) \tag{1.134}$$

From the symmetry properties of space and time with respect to an inertial frame, namely, homogeneity and isotropy of space and homogeneity of time, the Lagrangian for the free particle must be invariant with respect to (*i*) Translation in space (*ii*) Rotation in space about any axis, and (*iii*) Translation in time.

In other words, *L* cannot be an explicit function of $\vec{r}$ and *t*. Further, *L* should not depend upon the direction of the velocity of the particle. Thus, *L* can only be a function of the magnitude of the velocity of the particle. Thus, we may write

$$L = L(v^2) \tag{1.135}$$

According to Equation (1.125), Lagrange's equation of motion for the system is given by

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \vec{v}}\right) = \frac{\partial L}{\partial \vec{r}}$$

or,
$$\frac{d}{dt}\left(\frac{\partial L}{\partial \vec{v}}\right) = 0 \left(\because \frac{\partial L}{\partial \vec{r}} = 0\right)$$

or,
$$\frac{\partial L}{\partial \vec{v}} = \text{A constant of motion} \qquad (1.136)$$

Since L is a function of only velocity, the above equation leads to

$$\vec{v} = \text{Constant of motion} \qquad (1.137)$$

Equation (1.137) is the **law of inertia** according to which a particle which moves without the influence of any external agent has a constant velocity vector.

To find the exact form of the function $L(v^2)$, consider two inertial frames of reference S and S', where S' moves with an infinitesimal uniform velocity, say $\vec{\varepsilon}$, with respect to S. Let L and L' be the Lagrangians of the particle as observed from the frames S and S', respectively. Since the equations of motion remain the same in all inertial frames, we must have L and L', different by only a total time derivative of a function of coordinates and time. We have according to Equation (1.127)

$$L' = L\left(v'^2\right) = L\left(v^2 + 2\vec{v}\cdot\vec{\varepsilon} + \varepsilon^2\right)$$

$$= L\left(v^2\right) + 2\vec{v}\cdot\vec{\varepsilon}\,\frac{\partial L}{\partial v^2} + \text{terms with higher order of } \varepsilon$$

$$= L + 2\vec{v}\cdot\vec{\varepsilon}\,\frac{\partial L}{\partial v^2}$$

Thus,
$$L' - L = 2\vec{v}\cdot\vec{\varepsilon}\,\frac{\partial L}{\partial v^2} \qquad (1.138)$$

Now for $L' - L$ to be a total time derivative of a function of coordinates and time, we must have $L' - L$ as a linear function of $\vec{v}$, i.e., $\frac{\partial L}{\partial v^2}$ to be independent of the velocity. Further, $\frac{\partial \vec{\varepsilon}}{\partial t} = 0$. We thus find that L must be proportional to v^2, i.e., $L \propto v^2$. To assign a physical meaning to L, we take the constant of proportionality as $\frac{1}{2}m$, so that we can write

$$L = \frac{1}{2}mv^2 = \text{Kinetic energy of the particle.} \qquad (1.139)$$

In order to obtain the Lagrangian function for an assembly of particles (non-interacting or interacting) we need to consider the important properties which the Lagrangian of the system possesses.

Check Your Progress

6. What are the Lagrange's equations?

7. What is the equation for the Lagrangian function of a system?

8. What is gauge function?

9. Define law of inertia.

1.8 TWO BODY CENTRAL FORCE PROBLEM

The particles (1 and 2) exert forces on each other, and there are no external forces. The potential energy is U(|r|) where r = r1 − r2 ; the forces are central and spherically symmetric ; r = | r | ; U = U(r)
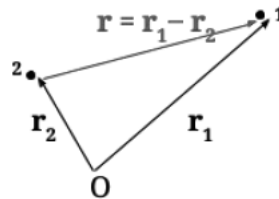
***Fig. 1.9** Particles 1 and 2*

Astronomical Problem

The potential energy is U(r) = − G m1 m2 /r .
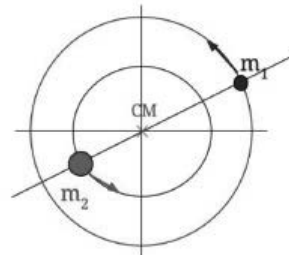
***Fig 1.10** Center of mass of m1 and m2*

Notes:

- The orbits are not circular in general.

- Sun and Earth, or another planet; m2m1

- Earth and Moon, or a satellite; m2>>m1
- Binary Star; m2 and m1 are comparable.

Atomic Problem

The hydrogen atom U(r) = − k e2 /r

Diatomic molecule e.g., O2

U(r) = A1/r12 − A2/r6
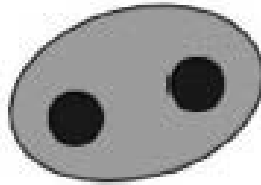
Fig 1.11 *Diatomic molecules*

1.9 KEPLER'S PROBLEM

Based on the observations made by Tycho Brahe, Kepler enunciated the following three laws for the motion of planets round the sun.

1st Law: Each planet moves in an elliptical path with the sun at one of the foci of the ellipse.

2nd Law: The area swept by the radius vector (the line joining the sun to the planet) in equal intervals of time is equal, i.e., the areal velocity of the planet is a constant.

3rd Law: The square of the time period of revolution of the planet round the sun is directly proportional to the cube of the semi-major axis of the ellipse.

Derivation of Kepler's Laws

Consider a planet of mass M_p moving under the gravitational attraction of the sun of mass M_s. The force on the planet towards the sun when its distance from the sun is r is given by

$$F(r) = -G\frac{M_s M_p}{r^2} = -\frac{k}{r^2} \qquad (1.140)$$

where $\qquad k = GM_s M_p = \text{a constant} \qquad (1.141)$

The potential energy $U(r)$ corresponding to the force $F(r)$ is

$$U(r) = -\int F(r)\,dr = k\int\frac{dr}{r^2} = -\frac{k}{r} \qquad (1.142)$$

In terms of the variable u, introduced in the previous section, we may write Equation (1.142) as

$$U\left(\frac{1}{u}\right) = -ku \qquad (1.143)$$

Substituting Equation (1.143) in the differential equation of motion given by Equation,

$$\frac{d^2u}{d\theta^2} + u = -\frac{m}{M^2}\frac{d}{du}U\left(\frac{1}{u}\right)$$

we get

$$\frac{d^2u}{d\theta^2} + u = -\frac{m}{M^2}\frac{d}{du}(-ku)$$

or $$\frac{d^2u}{d\theta^2} + u = \frac{km}{M^2} \qquad (1.144)$$

The most general solution of the Equation (1.144) is

$$u = \frac{mk}{M^2} + u_o\cos(\theta - \theta_o) \qquad (1.145)$$

where u_o and θ_0 are constants. By orienting the coordinate system properly, let us, for convenience, choose the constant θ_0 equal to zero so that the Equation (1.145) takes the form
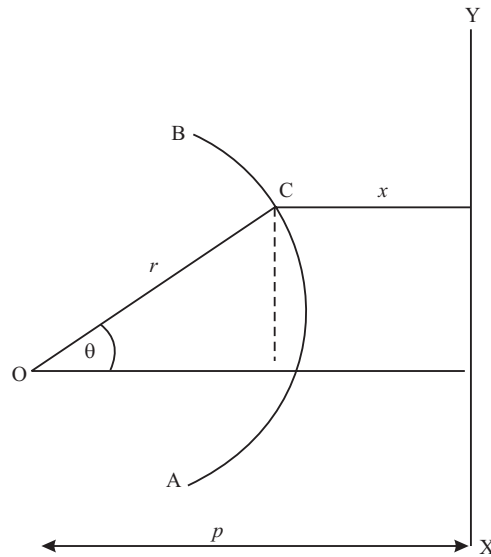
Fig. 1.12 *Conic Section*

$$\frac{1}{r} = \frac{km}{M^2} + u_o\cos\theta$$

or $$r = \frac{1}{\dfrac{mk}{M^2} + u_o\cos\theta}$$

or $$r = \frac{\dfrac{M^2}{mk}}{1 + \dfrac{u_oM^2}{mk}\cos\theta} \qquad (1.146)$$

We find that for $\theta = 0$, r is the maximum, while for $\theta = 0 + \pi$, r is the minimum. We can thus interpret θ_0, which we have taken as zero, to represent one of the angles corresponding to a turning point in the path of motion.

Equation (1.146) can be compared with the equation of a conic section which is a curve AB as shown in the Figure 1.12. In the figure, O is a fixed point called the focus and XY a fixed line called the directrix of the conic section. Let C be any arbitrary point on the curve AB.

$$OC = r\,(\text{say})$$

Let CD be the perpendicular from the point C on the directrix.

$$CD = x\,(\text{say})$$

Let the line OC make an angle θ with the line drawn normal from the focus to the directrix.

The curve AB is such that the ratio $\dfrac{r}{x}$ is a constant. This constant ratio is called the eccentricity of the conic section and is usually denoted by the symbol ε.

Let p be the distance of the directrix from the focus. We then get according to the Figure 1.12.

$$p = x + r\cos\theta = \frac{r}{\varepsilon} + r\cos\theta$$

Semilatus rectum, which we denote by the symbol ρ of the conic section is defined as

$$\rho = \varepsilon p\ (= \text{constant})$$

The above gives

$$p = \frac{\rho}{\varepsilon}$$

Thus, we get

$$\frac{\rho}{\varepsilon} = \frac{r}{\varepsilon} + r\cos\theta = \frac{r}{\varepsilon}[1 + \varepsilon\cos\theta]$$

or

$$r = \frac{\rho}{1 + \varepsilon\cos\theta} \qquad (1.147)$$

Equation (1.147) is the general equation of a conic section. Comparing Equation (1.146) with Equation (1.147), we find that if motion takes place under a central attractive force varying inversely as the square of the distance from the force centre then the path is a conic section having the focus at the force centre; the eccentricity and the semilatus rectum of the conic section being given by

$$\varepsilon = \frac{u_o M^2}{mk} \qquad (1.148)$$

$$\rho = \frac{M^2}{mk} \qquad (1.149)$$

To know exactly the eccentricity of the conic section in which the motion takes place we are required to find the constant u_o in terms of known quantities.

For the motion which is under consideration, the total energy is given by

$$E = \frac{1}{2}m\dot{r}^2 + \frac{M^2}{2mr^2} - \frac{k}{r} = \text{constant} \qquad (1.150)$$

At the turning point corresponding to $r = r_{min}$, we get according to Equation (1.150)

$$E = \frac{M^2}{2mr_{min}^2} - \frac{k}{r_{min}} \qquad (1.151)$$

From Equation (1.147) we have

$$r_{min} = \frac{\rho}{1+\varepsilon} = \frac{M^2}{mk(1+\varepsilon)} \qquad (1.152)$$

Using Equation (1.152) in Equation (1.151) we obtain

$$E = \frac{M^2 m^2 k^2 (1+\varepsilon)^2}{2mM^4} - \frac{kmk(1+\varepsilon)}{M^2}$$

or

$$E = \frac{1}{2}mk^2 \frac{(1+\varepsilon)^2}{M^2} - \frac{mk^2}{M^2}(1+\varepsilon)$$

or

$$E = \frac{mk^2}{2M^2}\left[1+\varepsilon^2+2\varepsilon-2-2\varepsilon\right] = \frac{mk^2}{2M^2}\left(\varepsilon^2-1\right)$$

or

$$\varepsilon = \left[1+\frac{2M^2 E}{mk^2}\right]^{\frac{1}{2}} \qquad (1.153)$$

Substituting for ε given by Equation (1.153), the equation for the conic section in which the motion takes place is given by

$$r = \frac{\dfrac{M^2}{mk}}{1+\left[1+\dfrac{2M^2 E}{mk^2}\right]^{\frac{1}{2}}\cos\theta} \qquad (1.154)$$

From Equation (1.153) we find that the eccentricity ε and hence the nature of the conic section is primarily decided by the total energy E. We the get

(*i*) For $E > 0$, i.e., the total energy being positive, the eccentricity is greater than 1 and the conic section is a hyperbola,

(*ii*) For $E = 0$, the eccentricity is 1 and the conic section is a parabola,

(*iii*) For $E < 0$, i.e., the total energy being negative, the eccentricity is less than 1 and the conic section is an ellipse,

(*iv*) For eccentricity equal to 0, the conic section is a circle.

For the motion of the planet in the gravitational field of the sun which is being considered presently we have the following:

Kinetic energy and potential energy of the planet when it is at a distance r from the sun and has velocity v are

$$T = \frac{1}{2}M_p v^2$$

and
$$U = -G\frac{M_s M_p}{r^2}$$

Clearly, the total energy of the planet is

$$E = T + U = \frac{1}{2}M_p v^2 - G\frac{M_s M_p}{r} \qquad (1.155)$$

The necessary centripetal force for the planet to move along the conic is provided by the gravitational force of attraction on the planet due to the sun. Thus, we have

$$\frac{M_p v^2}{r} = G\frac{M_s M_p}{r^2}$$

or
$$M_p v^2 = G\frac{M_s M_p}{r} \qquad (1.156)$$

Using Equation (1.156) in Equation (1.157), we get

$$E = \frac{1}{2}G\frac{M_s M_p}{r} - G\frac{M_s M_p}{r} = -\frac{1}{2}G\frac{M_s M_p}{r} \qquad (1.157)$$

We find the total energy E of the planet to be negative. Clearly, the eccentricity of the conic section is less than 1, and consequently the planet goes round in an elliptic path with the sun at one of its foci. This is Kepler's first law.

We have seen that in the case of motion under central force, the angular momentum is a constant of the motion Equation

$$p_\theta = \frac{\partial L}{\partial \dot\theta} = mr^2\dot\theta = \text{constant} = M(\text{say})$$
$$mr^2\dot\theta = \text{constant} = M \qquad (1.158)$$
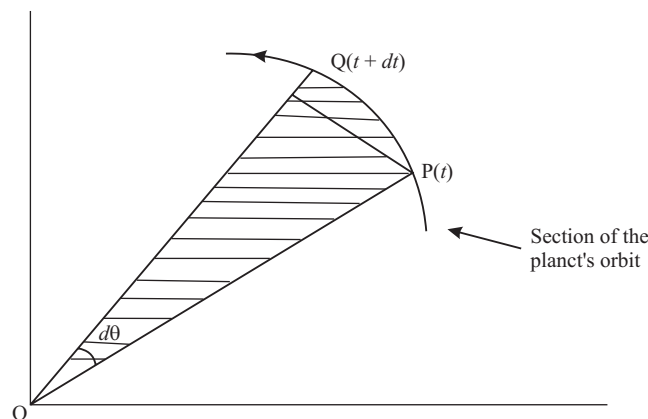
***Fig. 1.13** Positions of the Planet*

Consider Figure 1.13, in which the positions of the planet on its path of motion at two instants of time t and $t + dt$ are shown. During the interval

dt the area swept *dA* by the radius vector is the area of the shaded region. Since *dt* is infinitesimally small, the arc *PQ* can be considered as a straight line. Hence, we get

$$dA = \text{Area of the triangle } OPQ$$

$$= \frac{1}{2}rrd\theta = \frac{1}{2}r^2 d\theta$$

Thus, the areal velocity of the planet is

$$\frac{dA}{dt} = \frac{1}{2}r^2\frac{d\theta}{dt} = \frac{1}{2}r^2\dot{\theta}$$

In view of Equation (1.158), the above becomes

$$\frac{dA}{dt} = \frac{1}{2}\frac{M}{m} = \text{constant} \tag{1.159}$$

The above is the Kepler's second law of planetary motion.

Kepler's third law can be proved as follows:

If *a* be the semi-major axis of the ellipse in which the planet moves, we get by definition

$$a = \frac{\rho}{1-\varepsilon^2} \tag{1.160}$$

Substituting for ρ given by Equation (1.149), we get

$$a = \frac{M^2}{mk\left(1-\varepsilon^2\right)}$$

or

$$1 - \varepsilon^2 = \frac{M^2}{mka} \tag{1.161}$$

Using Equation (1.161) in the expression for *E* given by Equation (1.153), we get

$$E = -\frac{mk^2}{2M^2}\frac{M^2}{mka} = -\frac{k}{2a} \tag{1.162}$$

The semi-minor axis *b* of the ellipse is related to *a* and ε as

$$b = a\left(1-\varepsilon^2\right)^{\frac{1}{2}} \tag{1.163}$$

Using Equation (1.161) the above becomes

$$b = a\frac{M}{\left(mka\right)^{\frac{1}{2}}} = \frac{Ma^{\frac{1}{2}}}{\left(mk\right)^{\frac{1}{2}}} \tag{1.164}$$

Let *T* be the time period of revolution of the planet in its elliptic orbit. We then have

$$\text{Area of the ellipse } = \int_0^T \left(\frac{dA}{dt}\right) dt$$

or $$\pi ab = \frac{M}{2m} T$$

or $$T = \frac{2m}{M} \pi ab$$

Substituting for a and b in the above, we get

$$T = \frac{2m}{M} \pi a \frac{Ma^{\frac{1}{2}}}{(mk)^{\frac{1}{2}}} = \frac{2\pi m a^{\frac{3}{2}}}{(mk)^{\frac{1}{2}}}$$

Substituting for k, the above gives

or $$T^2 = \frac{4\pi^2}{(M_p + M_s)GM_sM_p} a^3$$

or $$T^2 = \frac{4\pi^2}{G(M_s + M_p)} a^3 \qquad (1.165)$$

Thus, we find that

$$T^2 \propto a^3 \qquad (1.166)$$

The above is the Kepler's third law of planetary motion.

Check Your Progress

10. Define Kepler's first law of planetary motion.

11. State Kepler's second law of planetary motion.

12. What is meant by Kepler's third law of planetary motion?

1.10 INVERSE SQUARE LAW OF FORCE

In science, an inverse-square law is any scientific law stating that a specified physical quantity is inversely proportional to the square of the distance from the source of that physical quantity. The fundamental cause for this can be understood as geometric dilution corresponding to point-source radiation into three-dimensional space.

Mathematically notated:

$$\text{intensity} \propto \frac{1}{\text{distance}^2}$$

It can also be mathematically expressed as:

$$\frac{\text{intensity}_1}{\text{intensity}_2} = \frac{\text{distance}_2^2}{\text{distance}_1^2}$$

or as the formulation of a constant quantity:

$$\text{intensity}_1 \times \text{distance}_1^2 = \text{intensity}_2 \times \text{distance}_2^2$$

The divergence of a vector field which is the resultant of radial inverse-square law fields with respect to one or more sources is everywhere proportional to the strength of the local sources, and hence zero outside sources. Newton's law of universal gravitation follows an inverse-square law, as do the effects of electric, light, sound, and radiation phenomena.

1.11 DEFINITION OF SCATTERING

The 'Scattering Theory' is significantly used for studying and understanding the scattering of waves and particles in mathematics and physics. Typically the wave scattering corresponds to the collision and scattering of a wave with some material object, for example formation of rainbow is resultant of sunlight scattered by rain drops. Latest technology of ultrasonic testing is another example of scattering theory which is used in medical imaging, non-destructive testing of metals and quantum field theory.

Rayleigh scattering is one commonly known type of scattering which mainly consists of scattering from atmospheric gases, it occurs when the particles causing scattering are smaller in size than the radiation wavelengths in contact with them.

Mie scattering, and non-selective scattering are the two other types of wave scattering. Principally, the Mie scattering is considered to be elastic scattered light of particles that have a diameter similar to or larger than the wavelength of the incident light. The Mie signal is proportional to the square of the particle diameter, where as in case of non-selective scattering also known as Raman scattering, it occurs in all wavelengths of electromagnetic radiation equally in the atmosphere and is usually caused by particles which are much larger than the energy wavelengths.

Definitions of Scattering

1. Scattering, in physics, is defined as a change in the direction of motion of a particle because of a collision with another particle. As defined in physics, a collision can occur between particles that repel one another, such as two positive (or negative) ions, and need not involve direct physical contact of the particles.

2. Scattering occurs when light or other energy waves pass through an imperfect medium, such as air filled with particles of some sort, and are deflected from a straight path. The light is deflected off of its straight path and scatters in many directions.

3. Scattering is a general physical process where some forms of radiation, such as light, sound, or moving particles, are forced to deviate from a straight trajectory by one or more paths due to localized non-uniformities in the medium through which they pass.

4. As per the Encyclopaedia Britannica, the 'Scattering, in physics, a change in the direction of motion of a particle because of a collision with another particle. A collision can occur between particles that repel one another, such as two positive (or negative) ions, and need not involve direct physical contact of the particles.

The physicist Ernest Rutherford passed a stream of alpha particles through a thin sheet of gold foil. The alpha particles were emitted by a radioactive material and had enough energy to penetrate an atom; although most passed right through the gold foil, some were deflected in a way that indicated that the scattering was produced by a Coulomb force. Because the alpha particles are positively charged and the electrons in the atom are negatively charged, it followed that there must be a large positive charge inside the atom to create the Coulomb force by interacting with the alpha particles. In this way the nucleus of the atom was discovered.

Elastic and Inelastic Scattering

The term 'Elastic Scattering' implies that the internal states of the scattering particles do not change, and hence they emerge unchanged from the scattering process. In inelastic scattering, by contrast, the particles' internal state is changed, which may amount to exciting some of the electrons of a scattering atom, or the complete annihilation of a scattering particle and the creation of entirely new particles.

When two atoms are scattered off one another, one can understand them as being the bound state solutions of some differential equation. Thus, for example, the hydrogen atom corresponds to a solution to the Schrödinger equation with a negative inverse-power, i.e., attractive Coulombic, central potential. The scattering of two hydrogen atoms will disturb the state of each atom, resulting in one or both becoming excited, or even ionized, representing an inelastic scattering process.

1.11.1 Scattering in a Central Force Field

In the case of a repulsive central potential, it becomes obvious that the two particles will not orbit each other - they will at most approach each other, before the repulsive potential causes them to move away from each other, and never meet again. This type of behaviour is typically referred to as scattering, an example of which is illustrated in Figure 1.14. This type of scenario is incredibly important in a wide range of physics, especially in condensed matter systems (where neutrons being scattered off of a material reveal information about the microscopic details of the material) and in high energy physics (where scattering elementary particles against each other can reveal information about the existence of new fundamental particles). For this reason, we want to understand how to describe the basic physics of such a system.
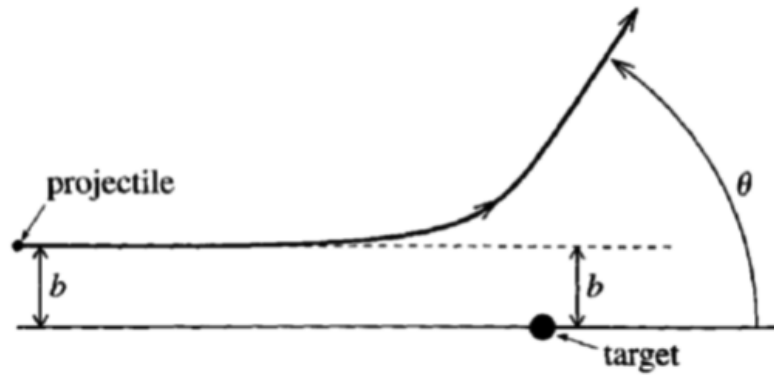
Fig. 1.14 *A projectile, with impact parameter b, being scattered by a stationary target*

After being scattered, the projectile travels off at some angle θ, never to return to the target. We will also assume, as before, that the two particles interact through a central potential, which in the case of a stationary target, essentially acts as an external central potential on the smaller body. For this reason, we know that we can still make use of linear momentum, angular momentum, and energy conservation in solving our problem, which means that all of the usual results from the study of the two-body problem should still be applicable. In the case that a collision between two particles conserves total energy, we typically refer to the scattering as being elastic.

Check Your Progress

13. What is scattering theory?
14. How does scattering occur?

1.12 ANSWERS TO 'CHECK YOUR PROGRESS'

1. The corresponding change in the configuration of the system subjected to arbitrary displacement is independent of time, i.e., no actual displacement of the system occurs with respect of time. Such displacements in the configuration space are called virtual displacements.

2. dW = 0

3. The principle can be stated as follows: The work done in infinitesimal reversible virtual displacements, consistent with the constraints, from the equilibrium configuration of a system is zero.

4. The principle may be stated as follows: For any dynamical system, the total work done by the effective force is zero in the course of reversible infinitesimal virtual displacement compatible with the constraints imposed on the system.

5. We can identify L as the Lagrangian function of the system. Thus, for a conservative system we obtain, L = Kinetic energy of the system – Potential energy of the system

6. The following set of s number of second order differential equations satisfied by the Lagrangian of the system are called the Lagrange's equations of motion.

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} = 0; \quad (k = 1, 2, \ldots, s)$$

7. The Lagrangian function of a system having s degrees of freedom and described by the generalized coordinates q1, ..., qs and the generalized velocities q1, ..., qs is given by

L = L(q1, ..., qs, q1, ..., qs, t) = L(qk, qk, t)

8. The arbitrary function F (q1, ..., qs, t) = F (qk, t) is called gauge function.

9. Law of inertia states that, a particle which moves without the influence of any external agent has a constant velocity vector.

10. Each planet moves in an elliptical path with the sun at one of the foci of the ellipse.

11. The area swept by the radius vector (the line joining the sun to the planet) in equal intervals of time is equal, i.e., the areal velocity of the planet is a constant.

12. The square of the time period of revolution of the planet round the sun is directly proportional to the cube of the semi-major axis of the ellipse.

13. The 'Scattering Theory' is significantly used for studying and understanding the scattering of waves and particles in mathematics and physics. Typically the wave scattering corresponds to the collision and scattering of a wave with some material object, for example formation of rainbow is resultant of sunlight scattered by rain drops. Latest technology of ultrasonic testing is another example of scattering theory which is used in medical imaging, nondestructive testing of metals and quantum field theory.

14. Scattering occurs when light or other energy waves pass through an imperfect medium, such as air filled with particles of some sort, and are deflected from a straight path. The light is deflected off of its straight path and scatters in many directions.

1.13 SUMMARY

- The Lagrangian of a system cannot be defined uniquely, but can be defined only within an additive total time derivative of any function of coordinates relevant to the system and time.

- If the Lagrangian is multiplied by any arbitrary constant then the equations of motion remain unaltered.

- A particle which moves without the influence of any external agent has a constant velocity vector.

- The corresponding change in the configuration of the system is independent of time, i.e., no actual displacement of the system occurs with respect to time.

- The work done in infinitesimal reversible virtual displacements, consistent with the constraints, from the equilibrium configuration of a system is zero.

- The variational principle finds immense usefulness in treating mechanical system on the one hand, while on the other hand, it considers the motion of the system as a whole between the given time limits along some small variation in the motion of the system between the same time limits from the actual motion.

- The 'Scattering Theory' is significantly used for studying and understanding the scattering of waves and particles in mathematics and physics. Typically the wave scattering corresponds to the collision and scattering of a wave with some material object.

1.14 KEY TERMS

- **Law of inertia:** It states that a particle which moves without the influence of any external agent has a constant velocity vector.

- **Virtual displacement:** A presumed infinitesimal change of system coordinates occurring while time is confined constant is known as virtual displacement.

- **Central force:** Central force is that force which acts either towards or away from a fixed point (called the centre of the force) and depends only on the distance from the fixed point.

- **Harmonic oscillator:** It is a system that, when displaced from its equilibrium position, encounters a restoring force proportional to the displacement.

- **Simple harmonic motion:** It is an oscillatory motion under a retarding force proportional to the quantity of displacement from an equilibrium position.

• **Scattering theory:** This is significantly used for studying and understanding the scattering of waves and particles in mathematics and physics.

• **Elastic scattering:** It implies that the internal states of the scattering particles do not change, and hence they emerge unchanged from the scattering process.

• **Inelastic scattering:** In this the particles' internal state is changed, which may amount to exciting some of the electrons of a scattering atom, or the complete annihilation of a scattering particle and the creation of entirely new particles.

1.15 SELF-ASSESSMENT QUESTIONS AND EXERCISES

Short Answer Questions

1. Give a brief account of important properties of Lagrangian function.

2. State Lagrangian of a particle moving freely in space.

3. Write a short note on virtual work.

4. State the motion of a linear harmonic oscillator.

5. Mention Lagrangian of a charged particle moving in an electromagnetic field.

6. State Rayleigh's dissipation function.

7. Derive Kepler's first law of planetary motion.

8. What is scattering theory?

9. Define elastic scattering.

Long Answer Questions

1. Derive Lagrange's equations for simple systems.

2. Deduce different mathematical forms of D'Alembert's principle.

3. Discuss applications of Lagrangian formulations.

4. Describe Lagrange's equations for conservative systems.

5. Discuss Kepler's laws of planetary motion.

6. Discuss the significance of term scattering theory giving appropriate examples.

7. Explain elastic and inelastic scattering.

1.16 FURTHER READING

Rao, K. Sankara. 2009. *Classical Mechanics*. New Delhi: PHI Learning Private Limited.

Upadhyaya, J.C. 2010. *Classical Mechanics, 2nd Edition*. New Delhi: Himalaya Publishing House.

Goldstein, Herbert. 2011. *Classical Mechanics, 3rd Edition*. New Delhi: Pearson Education India.

Gupta, S.L. 1970. *Classical Mechanics*. New Delhi: Meenakshi Prakashan.

Takwala, R.G. and P.S. Puranik. 1980. New Delhi: Tata McGraw Hill Publishing.

UNIT 2   HAMILTONIAN MECHANICS AND RIGID BODY

Structure

2.0   INTRODUCTION

William Rowan Hamilton formulated the principle of stationary action which expresses that the dynamics of a physical system is established by a variational problem for a functional based on a single function, the Lagrangian, which holds all physical information relating the system and the forces acting on it. Hamilton's principle is applicable to the electromagnetic and gravitational fields also. It contributes in quantum mechanics, quantum field theory and criticality theories significantly. In this unit you will study Hamiltonian of the system and the Legendre transformation relations for the change of basis. Hamilton's principle for conservative system and principle of least action is also discussed.

2.1   OBJECTIVES

After going through this unit, you will be able to:

- Explain Hamiltonian of the system and the method of Legendre transformation

- Describe Hamilton's principle for a conservative system

- Understand principle of least action and Lagrangian formulation of Mechanics

- Analyse Lagrange brackets and Poisson brackets

• Explain Eulerian angles, Euler's equations of motion of a rigid body and torque free motion of a rigid body

2.2 AN INTRODUCTION OF HAMILTONIAN EQUATION

Legendre transformation refers to the mathematical method for changing the basis of the description of a system from one set of independent variables to another set of independent variables.

Consider a function $f = f(x, y)$ of two independent variables x and y.

The total differential of f is

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy \tag{2.1}$$

Let us define

$$u = \frac{\partial f}{\partial x} \tag{2.2}$$

and

$$v = \frac{\partial f}{\partial y} \tag{2.3}$$

We may then write Equation (2.1) as

$$df = udx + vdy \tag{2.4}$$

Let us now consider u to be an independent variable and x a dependent variable in order to change our basis from the variables (x, y) to the variables (u, y).

Let $f' = f'(u, y)$ be a function of u and y defined according to

$$f' = f - ux \tag{2.5}$$

We then have

$$df' = df - udx - xdu \tag{2.6}$$

Using Equation (2.4) in Equation (2.6), we obtain

$$df' = udx + vdy - udx - xdu$$

or

$$df' = vdy - xdu \tag{2.7}$$

Since $f' = f'(u, y)$, we have

$$df' = \frac{\partial f'}{\partial u} du + \frac{\partial f'}{\partial y} dy \tag{2.8}$$

Comparing Equation (2.7) and (2.8) we obtain

$$x = -\frac{\partial f'}{\partial u} \tag{2.9}$$

and

$$v = \frac{\partial f'}{\partial y} \tag{2.10}$$

The relations given by Equation (2.9) and (2.10) are called the Legendre

transformation relations for the change of basis from (*x, y*) to (*u, y*). It is possible to extend the above method if we need to transform more than one variable.

Consider a mechanical system having *s* degrees of freedom. Let q_1, ….., q_s be the generalized coordinates that describe the system.

In the Lagrangian formulation of mechanics, the independent variables are the *s* generalized coordinates and time. The Lagrangian function *L* that characterizes the system is, in general, a function of the generalized coordinates, the generalized velocities and time, i.e.,

$$L = L(q_1, ..., q_s, \dot{q}_1, ..., \dot{q}_s, t) = L(q, \dot{q}, t) \qquad (2.11)$$

In Equation (2.11), *q* stands for all the coordinates and $\dot{q}$ stands for all the velocities.

We note that although the generalized velocities appear in the expression for *L* explicitly, they cannot be treated as independent variables because of being equal to the total time derivatives of the generalized coordinates. Hamilton developed an alternative formulation of mechanics by considering the independent variables for the system as the generalized co-ordinates (q_1, ….., q_s), the generalized momenta p_1, ….., p_s and time *t*. In this formulation, the generalized velocities are dependent functions such as

$$\dot{q}_k = \dot{q}_k(q_1, ..., q_s, p_1, ..., p_s, t) \qquad (2.12)$$

We may note that the generalized momenta are derived variables defined in terms of the Lagrangian *L* as

$$p_k = \frac{\partial L(q, \dot{q}, t)}{\partial \dot{q}_k} \qquad (2.13)$$

It follows from the above that to go over from the Lagrangian formulation to the Hamiltonian formulation, we need to change our basis of description of the system from $(q, \dot{q}, t)$ set to the (q, p, t) set. Such a change of basis can be carried out by the method of Legendre transformation discussed in the previous section.

A new function *H* = *H* (q_1, ….., q_s, p_1, ….., p_s, *t*) = *H* (*q, p, t*), which also characterizes the system under consideration, is defined in terms of the Lagrangian function of the system $L(q, \dot{q}, t)$ in a manner analogous to Equation (2.8) as

$$H(q, p, t) = \sum_{k=1}^{s} p_k \dot{q}_k - L(q, \dot{q}, t) \qquad (2.14)$$

The function *H*(*q, p, t*) given by Equation (2.14) is known as the Hamiltonian of the system.

2.2.1 Derivation of Hamiltonian Equation from Vibrational Principles

Consider a mechanical system of s degrees of freedom. Let $(q^1, \ldots, q^s)$ be the generalized coordinates, $(\dot{q}_1, \ldots, \dot{q}_s)$ be the generalized velocities and $(p^1, \ldots, p^s)$ be the generalized momenta for the system. The Lagrangian L of the system then is given by

$$L = L\left(q_1, \ldots, q_s, \dot{q}_1, \ldots, \dot{q}_s, t\right) = L\left(q_k, \dot{q}_k, t\right) \quad (2.15)$$

Hamilton's variational principle, or the Principle of Least Action, is stated as

$$\delta S = \delta \int_{t_1}^{t_2} L\left(q_k, \dot{q}_k, t\right) dt = 0 \qquad (2.16)$$

The Hamiltonian function $H = H(q^1, \ldots, q^s, p^1, \ldots, p^s, t) = H(q^k, p^k, t)$ is related to the Lagrangian L as

$$H = \sum_{k=1}^{s} p_k \dot{q}_k - L \qquad (2.17)$$

Using Equation (3) in Equation (2) we obtain

$$\delta S = \delta \int_{t_1}^{t_2} \left[\sum_k p_k \dot{q}_k - H\left(q_k, p_k, t\right) \right] dt = 0 \qquad (2.18)$$

or

$$\delta \sum_k \int_{t_1}^{t_2} p_k \dot{q}_k \, dt - \delta \int_{t_1}^{t_2} H dt = 0$$

or

$$\delta \sum_k \int_{t_1}^{t_2} p_k dq_k - \delta \int_{t_1}^{t_2} H dt = 0 \qquad (2.19)$$

Equations (2.18) and (2.19) are referred to as the **modified Hamilton's Principle**.

Let us label each path of the system in its configuration space between time limits t^1 and t^2 by a parameter α. We can then write the δ variation for the action S as

$$\delta S = d\alpha \cdot \frac{\partial S}{\partial \alpha} \qquad (2.20)$$

The above gives, for generality,

$$\delta \equiv d\alpha \cdot \frac{\partial}{\partial \alpha} \qquad (2.21)$$

Using Equation (2.18) in the Equation (2.20), we obtain

$$\delta S = d\alpha . \frac{\partial}{\partial \alpha} \left[\int_{t_1}^{t_2} \left\{ \sum_k p_k \dot{q}_k - H(q_k, p_k, t) \right\} dt \right] = 0 \qquad (2.22)$$

It is possible to introduce the differential operator $\dfrac{\partial}{\partial \alpha}$ inside the integral because the two time limits t^1 and t^2 are the same for all the paths and hence are independent of α. We thus obtain

$$\delta S = d\alpha \cdot \int_{t_1}^{t_2} \sum_k \left\{ \frac{\partial p_k}{\partial \alpha} \dot{q}_k + p_k \frac{\partial \dot{q}_k}{\partial \alpha} - \frac{\partial H}{\partial q_k} \frac{\partial q_k}{\partial \alpha} - \frac{\partial H}{\partial p_k} \frac{\partial p_k}{\partial \alpha} - \frac{\partial H}{\partial t} \frac{\partial t}{\partial \alpha} \right\} dt = 0 \qquad (2.23)$$

We now have
$$\int_{t_1}^{t_2} p_k \frac{\partial \dot{q}_k}{\partial \alpha} dt = \int_{t_1}^{t_2} p_k \frac{d}{dt}\left(\frac{\partial q_k}{\partial \alpha} \right) dt$$

Evaluating the integral on the right hand side of the above by parts we obtain

$$\int_{t_1}^{t_2} p_k \frac{\partial \dot{q}_k}{\partial \alpha} dt = \left\{ p_k \frac{\partial q_k}{\partial \alpha} \right\}_{t_1}^{t_2} - \int_{t_1}^{t_2} \dot{p}_k \cdot \frac{\partial q_k}{\partial \alpha} dt = -\int_{t_1}^{t_2} \dot{p}_k \cdot \frac{\partial q_k}{\partial \alpha} dt \qquad (2.24)$$

since $\dfrac{\partial q_k}{\partial \alpha} = 0$ at t^1 and at t^2.

Further, we have
$$\frac{\partial t}{\partial \alpha} = 0 \qquad (2.25)$$

since the time of travel along all the paths is the same.

Using the results given by Equation (2.24) and (2.25) in Equation (2.23) we obtain

$$d\alpha . \int_{t_1}^{t_2} \sum_k \left[\frac{\partial p_k}{\partial \alpha} \dot{q}_k - \dot{p}_k \frac{\partial q_k}{\partial \alpha} - \frac{\partial H}{\partial q_k} \frac{\partial q_k}{\partial \alpha} - \frac{\partial H}{\partial p_k} \frac{\partial p_k}{\partial \alpha} \right] dt = 0 \qquad (2.26)$$

In view of Equation (2.21) we may write

$$d\alpha \cdot \frac{\partial p_k}{\partial \alpha} = \delta p^k$$

$$d\alpha \cdot \frac{\partial q_k}{\partial \alpha} = \delta q^k \qquad (2.27)$$

Using Equation (2.27) in Equation (2.26) we get

$$\int_{t_1}^{t_2} \sum_k \left(\delta p_k \dot{q}_k - \dot{p}_k \delta q_k - \frac{\partial H}{\partial q_k} \delta q_k - \frac{\partial H}{\partial p_k} \delta p_k \right) dt = 0$$

or,
$$\int_{t_1}^{t_2} \sum_k \left\{ \left(\dot{q}_k - \frac{\partial H}{\partial p_k} \right) \delta p_k - \left(\dot{p}_k + \frac{\partial H}{\partial q_k} \right) \delta q_k \right\} dt = 0 \quad (2.28)$$

Since q^k and p^k are independent variables, their variations δq^k and δp^k are also independent. Hence, for Equation (2.28) to hold, the coefficients of δp^k and δq^k must separately vanish. We hence obtain

$$\dot{q}_k - \frac{\partial H}{\partial p_k} = 0 \qquad \text{or} \qquad \dot{q}_k = \frac{\partial H}{\partial p_k} \qquad (2.29)$$

$$\dot{p}_k + \frac{\partial H}{\partial q_k} = 0 \qquad \text{or} \qquad \dot{p}_k = -\frac{\partial H}{\partial q_k}$$

For each k, we have two equations of the form given by Equation (2.29). These $2s$ number of first-order differential equations are **Hamilton's equations of motion, also called Hamilton's canonical equations**.

Check Your Progress

1. What do you understand by Legendre transformation?
2. Write the equations for the Legendre transformation relations for the change of basis.
3. What is the Lagrangian function?
4. Give the function which is known as the Hamiltonian of the system.

2.3 PRINCIPLE OF LEAST ACTION

Lagrangian formulation of mechanics, which is an alternative to Newtonian formulation, is based on one of the fundamental variational principles given by Hamilton known as the Hamilton's variational principle. It is important to note that the principle is stated in a form which is independent of any coordinate system and as such the principle can be used for dealing with non-mechanical systems and fields as well.

According to Hamilton, every mechanical system possesses a characteristic function of coordinates, velocities and time called the Lagrangian of the system usually denoted by the symbol L. If for a dynamical system having s-degrees of freedom, $q_1,$ ....., q_s and $\dot{q}_1, ..., \dot{q}_s$ be respectively the generalized coordinates and generalized velocities (both the coordinates and the velocities may be implicit as well as explicit functions of time), then the Lagrangian of the system is given by

$$L = L\left(q_1, ..., q_s, \dot{q}_1, ..., \dot{q}_s, t \right) \qquad (2.30)$$

At any instant of time t, the configuration of the system can be represented by a point called the system point in the s-dimensional mathematical space, namely, the configuration space of the system. As time passes, the system point moves in the configuration space and traces out a definite curve or path during a definite interval of time. Hamilton's principle is concerned with the trajectory or the path which is followed by the system point.

The principle states that of all possible paths along which the system may move from one point to another in its configuration space between two given time instants, say t1 and t2, which are consistent with the constraints imposed on the system, if any, the actual path which the system follows is the one for which the time integral of the Lagrangian of the system is an extremum (either maximum or minimum).

Mathematically, the principle is stated as

$$S = \int_{t_1}^{t_2} L dt = \text{An extremum} \tag{2.31}$$

The line integral $\int_{t_1}^{t_2} L dt$ which has been denoted above by the symbol

S, is called the Hamilton's principle function, or action integral, or simply the action during the time interval from t_1 to t_2.

In most of the dynamical problems, the minimum condition for the action S is satisfied. For this reason, the principle is also called Hamilton's principle of least action.

In terms of calculus of variation we can express Hamilton's principle given by Equation (2.31) as

$$\delta S = \delta \int_{t_1}^{t_2} L dt = 0$$

or $$\delta S = \delta \int_{t_1}^{t_2} L\left(q_1, ..., q_s, \dot{q}_1, ..., \dot{q}_s, t\right) = 0 \tag{2.32}$$

Check Your Progress

5. Give mathematical statement for Hamilton's principle for a conservative system.

6. What do you understand by Hamilton's principle for a conservative system?

7. Why Hamilton's principle for a conservative system is also called Hamilton's principle of least action?

2.4 EQUATION OF CANONICAL TRANSFORMATION

Canonical equations of motion can be derived alternatively, using the definitions of Lagrangian function, Hamiltonian function and Lagrange's equations of motion.

We have the Lagrangian function of the system under consideration given by

$$L = L(q_1, \ldots, q_s, \dot{q}_1, \ldots, \dot{q}_s, t)$$

The total differential of L is thus

$$dL = \sum_k \frac{\partial L}{\partial q_k} dq_k + \sum_k \frac{\partial L}{\partial \dot{q}_k} d\dot{q}_k + \frac{\partial L}{\partial t} dt \qquad (2.33)$$

The Hamiltonian function of the system is

$$H = H(q^1, \ldots, q^s, p^1, \ldots, p^s, t)$$

The total differential of H is thus

$$dH = \sum_k \frac{\partial H}{\partial q_k} dq_k + \sum_k \frac{\partial H}{\partial p_k} dp_k + \frac{\partial H}{\partial t} dt \qquad (2.34)$$

Further, H is related to L according to

$$H = \sum p_k \dot{q}_k - L$$

So, we obtain

$$dH = \sum p_k d\dot{q}_k + \sum \dot{q}_k dp_k - dL$$

Substituting for dL from Equation (2.33) in the above, we get

$$dH = \sum p_k d\dot{q}_k + \sum \dot{q}_k dp_k - \sum \frac{\partial L}{\partial q_k} dq_k - \sum \frac{\partial L}{\partial \dot{q}_k} d\dot{q}_k - \frac{\partial L}{\partial t} dt \qquad (2.35)$$

Using $p_k = \dfrac{\partial L}{\partial \dot{q}_k}$, the above equation reduces to

$$dH = \sum \frac{\partial L}{\partial \dot{q}_k} d\dot{q}_k + \sum \dot{q}_k dp_k - \sum \frac{\partial L}{\partial q_k} dq_k - \sum \frac{\partial L}{\partial \dot{q}_k} d\dot{q}_k - \frac{\partial L}{\partial t} dt$$

or $$dH = \sum \dot{q}_k dp_k - \sum \frac{\partial L}{\partial q_k} dq_k - \frac{\partial L}{\partial t} dt \qquad (2.36)$$

The Lagrange's equation is given by

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k} \right) = \frac{\partial L}{\partial q_k}$$

or $$\frac{d}{dt}(p_k) = \frac{\partial L}{\partial q_k} \quad \text{or} \quad \dot{p}_k = \frac{\partial L}{\partial q_k} \qquad (2.37)$$

Equation (2.37) used in Equation (2.36) gives

$$dH = \sum \dot{q}_k dp_k - \sum \dot{p}_k dq_k - \frac{\partial L}{\partial t} dt \qquad (2.38)$$

Comparing the coefficients of dp^k, dq^k and dt on the right hand sides of Equation (2.34) and (2.38), we obtain

$$\dot{q}_k = \frac{\partial H}{\partial p_k} \quad (k = 1, 2 \ldots, s) \qquad (2.39)$$

$$\dot{p}_k = -\frac{\partial H}{\partial q_k}$$

and
$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \qquad (2.40)$$

Equation (2.39) is the Hamilton's canonical equation of motion as obtained earlier.

2.5 LAGRANGE BRACKETS

Consider a mechanical system of s degree of freedom. Let the system be described by generalized coordinates q_1, ..., qs and conjugate momenta p_1, ..., ps.

Let $f = f(q_1, \ldots, qs, p_1, \ldots, ps) = f(q, p)$ and $g = g(q_1, \ldots, qs, p_1, \ldots, ps) = g(q, p)$ be two dynamical variables of the system.

The Lagrange bracket of f and g with respect to the basis (q, p) is written as $\{f, g\}q_{,} p$ and is defined as

$$\{f, g\}q_{,} p = \sum_k \left(\frac{\partial q_k}{\partial f} \frac{\partial p_k}{\partial g} - \frac{\partial p_k}{\partial f} \frac{\partial q_k}{\partial g} \right) = 0 \qquad (2.41)$$

We may note the following

(a) Taking $f = qi$ and $g = qj$ Equation (2.41) gives

$$\{qi, qj\} = \sum_k \left(\frac{\partial q_k}{\partial q_i} \frac{\partial p_k}{\partial q_j} - \frac{\partial p_k}{\partial q_i} \frac{\partial q_k}{\partial q_j} \right) = 0 \qquad (2.42)$$

(b) Taking $f = pk$ and $g = pj$ Equation (2.41) gives

$$\{pk, pj\} = \sum_k \left(\frac{\partial q_k}{\partial p_k} \frac{\partial p_k}{\partial p_j} - \frac{\partial p_k}{\partial p_k} \frac{\partial q_k}{\partial p_j} \right) = 0 \qquad (2.43)$$

(*c*) Taking $f = qk$ and $g = pj$ Equation (2.41) gives

$$\{qk, pj\} = \sum_k \left(\frac{\partial q_k}{\partial q_k} \frac{\partial p_k}{\partial p_j} - \frac{\partial p_k}{\partial q_k} \frac{\partial q_k}{\partial p_j} \right) = \delta_{kj} \qquad (2.44)$$

We may further note the following important **properties of Lagrange bracket**.

(*d*) Lagrange bracket is invariant under canonical transformation from the set of variables (*q, p*) to the set of variables (*Q, P*), i.e.,

$$\{f, g\}q,p = \{f, g\}Q,P \qquad (2.45)$$

(*e*) Lagrange bracket is non-commutative. Thus, we have

$$\{f, g\} = -\{g, f\} \qquad (2.46)$$

(*f*) The following general theorem that relates Lagrange bracket and Poisson bracket is found to hold

$$\sum_{k=1}^{2n} \{f_k f_i\}\left[f_k, f_j \right] = \delta ij \qquad (2.47)$$

where $f, f_2, ..., f_2n$ is a set of 2*n* independent functions, each of which is itself a function of *n* coordinates $q_1, ..., qn$ and *n* momenta $p_1, ..., pn$.

2.6 POISSON BRACKETS

Consider a mechanical system of *s* degrees of freedom. Let $q_1, ..., qs$ be the generalized coordinates, and $p_1, ..., ps$ be the generalized momenta in terms of which the system is described. Let *F* be any dynamical variable of the system which is a function of the coordinates, momenta and time, i.e.,

$$F = F (q_1, ..., qs, p_1, ..., ps, t) = F (q, p, t) \qquad (2.48)$$

The total time derivative of *F* is given by

$$\frac{dF}{dt} = \sum_k \frac{\partial F}{\partial q_k} \dot{q}_k + \sum_k \frac{\partial F}{\partial p_k} \dot{p}_k + \frac{\partial F}{\partial t} \qquad (2.49)$$

Using the Hamilton's canonical equations given by

$$\dot{q}_k = \frac{\partial H}{\partial p_k}; \; \dot{p}_k = \frac{-\partial H}{\partial q_k}$$

in Equation (2.49), we obtain

$$\frac{dF}{dt} = \sum_k \frac{\partial F}{\partial q_k} \frac{\partial H}{\partial p_k} - \sum \frac{\partial F}{\partial p_k} \frac{\partial H}{\partial q_k} + \frac{\partial F}{\partial t}$$

or

$$\frac{dF}{dt} = \sum_k \left[\frac{\partial F}{\partial q_k} \frac{\partial H}{\partial p_k} - \frac{\partial F}{\partial p_k} \frac{\partial H}{\partial q_k} \right] + \frac{\partial F}{\partial t}$$

or

$$\frac{dF}{dt} = \left[F, H \right] + \frac{\partial F}{\partial t} \qquad (2.50)$$

The quantity within the parenthesis on the right hand side of Equation (2.50) turns out to be of fundamental importance in the formal development of mechanics and is known as the Poisson bracket (PB) of F and H. It is usual to write it as $[F, H]q\,p$. Thus,

$$[F, H]q\,p = \sum_k \left[\frac{\partial F}{\partial q_k} \frac{\partial H}{\partial p_k} - \frac{\partial F}{\partial p_k} \frac{\partial H}{\partial q_k} \right] \qquad (2.51)$$

In general, for any two arbitrary physical quantities f and g, which are functions of coordinates, momenta and time, the Poisson bracket is defined as

$$[f, g]q\,p = \sum_k \left[\frac{\partial f}{\partial q_k} \frac{\partial g}{\partial p_k} - \frac{\partial f}{\partial p_k} \frac{\partial g}{\partial q_k} \right] \qquad (2.52)$$

Some of the special cases of Equation (2.50) give

(i) $\dot{q}_k = [qk, H]$ (choosing $F = qk$) $\qquad (2.53)$

(ii) $\dot{p}_k = [p_k, H]$ (choosing $F = p_k$) $\qquad (2.54)$

Again, from Equation (2.52) it follows that the PB of a quantity with itself is zero. Hence, we obtain from Equation (2.50)

(iii) $\dot{H} = \dfrac{dH}{dt} = \dfrac{\partial H}{\partial t} + [H, H] = \dfrac{\partial H}{\partial t}$,   (choosing $F = H$) $\qquad (2.55)$

We may note the following identities from the general definition of Poisson bracket given by Equation (2.52).

$$[f, g] = -[g, f]$$

$$[f, c] = 0 \qquad (2.56)$$

$$[cf, g] = c[f, g]$$

In the above, c is a constant

$$[f, g_1 + g_2] = [f, g_1] + [f, g_2]$$

$$[f, g_1 g_2] = g_1[f, g_2] + [f, g_1]g_2 \qquad (2.57)$$

$$\frac{\partial}{\partial t}[f, g] = \left[f, \frac{\partial g}{\partial t} \right] + \left[\frac{\partial f}{\partial t}, g \right] \qquad (2.58)$$

Furthermore, some of the special cases of Equation (2.55) and (2.56) are easily seen to follow.

(*a*) Taking $g = qj$ in Equation (2.52), we get

$$[f, qj] = \sum_k \left[\frac{\partial f}{\partial q_k} \frac{\partial q_j}{\partial p_k} - \frac{\partial f}{\partial p_k} \frac{\partial q_j}{\partial q_k} \right]$$

$$= -\sum_k \frac{\partial f}{\partial p_k} \delta_{jk} \qquad \left[\because \frac{\partial q_j}{\partial p_k} = 0 \; ; \; \begin{array}{l} \delta_{jk} = 0 \ \text{if} \ j \neq k \\ = 1 \ \text{if} \ j = k \end{array} \right]$$

Thus, $\qquad [f, qj] = -\dfrac{\partial f}{\partial p_j}$ $\qquad\qquad\qquad$ (2.59)

(*b*) Taking $f = qk$ and $g = qj$, we obtain from Equation (2.52)

$$[qk, qj] = -\frac{\partial q_k}{\partial p_j} = 0 \qquad\qquad\qquad (2.60)$$

Similarly, taking $g = qj$, we obtain

$$[pk, qj] = -\frac{\partial p_k}{\partial p_j} = -\delta_{kj} \qquad\qquad\qquad (2.61)$$

(*c*) If $g = pj$, we get from Equation (2.56)

$$[f, pj] = \sum_k \left[\frac{\partial f}{\partial q_k} \frac{\partial p_j}{\partial p_k} - \frac{\partial f}{\partial p_k} \frac{\partial p_j}{\partial q_k} \right] = \sum \frac{\partial f}{\partial q_k} \delta_{jk}$$

or $\qquad\qquad [f, pj] = \dfrac{\partial f}{\partial q_j}$ $\qquad\qquad\qquad$ (2.62)

If $g = pj$, then

$$[pi, pj] = \frac{\partial p_i}{\partial p_j} = 0 \qquad\qquad\qquad (2.63)$$

If $f = qi$, then

$$[qi, pj] = \frac{\partial q_i}{\partial q_j} = \delta_{ij} \qquad\qquad\qquad (2.64)$$

The above results are summarized as

$$[qi, qj] = 0 = [pi, pj] \qquad\qquad\qquad (2.65)$$
$$[qi, pj] = \delta ij$$

Equations (2.65) are known as the **Fundamental, or basic, Poisson brackets**.

Constants or Integrals of Motion

Consider Equation (2.50) which gives the total time derivative of the dynamical variable $F = F(q, p, t)$

$$\frac{dF}{dt} = \sum_k \left(\frac{\partial F}{\partial q_k} \frac{\partial H}{\partial p_k} - \frac{\partial F}{\partial p_k} \frac{\partial H}{\partial q_k} \right) + \frac{\partial F}{\partial t}$$

or
$$\frac{dF}{dt} = [F,H]_{q,p} + \frac{\partial F}{\partial t} \qquad (2.66)$$

If F has no explicit time dependence, we have $\frac{\partial F}{\partial t} = 0$ and Equation

(2.66) reduces to

$$\frac{dF}{dt} = [F, H]q_{,}p \qquad (2.67)$$

If the variable F under consideration is such that its Poisson bracket with the Hamiltonian of the system vanishes, i.e., $[F, H]q_{,}p = 0$, we get from Equation (2.67)

$$\frac{dF}{dt} = 0 \qquad (2.68)$$

or
$$F = \text{A constant of motion}$$

We thus find that a dynamical variable of a mechanical system is a constant of motion or an integral of motion, provided that

 (*i*) it has no explicit time dependence, and

 (*ii*) its Poisson bracket with the Hamiltonian of the system vanishes.

2. Canonical Transformation and Poisson Bracket

Let us consider a mechanical system of s degrees of freedom described by generalized coordinates $q_1, \dots, q_s$ and generalized momenta $p_1, \dots, p_s$.

Consider two dynamical variables f anf g which are functions of the q's and p's. The Poisson bracket of f anf g is, by definition,

$$[f,g]_{q,p} = \sum_k \left(\frac{\partial f}{\partial q_k} \frac{\partial g}{\partial p_k} - \frac{\partial f}{\partial p_k} \frac{\partial g}{\partial q_k} \right) \qquad (2.69)$$

Let us consider a canonical transformation of the variables q's and p's, respectively to Q's and P's. In terms of the transformed variables, the Poisson bracket of f and g is

$$[f,g]_{Q,P} = \sum_k \left(\frac{\partial f}{\partial Q_k} \frac{\partial g}{\partial P_k} - \frac{\partial f}{\partial P_k} \frac{\partial g}{\partial Q_k} \right) \qquad (2.70)$$

The Poisson bracket given by Equation (2.70) can alternatively be written as

$$[f,g]_{Q,P} = \sum_{k,j} \left[\frac{\partial f}{\partial Q_k} \left(\frac{\partial g}{\partial q_j} \frac{\partial q_j}{\partial P_k} + \frac{\partial g}{\partial p_j} \frac{\partial p_j}{\partial P_k} \right) - \frac{\partial f}{\partial P_k} \left(\frac{\partial g}{\partial q_j} \frac{\partial q_j}{\partial Q_k} + \frac{\partial g}{\partial p_j} \frac{\partial p_j}{\partial Q_k} \right) \right] \qquad (2.71)$$

On rearranging the terms, the above becomes

$$[f,g]_{Q,P} = \sum_k \left\{ \frac{\partial g}{\partial q_j} [f,q_j]_{Q,P} + \frac{\partial g}{\partial p_j} [f,p_j]_{Q,P} \right\} \qquad (2.72)$$

Replacing f by qj and g by f we obtain

$$\left[f,q_j\right]_{Q,P} = -\left[q_j,f\right]_{Q,P} = -\sum_i \left\{ \frac{\partial g_i}{\partial Q_i}\frac{\partial f}{\partial P_k} - \frac{\partial q_i}{\partial P_i}\frac{\partial f}{\partial Q_i} \right\}$$

$$= -\sum_{k,i}\left\{ \frac{\partial q_j}{\partial Q_i}\left(\frac{\partial f}{\partial q_k}\frac{\partial q_k}{\partial P_i} + \frac{\partial f}{\partial p_k}\frac{\partial p_k}{\partial P_i}\right) - \frac{\partial q_j}{\partial P_i}\left(\frac{\partial f}{\partial q_k}\frac{\partial q_k}{\partial Q_i} + \frac{\partial f}{\partial p_k}\frac{\partial p_k}{\partial Q_i}\right) \right\}$$

$$= -\sum_{k}\left\{ \frac{\partial f}{\partial q_k}\sum_i\left(\frac{\partial q_j}{\partial Q_i}\frac{\partial q_k}{\partial P_i} - \frac{\partial q_j}{\partial P_i}\frac{\partial q_k}{\partial Q_i}\right) + \frac{\partial f}{\partial p_k}\sum_i\left(\frac{\partial q_j}{\partial Q_i}\frac{\partial p_k}{\partial P_i} - \frac{\partial q_j}{\partial P_i}\frac{\partial p_k}{\partial Q_i}\right) \right\}$$

$$= -\sum_{k}\left\{ \frac{\partial f}{\partial q_k}\left[q_j,q_k\right]_{Q,P} + \frac{\partial f}{\partial p_k}\left[q_j,p_k\right]_{Q,P} \right\}$$

$$= -\sum_{k}\frac{\partial f}{\partial p_k}\delta_{jk}$$

using the properties $\left[q_i,q_k\right]=0$, $\left[q_j,p_k\right]=0$ if $j\neq k$, $[q_i,p_k]=1$ if $j=k]$

or
$$\left[f,q_j\right]_{Q,P} = -\frac{\partial f}{\partial p_j} \qquad (2.73)$$

Similarly, we get

$$\left[f,p_j\right]_{Q,P} = -\frac{\partial f}{\partial q_j} \qquad (2.74)$$

Using Equations (2.73) and (2.74) in Equation (2.71), we get

$$[f,g]_{Q,P} = \sum_j\left(-\frac{\partial g}{\partial q_j}\frac{\partial f}{\partial p_j} + \frac{\partial g}{\partial p_j}\frac{\partial f}{\partial q_j}\right)$$

or
$$[f,g]_{Q,P} = \sum_j\left(\frac{\partial f}{\partial q_j}\frac{\partial g}{\partial p_j} - \frac{\partial f}{\partial p_j}\frac{\partial g}{\partial q_j}\right) \qquad (2.75)$$

or
$$[f,g]_{Q,P} = [f,g]_{q,p} \qquad (2.76)$$

Thus, the Poisson bracket remains invariant under canonical transformation.

Check Your Progress

8. Define Lagrange brackets.

9. Write the equations of fundamental or basic Poisson brackets.

10. What do you mean by Poisson brackets?

2.7 ANGULAR MOMENTUM AND POISSON BRACKETS RELATION

The angular momentum is given as

$$\mathbf{J} = \mathbf{r} \times \mathbf{p} = \left(x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}} \right) \times \left(p_x\hat{\mathbf{i}} + p_y\hat{\mathbf{j}} + p_z\hat{\mathbf{k}} \right)$$

Or $$\mathbf{J} = \left(yp_z - zp_y \right)\hat{\mathbf{i}} + \left(zp_x - xp_z \right)\hat{\mathbf{j}} + \left(xp_y - yp_x \right)\hat{\mathbf{k}} \quad (2.77)$$

From equation (2.77),

$$J_x = yp_z - zp_y,\ J_y = zp_x - xp_z,\ J_z = xp_y - yp_x \quad (2.78)$$

Using the Poisson Bracket relation, $\left[A, p_k \right] = \dfrac{\partial A}{\partial q_k}$

$$[p_x, p_x] = [p_x, p_y] = [p_y, p_z] = [p_z, p_x] = [p_y, p_y] = [p_z, p_z] = 0 \quad (2.79)$$

Also, we can write,

$$[J_x, p_x] = \frac{\partial J_x}{\partial x} = \frac{\partial (yp_z - zp_y)}{\partial q_x} = 0 \quad (2.80)$$

$$[J_x, p_y] = \frac{\partial J_x}{\partial y} = \frac{\partial (yp_z - zp_y)}{\partial y} = p_z \quad (2.81)$$

$$[J_x, p_z] = \frac{\partial J_x}{\partial z} = \frac{\partial (yp_z - zp_y)}{\partial z} = -p_y \quad (2.82)$$

Similarly, the other relations can be written as follows:

$$[J_y, p_x] = \frac{\partial J_y}{\partial x} = \frac{\partial (zp_x - xp_z)}{\partial x} = -p_z \quad (2.83)$$

$$[J_y, p_y] = \frac{\partial J_y}{\partial y} = \frac{\partial (zp_x - xp_z)}{\partial y} = 0 \quad (2.84)$$

$$[J_y, p_z] = \frac{\partial J_y}{\partial z} = \frac{\partial (zp_x - xp_z)}{\partial z} = p_x \quad (2.85)$$

and

$$[J_z, p_x] = \frac{\partial J_z}{\partial x} = \frac{\partial (xp_y - yp_x)}{\partial x} = p_y \quad (2.86)$$

$$[J_z, p_y] = \frac{\partial J_z}{\partial y} = \frac{\partial (xp_y - yp_x)}{\partial y} = -p_x \quad (2.87)$$

$$[J_z, p_z] = \frac{\partial J_z}{\partial z} = \frac{\partial (xp_y - yp_x)}{\partial z} = 0 \quad (2.88)$$

Further, by using equation (2.77)

$$[J_x, J_y] = \Sigma \left(\frac{\partial J_x}{\partial q_i} \frac{\partial J_y}{\partial p_i} - \frac{\partial J_x}{\partial p_i} \frac{\partial J_y}{\partial q_i} \right) \tag{2.89}$$

Taking q_1=x, q_2=y, q_3=z and p_1=p_x, p_2=p_y, p_3=p_z

Equation (29) can be written as

or $\qquad [J_x, J_y] = 0 - 0 + 0 - 0 + (-p_y)(-x) - yp_x$

or $\qquad [J_x, J_y] = xp_y - yp_x$

or $\qquad [J_x, J_y] = J_z \tag{2.90}$

Similarly, we can derive

$$[J_y, J_z] = J_x \tag{2.91}$$

$$[J_z, J_x] = J_y \tag{2.92}$$

2.7.1 Equation of motion in Poisson Bracket form

$$A = q_i, \dot{q}_i = [q_i, H]$$
$$A = p_i, \dot{p}_i = [p_i, H]$$
$$F = H, \dot{H} = \frac{\partial H}{\partial t}$$

Properties of Poisson Bracket

Poisson Bracket has the property of antisymmetric, i.e.

[A,B]=-[B,A]

Also, the other identities of Poisson Bracket include:

[A,A]=0

[A,c]=0, where c is a constant.

[cA,B]=c[A,B]

[A_1+A_2,B]=[A_1,B]+[A_2,B]

$$[A, B_1 B_2] = B_1[A, B_2] + [A, B_1]B_2$$

$$\frac{\partial}{\partial t}[A, B] = \left[\frac{\partial A}{\partial t}, B\right] + \left[A + \frac{\partial B}{\partial t}\right]$$

$$[A, [B, C]] + [C, [A, B]] + [B, [C, A]] = 0, \textbf{(Jacobi's Identity)}$$

2.8 EULER EQUATION OF MOTION FOR A RIGID BODY AND ITS APPLICATIONS OF TORSION FREE SYMMETRIC RIGID BODY

The configuration of a rigid body which has six degrees of freedom is completely specified by locating the coordinates of a Cartesian system fixed in the body with respect to the coordinate axes of a system fixed in space external to the body.

In other words, it is usual to consider two coordinate systems for describing the motion of a rigid body:

(*i*) A space fixed system *XYZ* whose origin and whose axes are fixed in space, and

(*ii*) A body fixed system $X_1 X_2 X_3$ whose origin and axes are fixed within the body so that this system moves along with the body and it is usual to call it the moving system. The above two coordinate frames are shown in the Figure 2.1.
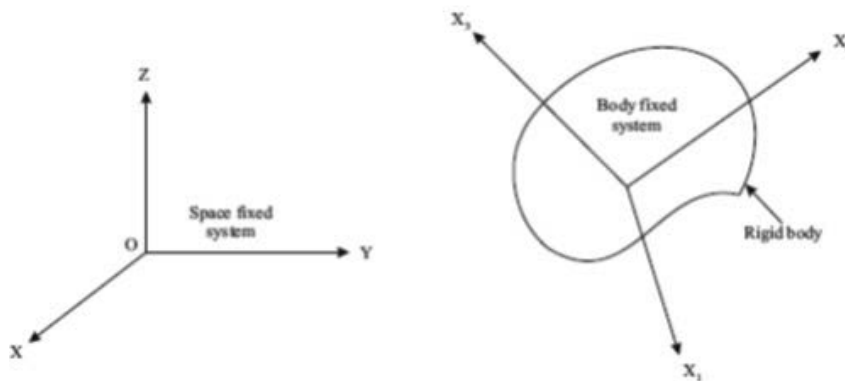
Fig. 2.1 *Space Fixed System and Body Fixed System*

As has been pointed out earlier, it is convenient to choose the origin of the moving system as the centre of mass of the body. Of the six generalized coordinates required to specify the configuration of the body, three are taken as the three Cartesian coordinates of the centre of mass of the body, i.e., the origin of the moving system with respect to the space-fixed system. About the remaining three let us look into the following:

Let

$\hat{i}, \hat{j}$ and $\hat{k}$ be, respectively, the unit vectors along the *X*-, *Y*- and *Z*-axes.

$\hat{i}', \hat{j}'$ and $\hat{k}'$ be respectively the unit vectors along the X_1-, X_2- and X_3-axes.

α_1, β_1, and γ_1 be the direction cosines of the X-axis relative to the X_1-, X_2- and X_3-axes, respectively.

α_2, β_2, and γ_2 be the direction cosines of the Y-axis relative to X_1-, X_2- and X_3-axes, respectively.

α_3, β_3 and γ_3 be the direction cosines of the Z-axis relative to X_1-, X_2- and X_3-axes, respectively.

We then have the relations

$$\hat{i}' = \alpha_1\hat{i} + \beta_1\hat{j} + \gamma_1\hat{k}$$

$$\hat{j}' = \alpha_2\hat{i} + \beta_2\hat{j} + \gamma_2\hat{k} \qquad (2.93)$$

$$\hat{k}' = \alpha_3\hat{i} + \beta_3\hat{j} + \gamma_3\hat{k}$$

We further have

$$\hat{i}'\cdot\hat{i}' = \alpha_1^2 + \beta_1^2 + \gamma_1^2 = 1$$

$$\hat{j}'\cdot\hat{j}' = \alpha_2^2 + \beta_2^2 + \gamma_2^2 = 1 \qquad (2.94)$$

$$\hat{k}'\cdot\hat{k}' = \alpha_3^2 + \beta_3^2 + \gamma_3^2 = 1$$

and

$$\hat{i}'\cdot\hat{j}' = \alpha_1\alpha_2 + \beta_1\beta_2 + \gamma_1\gamma_2 = 0$$

$$\hat{j}'\cdot\hat{k}' = \alpha_2\alpha_3 + \beta_2\beta_3 + \gamma_2\gamma_3 = 0 \qquad (2.95)$$

$$\hat{k}'\cdot\hat{i}' = \alpha_3\alpha_1 + \beta_3\beta_1 + \gamma_3\gamma_1 = 0$$

We find nine direction cosines connected by six relations. Thus, three remain unconnected. However, these three are not independent of each other and as such they cannot be taken as the remaining three generalized coordinates for the specification of the configuration of the rigid body.

Various sets of the remaining three generalized coordinates have been proposed. The most common and useful of them are the Eulerian angles. They refer to the angles corresponding to three successive rotations of the space-fixed system performed in a particular sequence or order, such that at the end, the axes of the space-fixed system coincide with those of the body-fixed system. Clearly, the Eulerian angles give the orientations of the axes of the body-fixed system relative to the space-fixed system.

In the following, we consider the three successive rotations of the space-fixed system to define Eulerian angles.

First Rotation

The space-fixed system (XYZ) is rotated about the Z-axis counter-clockwise by an angle ϕ, such that the X- and Y-axes, respectively, take the new positions

X' and Y' and the new Y–Z plane, namely the Y'–Z' plane contains the axis X_3 of the body-fixed system as shown in the Figure 2.2.
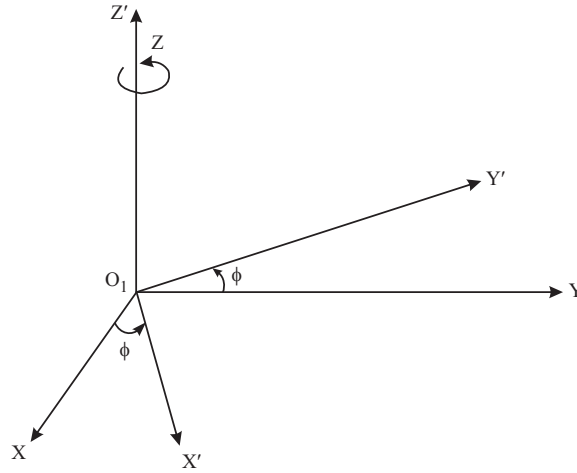
Fig. 2.2 *First Rotation*

Let $\hat{i}_1'', \hat{j}_1'$ and $\hat{k}_1'$ be, respectively, the unit vectors along the transformed

set of axes X', Y' and Z'.

We then have

$$\hat{i}_1'' = \hat{i}\cos\phi + \hat{j}\sin\phi$$

$$\hat{j}_1' = -\hat{i}\sin\phi + \hat{j}\cos\phi \qquad (2.96)$$

$$\hat{k}_1' = \hat{k}$$

The above equations can be written in the matrix form as

$$\begin{pmatrix} \hat{i}_1'' \\ \hat{j}_1' \\ \hat{k}_1' \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{i} \\ \hat{j} \\ \hat{k} \end{pmatrix} \qquad (2.97)$$

The matrix of transformation from XYZ to $X'\,Y'\,Z'$ is thus,

$$D = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad (2.98)$$

Second Rotation

The transformed system $X'Y'Z'$ is rotated about the X'-axis counter-clockwise by an angle θ, such that Z'-axis which is the same as the Z-axis coincides with the axis X_3 of the body-fixed system and the transformed $X''- Y''$ plane becomes the X_1–X_2 plane of the body-fixed system as shown in Figure 2.3.
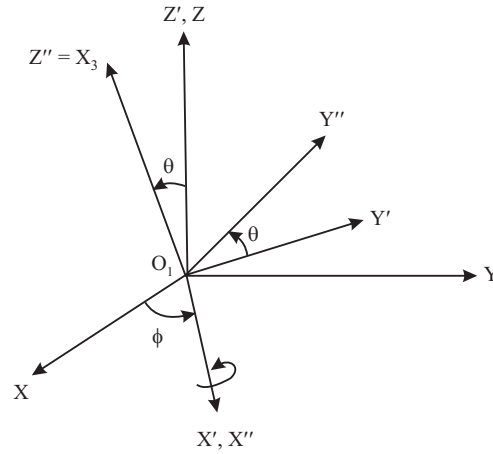
Fig. 2.3 *Second Rotation*

Let $\hat{i}_1''$, $\hat{j}_1''$, $\hat{k}_1''$ be, respectively, the unit vectors along the transformed

set of axes X", Y" and Z". We then have

$$\hat{i}_1'' = \hat{i}_1'$$

$$\hat{j}_1'' = \hat{j}_1' \cos\theta + \hat{k}_1' \sin\theta \qquad (2.99)$$

$$\hat{k}_1'' = -\hat{j}_1' \sin\theta + \hat{k}_1' \cos\theta$$

In matrix form, the above equations can be written as

$$\begin{pmatrix} \hat{i}_1'' \\ \hat{j}_1'' \\ \hat{k}_1'' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \end{pmatrix} \begin{pmatrix} \hat{i}_1' \\ \hat{j}_1' \\ \hat{k}_1' \end{pmatrix} \qquad (2.100)$$

The matrix of transformation from the $X'Y'Z'$ system to ($X''Y''Z''$) system is thus

$$C = \begin{pmatrix} 1 & 0 & 0 \\ \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \end{pmatrix} \qquad (2.101)$$

Third Rotation

The new system ($X''Y''Z''$) obtained after the second rotation is rotated about the Z'' ($= X_3$) axis counter-clockwise by an angle ψ such that the transformed axis X''' coincides with X_1-axis while the transformed axis Y''' coincides with X_2-axis of the body-fixed system as illustrated in Figure 2.4.

X', Z
Z'', Z''' = X₃
Y''' = X₂
Y''
θ
ψ
θ
Y'
O₁
φ
Y
φ
ψ
X
X', X''
X''' = X₁
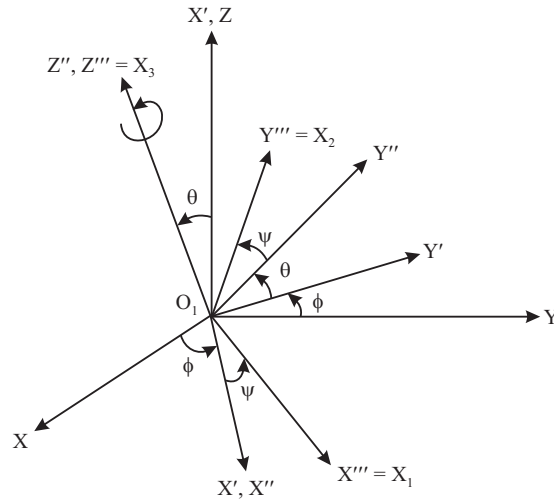
Fig. 2.4 *Third Rotation*

Let $\hat{i}_1'''$, $\hat{j}_1'''$, $\hat{k}_1'''$ be, respectively, the unit vectors along the transformed

axes X''', Y''' and Z'''. We get according to the operation performed

$$\hat{i}_1''' = \hat{i}_1'' = \hat{i}_1'' \cos\psi + \hat{j}_1'' \sin\psi$$

$$\hat{j}_1''' = \hat{j}_1''' = \hat{j}_1'' = -\hat{i}_1'' \sin\psi + \hat{j}_1'' \cos\psi \qquad (2.102)$$

$$\hat{k}_1''' = \hat{k}' = \hat{k}_1''$$

In matrix form we may write the above equations as

$$\begin{pmatrix} \hat{i}_1''' \\ \hat{j}_1''' \\ \hat{k}_1''' \end{pmatrix} = \begin{pmatrix} \hat{i}' \\ \hat{j}' \\ \hat{k}' \end{pmatrix} = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{i}_1'' \\ \hat{j}_1'' \\ \hat{k}_1'' \end{pmatrix} \qquad (2.103)$$

The matrix of transformation from $(X''Y''Z'')$ system to $X'''Y'''Z'''$ (or $X_1 X_2 X_3$) system is thus

$$B = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad (2.104)$$

The complete transformation from (XYZ) system to $(X_1 X_2 X_3)$ system is thus in matrix form given by

$$\begin{pmatrix} \hat{i}' \\ \hat{j}' \\ \hat{k}' \end{pmatrix} = B \begin{pmatrix} \hat{i}_1'' \\ \hat{j}_1'' \\ \hat{k}_1'' \end{pmatrix} = BC \begin{pmatrix} \hat{i}_1'' \\ \hat{j}_1'' \\ \hat{k}_1'' \end{pmatrix} = BCD \begin{pmatrix} \hat{i} \\ \hat{j} \\ \hat{k} \end{pmatrix}$$

or
$$\begin{pmatrix} \hat{i}' \\ \hat{j}' \\ \hat{k}' \end{pmatrix} = A \begin{pmatrix} \hat{i} \\ \hat{j} \\ \hat{k} \end{pmatrix} \qquad (2.105)$$

where $\qquad A = BCD \qquad (2.106)$

Using the matrices for B given by Equation (2.104), C given by Equation (2.101) and D given by Equation (2.98), we obtain

$$A = \begin{pmatrix} \cos\psi\cos\phi - \cos\theta\sin\phi\sin\psi & \cos\psi\sin\phi + \cos\theta\cos\phi\sin\psi & \sin\psi\sin\theta \\ -\sin\psi\cos\phi - \cos\theta\sin\phi\cos\psi & -\sin\psi\sin\phi + \cos\theta\cos\phi\cos\psi & \cos\psi\sin\theta \\ \sin\theta\sin\phi & -\sin\theta\cos\phi & \cos\theta \end{pmatrix} \quad (2.107)$$

Since all the elements of the matrix A are real, the matrix A itself is real.

Each of the matrices B, C and D corresponds to orthogonal transformation because the transformation of the axes is caused by simple rotations. We, thus, have

$$\widetilde{D} = D^{-1}, \ \widetilde{C} = C^{-1}, \ \widetilde{B} = B^{-1} \qquad (2.108)$$

We have the total transformation matrix $A = BCD$

Taking transpose of the above, we get

$$\widetilde{A} = \widetilde{BCD} = \widetilde{D}\widetilde{C}\widetilde{B} \qquad (2.109)$$

or
$$\widetilde{A} = D^{-1}C^{-1}B^{-1} = (BCD)^{-1} = A^{-1} \qquad (2.110)$$

Thus, A is also orthogonal.

Equations of Motion of a Rigid Body: Euler's Equations

There exist several methods to analyse the dynamics of a rigid body. One such method is due to Euler in which the analysis is made in terms of the body-fixed frame of reference or the moving coordinate system which rotates with the body. Simplification arises because relative to this frame, the moments of inertia and products of inertia are time-independent while relative to space-fixed system they are functions of time.

Euler's Equation for Force Free Motion

We know that a rigid body undergoes pure rotational motion about an axis passing through a fixed point in the body when a net external torque about that axis acts on the body. The external torque $\vec{\Gamma}$ and the angular momentum $\vec{J}$ of the body about the axis of rotation are related according to

$$\vec{\Gamma} = \frac{d\vec{J}}{dt} \qquad (2.111)$$

where the time derivative of $\vec{J}$ is calculated relative to the space-fixed

system external to the body.

If $\vec{\Omega}$ be the angular velocity of rotation of the body then the time

derivatives relative to space-fixed system of axes and body-fixed system of axes are given by

$$\left(\frac{d}{dt}\right)_{space} = \left(\frac{d}{dt}\right)_{body} + \vec{\Omega}\,X \tag{2.112}$$

In view of the relation expressed by Equation (2.112), we may write Equation (2.111) as

$$\vec{\Gamma} = \left(\frac{d\vec{J}}{dt}\right)_{body} + \vec{\Omega}\times\vec{J} \tag{2.113}$$

If, for convenience, we choose the axes of the body-fixed system as the principal axes of the body, we get

$$\vec{J} = \hat{i}I_1\Omega_1 + \hat{j}I_2\Omega_2 + \hat{k}I_3\Omega_3 \tag{2.114}$$

where Ω_1, Ω_2, Ω_3 are, respectively, the components of the angular velocity $\vec{\Omega}$ along the principal axes along which the unit vectors are $\hat{i}, \hat{j}, \hat{k}$, while I_1, I_2, I_3 are the principal moments of inertia.

Since, relative to the body-fixed system, the principal moments of inertia and the unit vectors are stationary or time-independent, we get from Equation (2.114)

$$\left(\frac{d\vec{J}}{dt}\right)_{body} = \hat{i}I_1\dot{\Omega}_1 + \hat{j}I_2\dot{\Omega}_2 + \hat{k}I_3\dot{\Omega}_3 \tag{2.115}$$

The component of $\vec{\Gamma}$ along the principal axis along which the unit vector is $\hat{i}$, is given by

$$\vec{\Gamma} = \hat{i}\cdot\vec{\Gamma} = \hat{i}\cdot\left[\left(\frac{d\vec{J}}{dt}\right)_{body} + \vec{\Omega}\times\vec{J}\right] \qquad \text{[using Equation (2.113)]}$$

$$= \hat{i}\cdot\left(\frac{d\vec{J}}{dt}\right)_{body} + \hat{i}\left(\vec{\Omega}\times\vec{J}\right)$$

$$= I_1\dot{\Omega}_1 + \hat{i} \cdot \left[\hat{i}\left(\Omega_2 j_3 - \Omega_3 j_2\right) + \hat{j}\left(\Omega_3 j_1 - \Omega_1 j_3\right) + \hat{k}\left(\Omega_1 j_2 - \Omega_2 j_1\right) \right]$$

or
$$\Gamma_1 = I_1\dot{\Omega}_1 + \left(\Omega_2 j_3 - \Omega_3 j_2\right)$$

or
$$\Gamma_1 = I_1\dot{\Omega}_1 + \Omega_2 I_3 \Omega_3 - \Omega_3 I_2 \Omega_2 \quad \text{[using Equation (2.114)]}$$

or
$$\Gamma_1 = I_1\dot{\Omega}_1 - \left(I_2 - I_3\right)\Omega_2\Omega_3 \tag{2.116}$$

Similarly, we obtain for the other two components of the torque $\vec{\Gamma}$ along the remaining principal axes as

$$\Gamma_2 = I_2\dot{\Omega}_2 - \left(I_3 - I_1\right)\Omega_3\Omega_1 \tag{2.117}$$

$$\Gamma_3 = I_3\dot{\Omega}_3 - \left(I_1 - I_2\right)\Omega_1\Omega_2 \tag{2.118}$$

Restricting our considerations to force-free motion of the rigid body, we get

(*i*) Potential energy of the body = 0

(*ii*) Kinetic energy of the body = Rotational kinetic energy (T_{rot})

so that the Lagrangian L of the body becomes

$$L = T_{\text{rot}}$$

Choosing the axes of the body-fixed or the moving system as the principal axes of the body, we get the Lagrangian as

$$L = \frac{1}{2}\left(I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2\right) \tag{2.119}$$

Further, for convenience, let us choose the generalized coordinates corresponding to the three rotational degrees of freedom as the Eulerian angles ψ, θ and ϕ. We may then write the Lagrangian as

$$L = L\left(\psi, \theta, \phi, \dot{\psi}, \dot{\theta}, \dot{\phi}\right) \tag{2.120}$$

The Lagrange's equation for the coordinate ψ is

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\psi}}\right) = \frac{\partial L}{\partial \psi}$$

or
$$\frac{d}{dt}\sum_{i=1}^{3}\frac{\partial L}{\partial \Omega_i}\frac{\partial \Omega_i}{\partial \dot{\psi}} - \sum_{i=1}^{3}\frac{\partial L}{\partial \Omega_i}\frac{\partial \Omega_i}{\partial \psi} = 0 \tag{2.121}$$

The components of the angular velocity $\vec{\Omega}$ can be expressed in terms of the Eulerian angles as

$$\Omega_1 = \dot{\phi}\sin\theta\sin\psi + \dot{\theta}\cos\psi$$

$$\Omega_2 = \dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi \qquad (2.122)$$

$$\Omega_3 = \dot{\phi}\cos\theta + \dot{\psi}$$

From the above we obtain

$$\frac{\partial\Omega_1}{\partial\psi} = \phi\sin\theta\cos\psi - \dot{\theta}\sin\psi = \Omega_2$$

$$\frac{\partial\Omega_2}{\partial\psi} = -\phi\sin\theta\sin\psi - \dot{\theta}\cos\psi = -\Omega_1 \qquad (2.123)$$

$$\frac{\partial\Omega_3}{\partial\psi} = 0$$

$$\frac{\partial\Omega_1}{\partial\dot{\psi}} = 0$$

and
$$\frac{\partial\Omega_2}{\partial\dot{\psi}} = 0 \qquad (2.124)$$

$$\frac{\partial\Omega_3}{\partial\dot{\psi}} = 1$$

From Equation (2.119), we get

$$\frac{\partial L}{\partial\Omega_1} = I_1\Omega_1$$

$$\frac{\partial L}{\partial\Omega_2} = I_2\Omega_2 \qquad (2.125)$$

$$\frac{\partial L}{\partial\Omega_3} = I_3\Omega_3$$

Euler's Equation in a Force Field

Since a rigid body, in general, has six degrees of freedom, its motion can be described in terms of six independent coordinates. Thus, the general equations of motion of a rigid body are six in number. In the Newtonian formulation, three of these are given by

$$\dot{\vec{p}} = \sum\dot{\vec{p}} = \sum\vec{f} = \vec{F}$$

where $\dot{\vec{p}}$ is the linear momentum of any particle, and $\vec{f}$ is the force acting on the particle.

The summation is carried over all the particles of the body. Here, the total force $\vec{F}$ includes, in principle, both external as well as internal forces.

We may start with the fundamental equations expressed as

$$\left(\frac{d\vec{P}}{dt}\right)_{fixed} = \vec{F} \text{ and } \left(\frac{d\vec{M}}{dt}\right)_{fixed} = \vec{k}$$

The designation 'fixed' is written explicitly since the above relations (from Newtonian mechanics) are valid only in an inertial frame of reference. The moving system X_1, X_2, X_3, is fixed in the rigid body and hence rotates with angular velocity $\vec{\Omega}$. If the radius vector of a point in the system changes

from $\vec{r}$ to $\vec{r} + \vec{\delta r}$ then the geometrical situation is correctly represented, if we write

$$\vec{\delta r} = \vec{\delta\theta} \times \vec{r}$$

where $\vec{\delta\theta}$ is a vector whose magnitude is equal to the infinitesimal rotation angle $\delta\theta$ and having direction along the instantaneous axis of rotation. Similarly, we can write

$$\vec{\delta v} = \vec{\delta\theta} \times \vec{v}$$

Let us consider an arbitrary vector $\vec{A}$. The change in this vector in time *dt* with respect to the fixed axis differs from the corresponding change with respect to the axis moving with the rigid body, only by the effects of the rotation of the body axes. In other words, we may write

$$\left(d\vec{A}\right)_{fixed} = \left(d\vec{A}\right)_{moving} + \left(d\vec{A}\right)_{rot}$$

$$\left(d\vec{A}\right)_{fixed} = \left(d\vec{A}\right)_{moving} + d\theta \times \vec{A}$$

The time rate of change of $\vec{A}$ is then

$$\left(\frac{d\vec{A}}{dt}\right)_{fixed} = \left(\frac{d\vec{A}}{dt}\right)_{moving} + \vec{\Omega} \times \vec{A}$$

The six equations of motion with respect to the body-fixed system are thus

$$\vec{F} = \left(\frac{d\vec{P}}{dt}\right)_{moving} + \vec{\Omega} \times \vec{P}$$

$$\vec{K} = \left(\frac{d\vec{M}}{dt}\right)_{moving} + \vec{\Omega} \times \vec{M}$$

Choosing the axes of the moving system to coincide with the principal axes of the body and taking $P = M\vec{V}$ and $M_i = I_i\Omega_i$, etc., we obtain

$$F_1 = M\left(\frac{dV_1}{dt} + \Omega_2 V_3 - \Omega_3 V_2\right)$$

$$F_2 = M\left(\frac{dV_2}{dt} + \Omega_3 V_1 - \Omega_1 V_3\right)$$

$$F_3 = M\left(\frac{dV_3}{dt} + \Omega_1 V_2 - \Omega_2 V_1\right)$$

$$K_1 = I_1\frac{d\Omega_1}{dt} \Omega_2\Omega_3\left(I_3 - I_2\right)$$

$$K_2 = I_2\frac{d\Omega_2}{dt} \Omega_3\Omega_1\left(I_1 - I_3\right)$$

$$K_3 = I_3\frac{d\Omega_3}{dt} \Omega_1\Omega_2\left(I_2 - I_1\right)$$

Substituting the above results [Equation (2.122) to (2.145)] in Equation (2.121), we obtain

$$\frac{d}{dt}\left(I_3\Omega_3\right) - \left[I_1\Omega_1\Omega_2 - I_2\Omega_2\Omega_1\right] = 0$$

or
$$\left(I_1 - I_2\right)\Omega_1\Omega_2 - I_3\Omega_3$$

or
$$I_3\dot{\Omega}_3 = \left(I_1 - I_2\right)\Omega_1\Omega_2 \qquad (2.126)$$

Similarly, we obtain the equations for $\dot{\Omega}_1$ and $\dot{\Omega}_2$ as

$$I_1\dot{\Omega}_1 = (I_2 - I_3)\Omega_2\Omega_3 \qquad (2.127)$$

$$I_2\dot{\Omega}_2 = (I_3 - I_1)\Omega_3\Omega_1 \qquad (2.128)$$

$$\dot{\vec{p}} = \sum \dot{\vec{p}} = \sum \vec{f} = \vec{F} \qquad (2.129)$$

$$\frac{\overrightarrow{dp}}{dt} = \vec{F} \text{ and } \left(\frac{d\vec{M}}{dt}\right)_{fixed} = \vec{K} \qquad (2.130)$$

$$\delta \vec{r} = \delta\vec{\theta} \times \vec{r} \qquad (2.131)$$

$$\delta \vec{v} = \delta\vec{\theta} \times \vec{v} \qquad (2.132)$$

$$\left(d\vec{A}\right)_{fixed} = \left(d\vec{A}\right)_{moving} + \left(d\vec{A}\right)_{rotational} \qquad (2.133)$$

$$\left(\frac{d\vec{A}}{dt}\right)_{fixed} = \left(\frac{d\vec{A}}{dt}\right)_{moving} + \vec{\Omega} \times \vec{A} \qquad (2.134)$$

$$\vec{F} = \left(\frac{d\vec{P}}{dt}\right)_{moving} + \vec{\Omega} \times \vec{P} \qquad (2.135)$$

$$\vec{K} = \left(\frac{d\vec{M}}{dt}\right)_{moving} + \vec{\Omega} \times \vec{M} \qquad (2.136)$$

Choosing the axes of the moving system to coincide with the principal axes of the body and taking $P_i = M\vec{V}_i$ and $M_i = I_i\Omega_i$, $M_i = I_i\Omega_i$, etc., we obtain

$$F_1 = M\left(\frac{dV_1}{dt} + \Omega_2 V_3 - \Omega_3 V_2\right)$$

$$F_2 = M\left(\frac{dV_2}{dt} + \Omega_3 V_1 - \Omega_1 V_3\right) \qquad (2.137)$$

$$F_3 = M\left(\frac{dV_3}{dt} + \Omega_1 V_2 - \Omega_2 V_1\right)$$

$$K_1 = I_1 \frac{d\Omega_1}{dt} \Omega_2 \Omega_3 (I_3 - I_2)$$

$$K_2 = I_2 \frac{d\Omega_2}{dt} \Omega_3 \Omega_1 (I_1 - I_3) \qquad (2.138)$$

$$K_3 = I_3 \frac{d\Omega_3}{dt} \Omega_1 \Omega_2 (I_2 - I_1)$$

Equations (2.137) and (2.138) are Euler's equations for the motion of a rigid body in a force field.

We may note that Equation (2.126) for Ω_3 is the Lagrange's equation for the coordinate ψ but Equation (2.127) and (2.138) are not Lagrange's equations for the coordinates θ and ϕ.

Torque Free Motion of a Rigid Body

Euler's equations obtained in the previous sections can be conveniently applied to describe the motion of the rigid body when no net force or no net torque acts on the body. We first consider the torque free motion.

Consider a rigid body rotating about an axis passing through the centre of mass of the body. Let us choose the centre of mass, which is a fixed point within the body, as the origin of the principal axes of the body. Considering no torque to be acting on the body, Euler's equations given by Equation (2.116), (2.117) and (2.118) reduce respectively to

$$I_1 \dot{\Omega}_1 = (I_2 - I_3) \, \Omega_2 \, \Omega_3 \qquad (2.139)$$

$$I_2 \dot{\Omega}_2 = (I_3 - I_1) \, \Omega_3 \, \Omega_1 \qquad (2.140)$$

$$I_3 \dot{\Omega}_3 = (I_1 - I_2) \, \Omega_1 \, \Omega_2 \qquad (2.141)$$

Multiplying Equation (2.139), (2.140) and (2.141) respectively by Ω_1, Ω_2 and Ω_3 and adding, we get

$$I_1 \dot{\Omega}_1 \Omega_1 + I_2 \dot{\Omega}_2 \, \Omega_2 + I_3 \dot{\Omega}_3 \, \Omega_3 = [I_2 - I_3 + I_3 - I_1 + I_1 - I_2] \, \Omega_1 \, \Omega_2 \, \Omega_3 = 0$$

or

$$\frac{d}{dt} \left(\frac{1}{2} I_1 \Omega_1^2 + \frac{1}{2} I_2 \Omega_2^2 + \frac{1}{2} I_3 \Omega_3^2 \right) = 0$$

or

$$\frac{1}{2} I_1 \Omega_1^2 + \frac{1}{2} I_2 \Omega_2^2 + \frac{1}{2} I_3 \Omega_3^2 = \text{a constant}$$

Thus,

$$T_{\text{rot}} = \frac{1}{2} \left(I_1 \Omega_1^2 + I_2 \Omega_2^2 + I_3 \Omega_3^2 \right) = \text{a constant} \qquad (2.142)$$

Equation (2.142) shows that the kinetic energy of rotation of the body is an integral of motion.

Since no torque acts on the body, the total angular momentum $\vec{J}$ of

the rotating body is another integral of motion $\left(\dfrac{d\vec{J}}{dt} = \text{torque} = 0 \text{ or } \vec{J} = \text{a constant} \right)$

. Thus, we have

$$\vec{J} = \hat{i}\, I_1\Omega_1 + \hat{j}\, I_2\Omega_2 + \hat{k}\, I_3\Omega_3 = \text{a constant} \qquad (2.143)$$

We have

$$\vec{\Omega} \cdot \vec{J} = \left[\hat{i}\Omega_1 + \hat{j}\Omega_2 + \hat{k}\Omega_3 \right] \cdot \left[\hat{i}I_1\Omega_1 + \hat{j}I_2\Omega_2 + \hat{k}\, I_3\Omega_3 \right]$$

$$= I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2 \qquad (2.144)$$

Combining Equation (2.142) and (2.144) we obtain

$$2T_{\text{rot}} = \vec{\Omega} \cdot \vec{J} = \text{a constant} \qquad (2.145)$$

(i) Inertia Ellipsoid

We may note that the motion of a rigid body depends on the structure of the body through the quantities (numbers) I_1, I_2 and I_3. Hence, any two bodies which have the same principal moments of inertia move in exactly the same manner although they may have different shapes. The simplest geometrical shape for a body having three given principal moments is that of a homogeneous ellipsoid. Hence, it often becomes convenient to describe the motion of a rigid body in terms of the motion of equivalent ellipsoid. Such a description of a rigid body was due to **Poinsot** which has the advantage of providing a geometrical description of the motion without trying to obtain a complete solution of the problem.

Poinsot's construction can be understood as explained below. The kinetic energy of the rotating rigid body relative to a coordinate system whose axes are the principal axes is given by

$$T_{\text{rot}} = \frac{1}{2}\vec{\Omega} \cdot \vec{J} = \frac{1}{2}(I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2)$$

We may write

$$2T = I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2 = I\Omega^2 \qquad (2.146)$$

where I is the moment of inertia of the body about the axis of rotation.

Let $\hat{n}$ be a unit vector in the direction of $\vec{\Omega}$, so that

$$\vec{\Omega} = \Omega\hat{n} \qquad (2.147)$$

Let the direction cosines of the axis of rotation be α, β and γ.

We then get

$$\hat{n} = \alpha i + \beta \hat{j} + \gamma \hat{k} \tag{2.148}$$

and we can write the moment of inertia of the body about this axis as

$$I = \alpha^2 I_{xx} + \beta^2 I_{yy} + \gamma^2 I_{zz} + 2I_{xy}\alpha\beta + 2I_{yz}\beta\gamma + 2I_{zx}\gamma\alpha$$

$$\tag{2.149}$$

Let us now define a vector $\vec{P}$ according to

$$\vec{P} = \frac{\hat{n}}{\sqrt{I}} \tag{2.150}$$

Using Equation (2.147) the above becomes

$$\vec{P} = \frac{\vec{\Omega}}{\Omega\sqrt{I}} = \frac{\vec{\Omega}}{\sqrt{I\Omega^2}}$$

In view of Equation (2.146) the above can be written as

$$\vec{P} = \frac{\vec{\Omega}}{\sqrt{2T}} = \frac{1}{\sqrt{2T}}(\hat{i}\Omega_1 + \hat{j}\Omega_2 + \hat{k}\Omega_3) \tag{2.151}$$

Further, we may write $\vec{P}$ in terms of its components P_1, P_2, P_3 as

$$\vec{P} = \hat{i}P_1 + \hat{j}P_2 + \hat{k}P_3 \tag{2.152}$$

Comparing Equation (2.151) and (2.152) we obtain

$$P_1 = \frac{\Omega_1}{\sqrt{2T}}, \qquad P_2 = \frac{\Omega_2}{\sqrt{2T}}, \qquad P_3 = \frac{\Omega_3}{\sqrt{2T}} \tag{2.153}$$

In view of Equation (2.153), Equation (2.146) gives

$$I_1 P_1^2 + I_2 P_2^2 + I_3 P_3^2 = 1 \tag{2.154}$$

Equation (2.154) is the equation of an ellipsoid and is called the equation of inertia ellipsoid.

(ii) Invariable Plane

Consider a rigid body rotating about a fixed point, say *O,* without the action of any external force or torque. The angular momentum vector $\vec{J}$ is a constant of motion and has a fixed direction in space as shown in Figure 2.5. The line along the fixed direction of $\vec{J}$ is called the invariable line. We have for force/torque free motion

$$\vec{\Omega}\cdot\vec{J} = 2T = \text{constant}$$

Clearly, the projection of $\vec{\Omega}$ along $\vec{J}$ is $\Omega \cos \theta$ which is constant and hence the tip of $\vec{\Omega}$ describes a plane called the invariable plane. To an observer fixed in the body-fixed coordinate system the angular velocity vector $\vec{\Omega}$ is found to precess about the angular momentum vector $\vec{J}$.
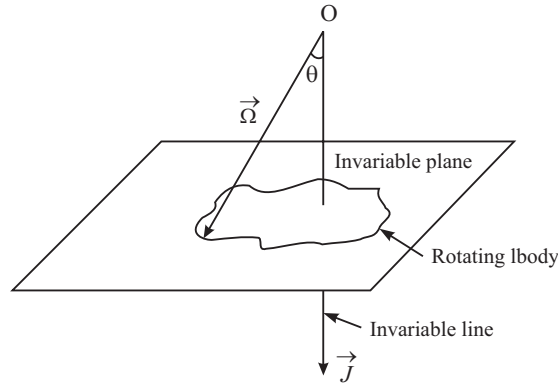
Fig. 2.5 *Fixed Direction*

For force-free motion of the rigid body we have

$$\vec{P}.\vec{J} \;=\; \frac{\vec{\Omega}.\vec{J}}{\sqrt{2T}} = \sqrt{2T} \;=\; \text{constant} \qquad\qquad (2.155)$$

The above shows that the tip of the vector $\vec{P}$ also describes an invariable plane. It can be seen that this invariable plane is the tangent plane at the point *P* of the inertia ellipsoid.

The distance between the origin of the ellipsoid and the tangent plane at the point *P* is

$$d = P \cos \theta = \frac{\vec{P}.\vec{J}}{J} = \frac{\vec{\Omega}.\vec{J}}{J\sqrt{2T}} = \frac{\sqrt{2T}}{J} \;=\; \text{constant} \qquad\qquad (2.156)$$

As a consequence we find that as the angular velocity vector $\vec{\Omega}$ and hence $\vec{P}$ changes with time, the inertia ellipsoid rolls on the invariable plane with the centre of the ellipsoid at a constant height above the plane.

The curve traced out on the invariable plane by the point of contact with the ellipsoid is called **herpolhode** and the corresponding curve described

on the ellipsoid is called **polhode**. We find that the polhode undergoes pure rolling on the herpolhode in the invariable plane (Refer to Figure 2.6).

In the case of a symmetrical rigid body rotating about the symmetry axis (z-axis) we have $I_1 = I_2$ and we find the inertia ellipsoid to be an ellipsoid of revolution. Vector $\vec{P}$ and hence vector $\vec{\Omega}$ remains constant in magnitude.

As a result, the polhode becomes a circle about the symmetry axis of the ellipsoid and herpolhode is a circle on the invariable plane. The angular velocity vector $\vec{\Omega}$ describes a cone called the **body cone**. As observed by

the observer in the space-fixed system $\vec{\Omega}$ moves also on the surface of a cone

called the space cone (Refer to Figure 2.7).
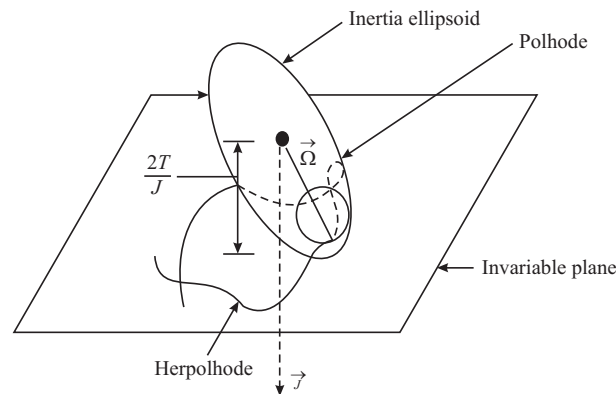
Fig. 2.6 *Invariable Plane*
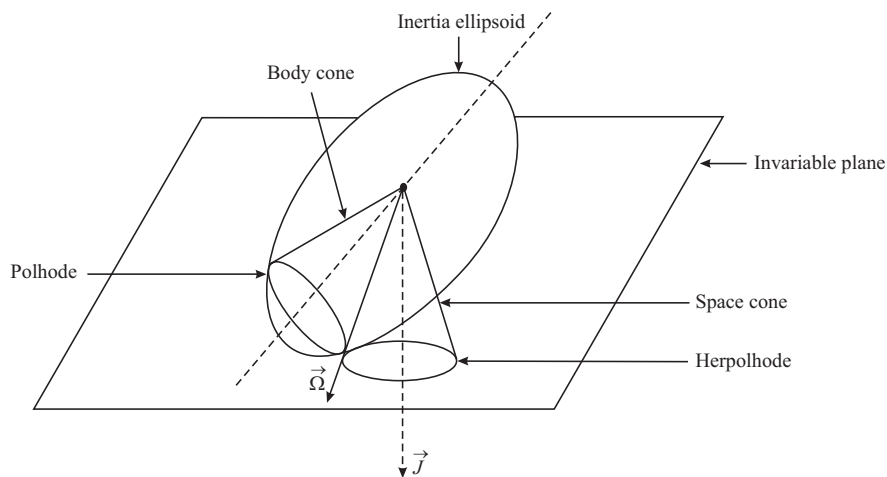
Fig. 2.7 *Space Cone*

Force Free Motion of a Symmetrical Rigid Body

As another application, we use in the following, Euler's equations to discuss force free motion of a symmetrical rigid body.

Choosing the symmetry axis as the principal z-axis we get $I_1 = I_2$ and Equation (2.116), (2.117) and (2.118) give

$$I_1 \dot{\Omega}_1 = (I_1 - I_3) \, \Omega_2 \, \Omega_3 \tag{2.157}$$

$$I_1 \dot{\Omega}_2 = (I_3 - I_1) \, \Omega_3 \, \Omega_1 = - (I_1 - I_3) \, \Omega_3 \, \Omega_1 \tag{2.158}$$

$$I_3 \dot{\Omega}_3 = 0 \tag{2.159}$$

Equation (2.158) yields

$$\Omega_3 = \text{constant} \tag{2.160}$$

i.e., the component of angular velocity along the symmetry axis is a constant.

Putting
$$\Omega_0 = \frac{I_1 - I_3}{I_1} \Omega_3 \text{ (which is a constant)} \tag{2.161}$$

we may write Equation (2.157) and (2.158) as

$$\dot{\Omega}_1 = \Omega_0 \, \Omega_2 \tag{2.162}$$

$$\dot{\Omega}_2 = - \Omega_0 \, \Omega_1 \tag{2.163}$$

Differentiating Equation (2.162) with respect to time we get

$$\ddot{\Omega}_1 = \Omega_0 \, \dot{\Omega}_2$$

Substituting Equation (2.163) in the above we obtain

$$\ddot{\Omega}_1 = - \Omega_0^2 \, \Omega_1$$

or
$$\ddot{\Omega}_1 + \Omega_0^2 \, \Omega_1 = 0 \tag{2.164}$$

Solution of Equation (2.164) can be put in the form

$$\Omega_1 = A \sin \Omega_0 t \tag{2.165}$$

where A is some constant. We have chosen the phase constant such that at $t = 0$, $\Omega_1 = 0$.

From Equation (2.165) we get

$$\dot{\Omega}_1 = A\Omega_0 \cos \Omega t \tag{2.166}$$

Using the above in Equation (2.162) we obtain

$$A\Omega_0 \cos\Omega_0 t = \Omega_0 \Omega_2$$

or $$\Omega_2 = A \cos \Omega_0 t \qquad (2.167)$$

Combining Equation (2.165) and (2.167) we obtain

$$\Omega_1^2 + \Omega_2^2 = A^2 \qquad (2.168)$$

which is the equation of a circle of radius A. Let us now consider the vector $\vec{\Omega}_P$ in the x-y plane as

$$\vec{\Omega}_P = \hat{i}\Omega_1 + \hat{j}\Omega_2 = \hat{i}\, A\sin\Omega t + \hat{j}\, A\cos\Omega t$$

The above gives

$$\vec{\Omega}_P \,| = A = \text{constant} \qquad (2.169)$$

We further find that the vector $\vec{\Omega}_P$ rotates about the symmetry axis with

constant angular frequency Ω given by Equation (2.61) as shown in Figure 2.8.

The angular velocity of the body given by

$$\vec{\Omega} = \hat{i}\Omega_1 + \hat{j}\Omega_2 + \hat{k}\Omega_3$$

can thus be written as

$$\vec{\Omega} = \vec{\Omega}_P + \Omega_3\hat{k} \qquad (2.170)$$

with $$|\vec{\Omega}| = \sqrt{\Omega_P^2 + \Omega_3^2} = \text{constant} \qquad (2.171)$$

We find the angular velocity vector $\vec{\Omega}$ to have a constant magnitude

and precessing about the axis of symmetry with the constant angular frequency Ω_0 (Refer to Figure 2.9). We find the vector $\vec{\Omega}$ to move on the

surface of a cone about the axis of symmetry with constant angular frequency Ω_0. This motion takes place with respect to the principal axes of the body which are themselves rotating in space with angular frequency Ω. Equation (2.161) shows that closer the values of I_1 and I_3, lower becomes the precessional frequency Ω_0 compared to rotational frequency Ω_3. We may determine Ω_P

and Ω_3 from a knowledge of the constant magnitudes of kinetic energy T and the angular momentum J given as

$$T = \frac{1}{2}I_1\Omega_P^2 + \frac{1}{2}I_3\Omega_3^2 \tag{2.172}$$

$$J^2 = I_1^2\Omega_P^2 + I_3^2\Omega_3^2 \tag{2.173}$$

The above results can be applied to the problem of rotation of the earth.
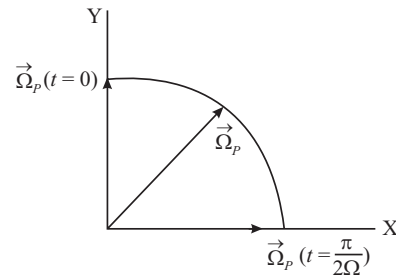
Fig. 2.8 Symmetric Axis with Constant Frequencies
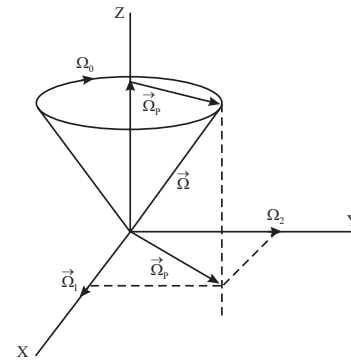
Fig. 2.9 Symmetric Axis with Angular Frequencies

We know that the earth is almost symmetric about the north-south (polar) axis and slightly bulged at the equator. As a consequence we have I_1 slightly less than I_3. On calculation we obtain

$$\frac{I_3 - I_1}{I_1} = \frac{1}{306}$$

and

$$\Omega_3 = \frac{2\pi}{24 \times 60 \times 60} \text{ rad s}^{-1}$$

The time period of precession of the axis of rotation of the earth is thus

$$T = \frac{2\pi}{\Omega_0} = \frac{2\pi}{\Omega_3}\frac{I_1}{I_3 - I_1} = \frac{2\pi}{\Omega_3} \times 306$$

or

$$T = 306 \text{ days}$$

Thus, an observer on the earth should find the axis of rotation of the earth to trace out a circle about the north pole every 306 days which agrees well with observation.

Motion of Symmetric Top Under the Action of Gravity

Consider the motion of a **symmetric top** spinning about the axis of symmetry namely the Z′-axis of the body-fixed system. Z′-axis is taken to be one of the principal axes, the other two principal axes being X′ and Y′ axes.

According to the above consideration; the principal moments of inertia about the X′ and the Y′ axes, namely I_1 and I_2 are equal.

Let the top have its pivot at its lower tip O which is the common origin of the body fixed and the space-fixed coordinate system X′Y′Z′ and XYZ respectively.

Let G be the centre of gravity of the top at a distance *l* from the point O.
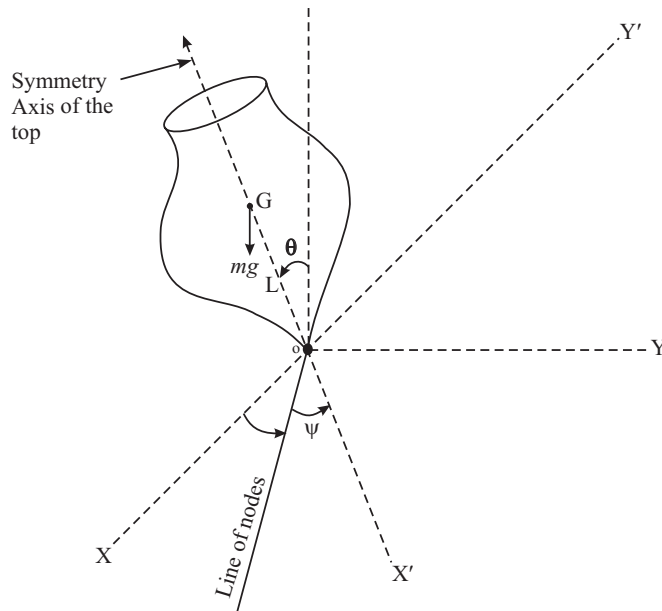
Fig. 2.10 *Symmetric Axis of Top under the Action of Gravity*

The only force that acts on the top of mass *m* is *mg* which acts vertically downwards from the point G. Let us consider the Z-axis of the space-fixed system to pointing vertically upwards while the X and Y axes to lie in the horizontal plane (Refer to Figure 2.10).

The most convenient generalized coordinates in terms of which the motion of the top can be described are the Euler angles ϕ, θ and ψ as shown in the Figure 2.10.

The Lagrangian function for the top under consideration is given by

$$L = T - V = \frac{1}{2}I_1(\Omega_1^2 + \Omega_2^2) + \frac{1}{2}I_3\Omega_3^2 - mgl\cos\theta \qquad (2.174)$$

Substituting for Ω_1, Ω_2 and Ω_3 in terms of Euler angles, the above becomes

$$L = \frac{1}{2}I_1\left[\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta\right] + \frac{1}{2}I_3\left[\dot{\psi} + \dot{\phi}\cos\theta\right]^2 - mgl\cos\theta \qquad (2.175)$$

The above expression for *L* shows that ψ and ϕ are cyclic coordinates. As a consequence the momenta conjugate to these coordinates, namely p_y and p_f are constants of motion or the first integral of motion.

We have

$$p_y = \frac{\partial L}{\partial \dot{\psi}} = I_3[\dot{\psi} + \dot{\phi}\cos\theta] = I_3\Omega_3 = I_1 a \qquad (2.176)$$

and $\qquad p_\text{f} = \dfrac{\partial L}{\partial \dot{\phi}} = [I_1 \sin^2 \theta + I_3 \cos^2 \theta]\, \dot{\phi} + I_3 \cos \theta\, \dot{\psi} = I_1 b$ $\qquad\qquad$ (2.177)

In the above the integrals of motion have been expressed in terms of new constants *a* and *b*.

A third integral of motion is the total energy of the top given by

$$E = T + V = \frac{1}{2} I_1 [\dot{\theta}^2 + \dot{\phi}^2 \sin^2 \theta] + \frac{1}{2} I_3 \Omega_3^2 + mgl \cos \theta \qquad (2.178)$$

Solving for $\dot{\phi}$ and $\dot{\psi}$, the equations (2.176) and (2.177) yield

$$\dot{\phi} = \frac{b - a \cos \theta}{\sin^2 \theta} \qquad (2.179)$$

and $$\dot{\psi} = \frac{I_1 a}{I_3} - \left[\frac{b - a \cos \theta}{\sin^2 \theta}\right] \cos \theta \qquad (2.180)$$

Using Equation (2.179) and (2.180) in Equation (2.178) we get

$$E = \frac{1}{2} I_1^2 a^2 + \frac{1}{2} I_1 \dot{\theta}^2 + \frac{1}{2} I_1 \frac{(b - a \cos \theta)^2}{\sin^2 \theta} + mgl \cos \theta \qquad (2.181)$$

For convenience we introduce a new quantity *E′* as

$$E' = E - \frac{1}{2} I_1^2 a^2 \qquad (2.182)$$

Using Equation (2.181) in Equation (2.182) we obtain

$$E' = \frac{1}{2} I_1 \dot{\theta}^2 + \frac{1}{2} I_1 \frac{(b - a \cos \theta)^2}{\sin^2 \theta} + mgl \cos \theta \qquad (2.183)$$

We may consider *E′* to be the sum of the kinetic energy $\frac{1}{2} I_1 \dot{\theta}^2$ and an effective potential energy function $V(\theta)$ defined as

$$V(\theta) = \frac{1}{2} I_1 \frac{(b - a \cos \theta)^2}{\sin^2 \theta} + mgl \cos \theta \qquad (2.184)$$

In view of Equation (2.184) we may write Equation (2.183) as

$$E' = \frac{1}{2} I_1 \dot{\theta}^2 + V(\theta) \qquad (2.185)$$

The above gives

$$\dot{\theta} = \left[\frac{2}{I_1} \{E' - V(\theta)\}\right]^{1/2}$$

or
$$dt = \frac{d\theta}{\left[\frac{2}{I_1}\{E' - V(\theta)\}\right]^{1/2}}$$

Integrating the above between time limits $t = 0$ to $t = t$, we get

$$\int_0^t dt = \int_{\theta(0)}^{\theta(t)} \frac{d\theta}{\left[\frac{2}{I_1}\{E' - V(\theta)\}\right]^{1/2}} \qquad (2.186)$$

The solution of the above equation to get θ and hence solve for ϕ and ψ as functions of t is however complicated and involves elliptic integrals.

For this reason, only the qualitative features of the motion of top from energy considerations are presented as in motion under central force.

1. Steady Precession of the Top

The variation of the effective potential $V(\theta)$ with θ is as shown in Figure 2.11. We find $V(\theta)$ to assume infinite values for $\theta = 0$ and $\theta = \pi$ and and for a particular value, namely $\theta = \theta_0$, $V(\theta)$ assumes the minimum value. Thus the physically acceptable value of θ lies between 0 and π. The minimum value of $V(\theta)$ correspondes to the condition $\frac{\partial V(\theta)}{\partial \theta} = 0$. Using the expression for $V(\theta)$ given by Equation (2.184), the above condition yields

$$I_1 a \frac{b - a\cos\theta_0}{\sin\theta_0} - I_1 \frac{(b - a\cos\theta_0)^2}{\sin^3\theta_0} - mgl\sin\theta_0 = 0$$

Solving the above we obtain

$$b - a\cos\theta_0 = \frac{a\sin^2\theta_0}{2\cos\theta_0}\left[1 \pm \sqrt{1 - \frac{4mgl\cos\theta_0}{I_1 a^2}}\right] \qquad (2.187)$$

$b - a\cos\theta_0$ is a real quantity. Thus, for $\theta_0 < n/2$, we get

$$\sqrt{1 - \frac{4mgl\cos\theta_0}{I_1 a^2}} \geq 0$$

or
$$I_1^2 a^2 \geq 4mglI_1\cos\theta_0$$

We have
$$I_1 a = I_3\Omega_3.$$

$$I_3^2\Omega_3^2 \geq 4mgl I_1\cos\theta_0$$

or
$$\Omega_3 \geq \frac{2}{I_3}\sqrt{mgl I_1\cos\theta_0} \qquad (2.188)$$
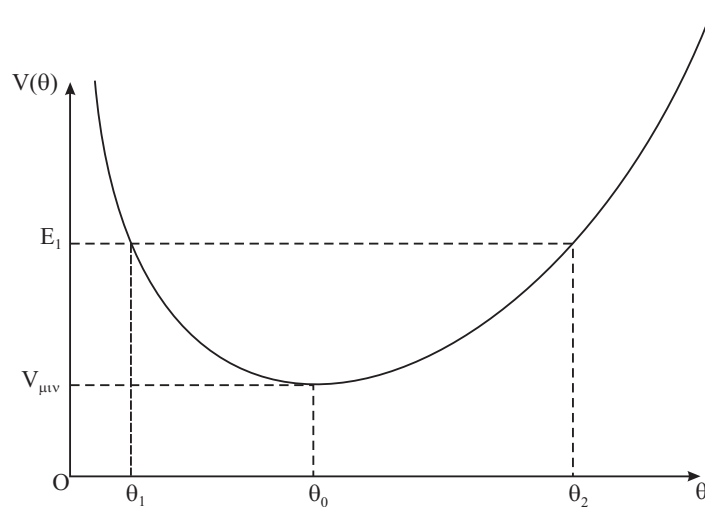
Fig. 2.11 *Steady Precession*

If the energy of the top is such that $E' = V_{minimum}$ then θ has only one value namely θ = $θ_0$. Thus corresponding to this energy the precession angle θ remain constants. This is referred to as **steady precession** in which case the symmetry axis of the top describes a right circular cone about the vertical Z-axis of the space-fixed system. The condition that must be satisfied for steady precession is that the value of $Ω_3$ be given according to Equation (2.188).

2. Nutational Motion of the Top

Referring to the effective potential energy diagram we find that if the energy of the top be such that $E' > V_{minimum}$, say $E = E_1'$, then the motion gets restricted between two values of θ, namely θ = $θ_1$ and θ = $θ_2$. This means that the symmetry axes OZ' of the rotating top varies its inclination θ with the vertical such that $θ_1 ≤ θ ≤ θ_2$. This kind of motion of the top is referred to as **nutation**.

We have

$$\dot{\phi} = \frac{b - a\cos θ}{\sin^2 θ}$$

Depending upon the values of $θ_1$ and $θ_2$, $\dot{\phi}$ may or may not change sign. If $\dot{\phi}$ does not change sign, the top precesses constantly about the vertical axis Z and the axis of rotation of the top, i.e., the Z' axis oscillates between θ = $θ_1$ and θ = $θ_2$. This is nutational motion which is an up and down motion of the top.

When the rotating top is released at the angle $θ_1$, it falls slightly due to the gravitational torque. As a consequence, it gains in the precessional as well as the nutational motion. The axis of the top reaches the maximum angle $θ_2$ with the vertical with the result of maximum precessional velocity and zero

nutational velocity. The precessional velocity at $\theta = \theta_2$ can be obtained from the Equation (2.179) having the value

$$\dot{\phi} = \frac{I_3 \Omega_3}{I_1} a = \frac{2mgl}{I_3 \Omega_3}$$

The following are a few important results:

1. Frequency of nutational motion is given by

$$\Omega = \frac{I_3 \Omega_3}{I_1} = a$$

2. Amplitude of nutation is given by

$$\theta_m = \frac{mgl \sin \theta_1}{I_3^2 \Omega_3^2}$$

This shows that nutation is less if the top spins fast.

Illustrative Examples

Example 2.1: If the moment of inertia of a cube about an axis that passes through the centre of mass and the centre of any one face is I_0, find the moment of inertia of the cube about an axis passing through the centre of mass and one corner of the cube.

Solution:
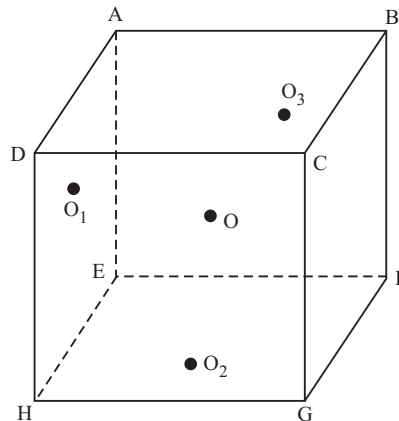
Fig. 2.12 Cube

As shown in the figure, let O be the centre of mass of the cube ABCDEFGH. Considering the point O as the origin, let us choose a Cartesian coordinate system with axes X, Y and Z passing through the centres O_1, O_2 and O_3 of the three adjacent faces ADHF, DCGH and ABCD, respectively. We then have, according to the problem

$$I_{xx} = I_{yy} = I_{zz} = I_0 \qquad (i)$$

and $\qquad\qquad I_{xy} = I_{yz} = I_{zx} = 0 \qquad (ii)$

Let an axis passing through the centre of mass of the cube and one corner, say C, have direction cosines α, β and γ, then the moment of inertia of the cube about this axis is given by

$$I = \alpha^2 I_{xx} + \beta^2 I_{yy} + \gamma^2 I_{zz} - 2\alpha\beta I_{xy} - 2\beta\gamma I_{yz} - 2\gamma\alpha I_{zx}$$

Using Equation (i) and (ii) in the above, we get

$$I = (\alpha^2 + \beta^2 + \gamma^2)I_0 \tag{iii}$$

The position vector of the corner C with respect to the origin O is given by

$$\vec{r} = \hat{i}a + \hat{j}a + \hat{k}a$$

Clearly, $\qquad |\vec{r}| = \sqrt{3}a$

We, thus, have

$$\alpha = \frac{a}{\sqrt{3}a} = \frac{1}{\sqrt{3}}, \quad \beta = \frac{a}{\sqrt{3}a} = \frac{1}{\sqrt{3}}, \quad \gamma = \frac{a}{\sqrt{3}a} = \frac{1}{\sqrt{3}} \tag{iv}$$

Substituting Equation (*iv*) in Equation (*iii*), we obtain

$$I = \left(\frac{1}{3} + \frac{1}{3} + \frac{1}{3}\right)I_o = I_o$$

Example 2.2: A thin uniform circular disc of mass m and radius *r* lies in the X–Y plane (plane of the paper). A point mass $\dfrac{5m}{4}$ is attached to the rim of the disc as shown in the figure. The moment of inertia of the disc about an axis through its centre of mass and perpendicular to the plane of the paper is

$$I = \frac{mr^2}{4}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Find the moment of inertia tensor of the system of the disc and point mass about the point A in the coordinate system shown.
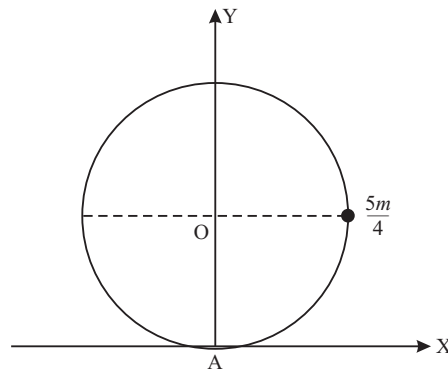
Fig. 2.13

Solution: The moment of inertia of the disc about the axes through the **centre of mass** O and perpendicular to the plane of the disc is given to be

$$I_{cm} = \frac{Mr^2}{4} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \qquad\qquad (i)$$

The mass point $\frac{5m}{4}$ attached to the disc contributes to the moments

and products of inertia about the origin A is

$$I_{ij} = \frac{5m}{4}\left(r_o^2 \delta_{ij} - x_i x_j\right) \qquad\qquad (ii)$$

where $\vec{r_0} = (x_1, x_2, x_3)$ is the radius vector of the mass point with respect to

the origin A

In the above

$$i, j = 1, 2, 3, \text{ and}$$
$$i, j = 1, 2, 3 \text{ and}$$
$$\delta_{ij} = 1 \text{ if } i = j$$
$$\delta_{ij} = 0 \text{ if } i \neq j \qquad\qquad (iii)$$

$$x_1 = x, x_2 = r \text{ and } x_3 = 0$$
$$r_0^2 = x_1^2 + x_2^2 + x_3^2 = 2r_0^2$$

We, thus, get the moment of inertia tensor of the mass point about the
point A as

$$\frac{I_{5m}}{4} = \frac{5m}{4}\begin{pmatrix} 2r^2 - r^2 & -r^2 & 0 \\ -r^2 & 2r^2 - r^2 & 0 \\ 0 & 0 & 2r^2 \end{pmatrix}$$

or

$$\frac{I_{5m}}{4} = \frac{5mr^2}{4}\begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \qquad\qquad (iv)$$

$$\delta_{ij} = 1 \text{ if } i = j \qquad\qquad (iii)$$
$$\delta_{ij} = 0 \text{ if } i \neq j$$
$$i, j = 1, 2, 3 \qquad\qquad (iv)$$

Using Equation (ii), (iii) and (iv), we obtain

$$I_{11} = m\left[\left(r_1^2 - x_{11}^2\right) + \left(r_2^2 - x_{21}^2\right) + \left(r_3^2 - x_{31}^2\right) + \left(r_4^2 - x_{41}^2\right)\right]$$

or

$$I_{11} = m[(a^2 - a^2) + (a^2 - a^2) + (4a^2 - 0) + (4a^2 - 0)]$$
$$I_{11} = 8ma^2$$

Proceeding as above, we obtain

$$I_{12} = 0 = I_{13}$$
$$I_{21} = 0 = I_{23}$$
$$I_{31} = 0 = I_{32}$$

and
$$I_{22} = 2ma^2$$
$$I_{33} = 10ma^2$$

Thus, the matrix of the inertial tensor is

$$I = \begin{pmatrix} 8ma^2 & 0 & 0 \\ 0 & 2ma^2 & 0 \\ 0 & 0 & 10ma^2 \end{pmatrix}$$

Example 2.3. In Example 2.2 find the moment of inertia of the system about an axis which is inclined equally to the positive *X*-, *Y*- and *Z*-axes.

Solution: Since the axis is equally inclined to the positive *X*-, *Y*- and *Z*-axes, its direction cosines α, β and γ are equal.

The moment of inertia of the system about the axis is given in terms of the elements of the inertial tensor $I_{ij}(i, j = 1, 2, 3)$ as

$$I = \alpha^2 I_{11} + \beta^2 I_{22} + \gamma^2 I_{33} - 2\alpha\beta I_{12} - 2\beta\gamma I_{23} - 2\gamma_2 I_{31}$$

Using the values of I_{ij} obtained in Example 2.2 and taking

$$\alpha = \beta = \gamma$$

We get, from the above equation,

$$I = \alpha^2 8ma^2 + \alpha^2 2ma^2 + \alpha^2 10ma^2 = 20ma^2\alpha^2$$

We further have

$$\alpha^2 + \beta^2 + \gamma^2 = 1$$
$$3\alpha^2 = 1$$

or
$$\alpha^2 = \frac{1}{3}$$

Substituting for α^2, we obtain

$$I = 20ma^2 \times \frac{1}{3} = \frac{20}{3}ma^2$$

Example 2.4: Four point masses, each equal to m, are situated in the *x–y* plane at positions $(a, 0, 0)$, $(-a, 0, 0)$, $(0, +2a, 0)$ and $(0, -2a, 0)$. The masses are joined by massless rods to form a rigid body.

Find the inertial tensor and express the tensor as a matrix using *X*-, *Y*- and *Z*-axes as the reference system.

Solution: The system to be considered is as shown in the figure. Let *O* be the origin of the reference system. The particles of mass *m* are situated at the points A(a, 0), B($-a$, 0), C(0, 2a) and D(0, $-2a$).
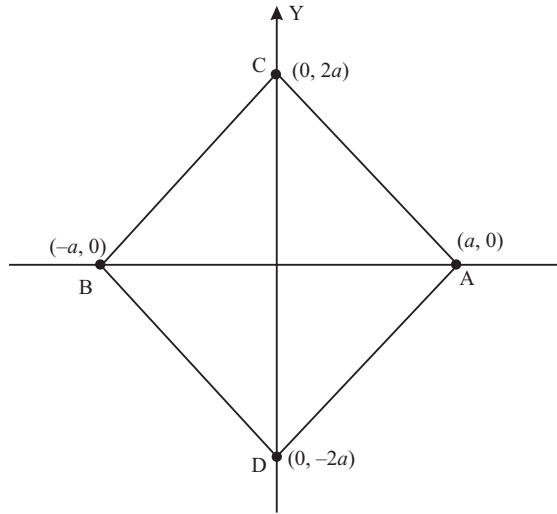
Fig. 2.14

By definition the elements I_{ij} of the inertial tensor are given by

$$I_{ij} = \sum_{k=1}^{4} m\left(r_k^2 \delta_{ij} - xk_i xk_j\right)$$

or $\quad I_{ij} = m\left[\left(r_1^2 \delta_{ij} - x_{1i} x_{1j}\right) + \left(r_2^2 \delta_{ij} - x_{2i} x_{2j}\right) + \left(r_3^2 \delta_{ij} - x_{3i} x_{3j}\right) + \left(r_4^2 \delta_{ij} - x_{4i} x_{4j}\right)\right]$ $\quad$ (i)

In the above,

$$r_1^2 = x_{11}^2 + x_{12}^2 + x_{13}^2 = a^2 + 0 + 0 = a^2$$

$$r_2^2 = x_{21}^2 + x_{22}^2 + x_{23}^2 = (-a)^2 + 0 + 0 = a^2$$

$$r_3^2 = x_{31}^2 + x_{32}^2 + x_{33}^2 = 0 + (2a)^2 + 0 = 4a^2 \qquad (ii)$$

$$r_4^2 = x_{41}^2 + x_{42}^2 + x_{43}^2 = 0 + (-2a)^2 + 0 = 4a^2$$

Example 2.5: Moments of inertia and products of inertia of a rigid body with respect to a coordinate system *XYZ* having the origin at some point within the body are I_{xx}, I_{yy}, I_{zz}, I_{xy}, I_{yz} and I_{zx}. Obtain the moment of inertia of the body about an axis inclined at angles α, β and γ with the *X*-, *Y*- and *Z*-axes, respectively.

Solution: The figure shows the origin *O* within the rigid body and the coordinate system *X Y Z*. Let *AB* be the axis about which the moment of inertia of the body is to be found.
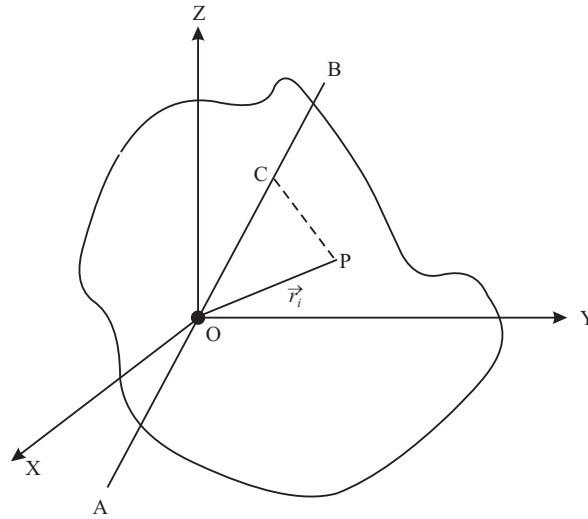
Fig. 2.15

Let the axis *AB* make angles of α, β and γ with the *X*-, *Y*- and *Z*-axes, respectively. If $\hat{n}$ be a unit vector along $\overrightarrow{OB}$, we get

$$\hat{n} = \hat{i}\cos\alpha + \hat{j}\cos\beta + \hat{k}\cos\gamma \qquad (i)$$

Let *P* be a particle of the body having mass m_i and radius vector $\vec{r_i}$.

Let *PC* be the perpendicular from *P* on the axis *AB*.

By definition, the moment of inertia of the particle about the axis *AB* $= m_i(PC)^2$.

Considering all the particles constituting the body, we get the moment of inertia of the body about the axis *AB* as

$$I = \sum m_i (PC)^2 = \sum m_i \left| \vec{r_i} \times \hat{n} \right|^2 \qquad (ii)$$

We have

$$\vec{r_i} \times \hat{n} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ x_i & y_i & z_i \\ \cos\alpha & \cos\beta & \cos\gamma \end{vmatrix}$$

$$= \hat{i}\left(y_i \cos\gamma - z_i \cos\beta\right) + \hat{j}\left(z_i \cos\alpha - x_i \cos\gamma\right) + \hat{k}\left(x_i \cos\beta - y_i \cos\alpha\right)$$

The above gives

$$\left| \vec{r_i} \times \hat{n} \right|^2 = \left(y_i \cos\gamma - z_i \cos\beta\right)^2 + \left(z_i \cos\alpha - x_i \cos\gamma\right)^2 + \left(x_i \cos\beta - y_i \cos\alpha\right)^2$$

$$= \left(y_i^2 + z_i^2\right)\cos^2\alpha + \left(x_i^2 + z_i^2\right)\cos^2\beta + \left(y_i^2 + x_i^2\right)\cos^2\gamma \qquad (iii)$$

$$-2y_i z_i \cos\gamma\cos\beta - 2z_i x_i \cos\alpha\cos\gamma - 2x_i y_i \cos\beta\cos\alpha$$

Using Equation (*iii*) in (*ii*), we obtain

$$I = \sum m_i \left(y_i^2 + z_i^2\right)\cos^2\alpha + \sum m_i \left(x_i^2 + z_i^2\right)\cos^2\beta + \sum m_i \left(y_i^2 + x_i^2\right)\cos^2\gamma$$

$$-2\sum m_i x_i y_i \cos\alpha\cos\beta - 2\sum m_i y_i z_i \cos\beta\cos\gamma - 2\sum m_i z_i x_i \cos\gamma\cos\alpha$$

Using the definitions of moments of inertia and products of inertia, the above can be expressed as

$$I = I_{xx}\cos^2\alpha + I_{yy}\cos^2\beta + I_{zz}\cos^2\gamma + 2I_{xy}\cos\alpha\cos\beta + 2I_{yz}\cos\beta\cos\gamma + 2I_{zx}\cos\gamma\cos\alpha$$

Using the theorem of parallel axes, the moment of inertia tensor disc about the axis through A (parallel to the axis through 0) is

$$I_A = I_{CM} + Mr^2$$

$$= \frac{mr^2}{4}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} + Mr^2\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad (iv)$$

$$= \frac{mr^2}{4}\left[\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} + \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}\right]$$

or $\qquad I_A = \dfrac{mr^2}{4}\begin{pmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 6 \end{pmatrix} \qquad (v)$

In view of Equation (*iv*) and (*v*) the moment of inertia tensor of the disc and the mass point about the axis through *A* is

$$I = I_A = \frac{I_{5m}}{4}$$

$$= \frac{mr^2}{4}\begin{pmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 6 \end{pmatrix} + \frac{5mr^2}{4}\begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

$$= \frac{mr^2}{4}\left[\begin{pmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 6 \end{pmatrix} + \begin{pmatrix} 5 & -5 & 0 \\ -5 & 5 & 0 \\ 0 & 0 & 10 \end{pmatrix}\right]$$

or
$$I = \frac{mr^2}{4}\begin{pmatrix} 10 & -5 & 0 \\ -5 & 10 & 0 \\ 0 & 0 & 16 \end{pmatrix}$$

(*b*) The frequency of the circular motion is

$$\omega = \sqrt{\frac{k}{mr}} \qquad\qquad (vi)$$

$$U\left(q_1, q_2, ..., q_s\right) = U\left(q_{1o}, q_{2o}, ..., q_{s,o}\right) + \sum_k \left(\frac{\partial u}{\partial q_k}\right)_o q_k + \frac{1}{2!}\sum_{j,k}\left(\frac{\partial^2 u}{\partial q_j \partial q_k}\right)_o q_j q_k + ...$$

$$(vii)$$

In the above expansion, we assume that $q_{0k} = 0$ and q_k is the displacement from $q_{0k} = 0$.

Using Equation (vi) in Equation (vii), we obtain

$$U\left(q_1, q_2, ..., q_s\right) = U\left(q_{1o}, q_{2o}, ..., q_{s,o}\right) + \frac{1}{2}\sum_{j,k}\left(\frac{\partial^2 u}{\partial q_j \partial q_k}\right)_o q_j q_k \qquad (viii)$$

Without loss of generality, we can set

$$U(q_{10}, q_{20}, ..., q_{s'0}) = 0 \qquad\qquad (ix)$$

Hence, we get from Equation (*viii*)

$$U(q_1, q_2, ..., q_s) = \frac{1}{2}\sum_{j,k}\left(\frac{\partial^2 u}{\partial q_j \partial q_k}\right)_o q_j q_k \qquad\qquad (x)$$

or
$$U(q_1, q_2, ..., q_s) = \frac{1}{2}\sum_{j,k} U_{jk} q_j q_k$$

where
$$U_{jk} = \left(\frac{\partial^2 V}{\partial q_j \partial q_k}\right)_o$$

is constant depending upon the equilibrium values of $q_k's$, i.e., on $q_{ok}'s$.

Check Your Progress

11. What do you understand by the Eulerian angles?

12. What do you understand by herpolhode and polhode?

2.9 ANSWERS TO 'CHECK YOUR PROGRESS'

---

1. Legendre transformation refers to the mathematical method for changing the basis of the description of a system from one set of independent variables to another set of independent variables.

2. The relations given by the following equations are called the Legendre transformation relations for the change of basis from (x, y) to (u, y):

$$x = -\frac{\partial f'}{\partial u} \quad \text{and} \quad v = \frac{\partial f'}{\partial y}$$

3. The Lagrangian function L that characterizes the system is, in general, a function of the generalized coordinates, the generalized velocities and time, i.e., $L = L(q_1, ..., q_s, q_1, ..., q_s, t) = L(q, q, t)$

4. The function $H(q, p, t)$ given by the following equation is known as the Hamiltonian of the system.

$$H(q, p, t) = \sum_{k=1}^{s} p_k \, \dot{q}_k - L(q, \dot{q}, t)$$

5. Mathematically, the principle is stated as

$$S = \int_{t_1}^{t_2} L dt = \text{An extremum}$$

The line integral $\int_{t_1}^{t_2} L dt$ which has been denoted above by the symbol S, is called the Hamilton's principle function, or action integral, or simply the action during the time interval from t_1 to t_2.

6. The principle states that of all possible paths along which the system may move from one point to another in its configuration space between two given time instants, say t1 and t2, which are consistent with the constraints imposed on the system, if any, the actual path which the system follows is the one for which the time integral of the Lagrangian of the system is an extremum (either maximum or minimum).

7. In most of the dynamical problems, the minimum condition for the action S is satisfied. For this reason, the principle is also called Hamilton's principle of least action.

8. Consider a mechanical system of s degree of freedom. Let the system be described by generalized coordinates q_1, ..., qs and conjugate momenta p_1, ..., ps. Let $f = f(q_1, ..., qs, p_1, ..., ps) = f(q, p)$ and $g = g(q_1, ..., qs, p_1, ...,ps) = g(q, p)$ be two dynamical variables of the system.

The Lagrange bracket of f and g with respect to the basis (q, p) is written as $\{f, g\}q.p$ and is defined as

$$\{f, g\}q\,p = \sum_{k} \left(\frac{\partial q_k}{\partial f} \frac{\partial p_k}{\partial g} - \frac{\partial p_k}{\partial f} \frac{\partial q_k}{\partial g} \right) = 0$$

9. $[qi, qj] = 0 = [pi, pj]$

$[qi, pj] = dij$

10. The quantity within the parenthesis on the right hand side of the following Equations (87) turns out to be of fundamental importance in the formal development of mechanics and is known as the Poisson bracket (PB) of F and H.

$$\frac{dF}{dt} = \sum_k \frac{\partial F}{\partial q_k} \frac{\partial H}{\partial p_k} - \sum_k \frac{\partial F}{\partial p_k} \frac{\partial H}{\partial q_k} + \frac{\partial F}{\partial t}$$

$$\frac{dF}{dt} = \sum_k \left[\frac{\partial F}{\partial q_k} \frac{\partial H}{\partial p_k} - \frac{\partial F}{\partial p_k} \frac{\partial H}{\partial q_k} \right] + \frac{\partial F}{\partial t}$$

$$\frac{dF}{dt} = [F, H] + \frac{\partial F}{\partial t}$$

11. They refer to the angles corresponding to three successive rotations of the space-fixed system performed in a particular sequence or order, such that at the end, the axes of the space-fixed system coincide with those of the body-fixed system.

12. The curve traced out on the invariable plane by the point of contact with the ellipsoid is called herpolhode and the corresponding curve described on the ellipsoid is called polhode.

2.10 SUMMARY

- Although the generalized velocities appear in the expression for L explicitly, they cannot be treated as independent variables because of being equal to the total time derivatives of the generalized coordinates.

- The generalized momenta are derived variables defined in terms of the Lagrangian L as

- Every mechanical system possesses a characteristic function of coordinates, velocities and time called the Lagrangian of the system usually denoted by the symbol L.

- Hamilton's principle is concerned with the trajectory or the path which is followed by the system point.

$$p_k = \frac{\partial L(q, \dot{q}, t)}{\partial \dot{q}_k}$$

2.11 KEY TERMS

- **Legendre transformation:** It refers to the mathematical method for changing the basis of the description of a system from one set of independent variables to another set of independent variables.

- **Mechanics:** The branch of science which deals with motion of objects under the action of forces, including the particular case in which a body continues at rest is called mechanics.

- **Mechanical system:** It is a system that manages power to complete a work that involves forces and movement. Here power means the rate of doing work or transferring heat.

- **Generalized coordinates:** The coordinates in a state space that together absolutely describe a system are called generalized coordinates. If they are selected so as to be independent of each other, the number of independent generalized coordinates matches the number of degrees of freedom of the system.

- **Angular momentum:** The product of the moment of inertia and the angular velocity is known as angular momentum.

- **Eulerian angles:** They refer to the angles corresponding to three successive rotations of the space-fixed system performed in a particular sequence or order, such that at the end, the axes of the space-fixed system coincide with those of the body-fixed system.

2.12 SELF-ASSESSMENT QUESTIONS AND EXERCISES

SHORT ANSWER QUESTIONS

1. Derive the Legendre transformation relations for the change of basis.
2. Write about the Hamilton's principle briefly.
3. What do you understand by Hamilton's canonical equations?
4. State the properties of Lagrange brackets.

LONG ANSWER QUESTIONS

1. Discuss Hamilton's principle for a conservative system. Also state why it is called the principle of least action.
2. Give an introduction to Hamilton's principle and derive the Hamiltonian of the system.
3. Give a detailed account of the Eulerian angles.
4. Discuss torque free motion of a rigid body.
5. Explain motion of symmetric top under the action of gravity.

2.13 FURTHER READING

Rao, K. Sankara. 2009. *Classical Mechanics*. New Delhi: PHI Learning Private Limited.

Upadhyaya, J.C. 2010. *Classical Mechanics,* 2nd Edition. New Delhi: Himalaya Publishing House.

Goldstein, Herbert. 2011. *Classical Mechanics,* 3rd Edition. New Delhi: Pearson Education India.

Gupta, S.L. 1970. *Classical Mechanics.* New Delhi: Meenakshi Prakashan.

Takwala, R.G. and P.S. Puranik. 1980. New Delhi: Tata McGraw Hill Publishing.

UNIT 3   HAMILTONIAN-JACOBI THEORY

Structure

3.0   INTRODUCTION

The Hamilton-Jacobi equation named for William Rowan Hamilton and Carl Gustav Jacob Jacobi, explains extremal geometry in generalizations of problems from the calculus of variations. It is an alternative formulation of classical mechanics, equivalent to other formulations such as Newton's laws of motion, Lagrangian mechanics and Hamiltonian mechanics. It is a first-order partial non-linear differential equation, especially applicable in understanding conserved quantities for mechanical systems, which may be possible even when the mechanical problem itself cannot be solved completely. In this unit you will study characteristic function and Hamilton-Jacobi equation. You will learn Hamilton's principal function, Abbreviated action and solution of mechanical problems using Hamilton-Jacobi method. Application of action angle variable and Jacobi identity is also explained.

3.1   OBJECTIVES

After going through this unit, you will be able to:

- Explain characteristic function and Hamilton-Jacobi equation
- Describe Hamilton's principal function and Hamilton's principal function
- Solve mechanical problems using Hamilton-Jacobi method and calculate motion of a body falling freely under gravity
- Describe action and angle variables, application of action angle variable to obtain the frequency of a linear harmonic oscillator

- Explain the Jacobi's identity
- Discuss the Kepler's problem in action and angle variables

3.2 HAMILTON-JACOBI EQUATION FOR HAMILTON'S PRINCIPLE FUNCTION

Consider a mechanical system of s degrees of freedom described by the generalized coordinates $q_1, \ldots, q_s$ and generalized momenta $p_1, \ldots, p_s$. The Hamiltonian for the system which involves time explicitly is

$$H = H(q, p, t) = H(q_1, \ldots, q_s, p_1, \ldots, p_s, t) \tag{3.1}$$

The Hamilton's canonical equations are

$$\dot{q}_k = \frac{\partial H}{\partial p_k}; \quad \text{and} \quad \dot{q}_k = -\frac{\partial H}{\partial q_k} \tag{3.2}$$

Let us affect a canonical transformation from the old set of variables $\{q, p\}$ to a new set $\{Q, P\}$. Let F be the generating function of the transformation. If K represents the transformed Hamiltonian then the following equations hold

$$\dot{Q}_k = \frac{\partial K}{\partial P_k}; \quad \dot{P}_k = -\frac{\partial K}{\partial Q_k} \tag{3.3}$$

and

$$K = H + \frac{\partial F}{\partial t} \tag{3.4}$$

Let us consider the above transformation to be such that the new variables are constants of motion (Q_k = Constant, P_k = Constant), so that

$$\dot{Q}_k = 0 \quad \text{and} \quad \dot{P}_k = 0 \tag{3.5}$$

According to Equation (3.3) and (3.5), we then get

$$\frac{\partial K}{\partial P_k} = 0 \quad \text{and} \quad \frac{\partial K}{\partial Q_k} = 0 \tag{3.6}$$

Equation (3.6) allows us to take the transformed Hamiltonian K identically equal to zero. Equation (3.4) then demands the generating function F to be such that

$$H(q, p, t) + \frac{\partial F}{\partial t} = 0 \tag{3.7}$$

Let us choose the generating function F to be a function of old coordinates, new constant momenta and time, i.e.,

$$F = F(q, P, t) \tag{3.8}$$

From Equation $F_1(q, q, t) = F_2(q, p, t) - \Sigma P_k Q_k$, we then have

$$p_k = \frac{\partial F}{\partial q_k} \tag{3.9}$$

In view of Equation (3.9) we may rewrite Equation (3.7) as

$$H\left(q_1, \ldots, q_s, \frac{\partial F}{\partial q_1}, \ldots, \frac{\partial F}{\partial q_s}, t\right) + \frac{\partial F}{\partial t} = 0 \qquad (3.10)$$

Equation (3.10) is referred to as the Hamilton–Jacobi equation. We may note that the Hamilton–Jacobi equation is a first-order partial differential equation in $(s + 1)$ variables, namely, $q_1, \ldots, q_s, t$.

By convention, solution of Equation (3.10) is denoted by S and is called the **Hamilton's Principal Function.**

The complete solution of Equation (3.10) involves $(s + 1)$ constants of integration. We may note that if we add a constant, say α, to the solution S, so that the solution may be written as $S + \alpha$ then we find that such a replacement of the solution satisfies Equation (3.10). Hence, we find that out of $(s + 1)$ constants mentioned above, one is an additive constant. We may thus write the general solution of the Hamilton–Jacobi equation as

$$S = S(q_1, \ldots, q_s, \alpha_1, \ldots, \alpha_s, t) + \alpha \qquad (3.11)$$

In the above, $\alpha_1, \alpha_2, \ldots, \alpha_s$ are the s independent, non-additive constants of integration. Clearly, solution of the Hamilton–Jacobi equation S is a function of s coordinates, s constants and time. This is precisely the same description as that of the generating function F considered above. The constants can be chosen as the new momenta which are constants of motion. Thus,

$$\alpha_k = p_k; \ (k = 1, \ldots, s) \qquad (3.12)$$

The new momenta which are constants can conveniently be chosen to be the momentum values $p_{01}, p_{02}, \ldots, p_{0s}$ at the initial time $t = t_0$.

Transformation equations given by Equation (3.9) can be rewritten as

$$p_k = \frac{\partial S}{\partial q_k} \qquad (3.13)$$

Equation (3.13) can be used to determine the relations between α_k, p_k and q_k at $t = t_0$.

We also have the transformation equation

$$Q_k = \frac{\partial S}{\partial P_k} = \frac{\partial S}{\partial \alpha_k} = \beta_k \ (\text{say}); \qquad k = 1, \ldots, s \quad (3.14)$$

We may choose the values of β_k as the momenta and they can also be expressed in terms of the initial values of the coordinates, namely, $q_{10}, \ldots, q_{50}$.

We may write Equation (3.14) also as

$$q_k = q_k(\alpha_k, \beta_k, t) \qquad (3.15)$$

Equation (3.15) along with the relations

$$p_k = p_k(\alpha_k, \beta_k, t) \qquad (3.16)$$

gives the solutions of the problem.

Identification of the Solution of Hamilton–Jacobi Equation

From Equation (3.8) we get

$$\frac{dF}{dt} = \sum \frac{\partial F}{\partial q_k} \dot{q}_k + \sum \frac{\partial F}{\partial p_k} \dot{p}_k + \frac{\partial F}{\partial t}$$

Using Equation (3.5) and (3.9) in the above, we obtain

$$\frac{dF}{dt} = \sum p_k \dot{q}_k + \frac{\partial F}{\partial t} \tag{3.17}$$

Equation (3.7) when used in Equation (3.17) gives

$$\frac{dF}{dt} = \sum p_k \dot{q}_k - H(q, p, t)$$

or

$$\frac{dF}{dt} = L(q, q, t) \tag{3.18}$$

where $L\left(q, \dot{q}, t\right) = \sum p_k \dot{q}_k - H(q, p, t)$, is the Lagrangian of the system.

Integration of Equation (3.18) over time leads to

$$\int \frac{dF}{dt} dt = \int L\left(q, \dot{q}, t\right)$$

or

$$F = S + \text{constant} \tag{3.19}$$

In the above, *S* represents the familiar action of the system. We thus identify the solution of the Hamilton–Jacobi equation, i.e., the Hamilton's principal function as the indefinite time integral of the Lagrangian of the system, plus a constant.

We may now summarize the method for solving a mechanical problem using Hamilton–Jacobi method. The steps are:

(*i*) Construct the Hamiltonian *H* for the system in terms of old coordinates *q*, old momenta *p* and time *t*.

(*ii*) Write the Hamilton–Jacobi equation and find its complete integral *S*, i.e., the Hamilton's principal function.

(*iii*) Differentiate *S* with respect to *s* arbitrary constants α_k and equate the derivatives to the new constants β_k to obtain *s* algebraic equations, such as

$$\frac{\partial S}{\partial \alpha_k} = \beta_k \tag{3.20}$$

(*iv*) Solve the *s* equations thus obtained to get the old coordinates q_k as functions of time *t* and the 2*s* arbitrary constants

$$q_k = q_k(\alpha_1, \alpha_2, \ldots, \alpha_s, \beta_1, \beta_2, \ldots, \beta_s, t) \tag{3.21}$$

(*v*) Obtain the old momenta according to

$$p_k = \frac{\partial S}{\partial q_k} \tag{3.22}$$

(*vi*) Use Equation (3.21) and their time derivatives to determine 2*s* initial conditions of the given mechanical system.

3.2.1 Abbreviated Action, or Hamilton's Characteristic Function

Consider a conservative mechanical system. The Hamiltonian of the system then has no explicit dependence on time and is a constant of motion representing the total energy *E* of the system. We have

$$S = \int_1^2 L dt = \int_1^2 \left[\sum p_k dq_k - H dt \right]$$

$$= \int_1^2 \sum p_k dq_k - \int_1^2 H dt$$

or $\qquad S = S_0 - E(t_2 - t_1) \tag{3.23}$

where $\qquad S_0 = \int_1^2 \sum p_k dq_k \tag{3.24}$

Substituting Equation (3.23) in the Hamilton–Jacobi equation given by Equation (3.10), we obtain

$$-E + H\left(q_1, \ldots, q_s, \frac{\partial S_o}{\partial q_1}, \ldots, \frac{\partial S_o}{\partial q_s} \right) = 0 \tag{3.25}$$

The solution S_0 of the time-independent partial differential Equation (3.25) is called abbreviated action, or Hamilton's characteristic function.

3.3 HARMONIC OSCILLATOR PROBLEM USING HAMILTON–JACOBI METHOD

A. One-dimensional Harmonic Oscillator Problem

Consider a one-dimensional harmonic oscillator of mass *m*. Let *q* and *p* be respectively the coordinate and momentum of the oscillator at any instant of time *t*.

The Hamiltonian function *H* of the oscillator is

$$H = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 q^2 \tag{3.26}$$

where $\omega = \sqrt{\dfrac{k}{m}}$, *k* being the force constant of the oscillator.

The Hamilton–Jacobi equation is given by

$$H(q, p) + \frac{\partial S}{\partial t} = 0$$

Using $p = \dfrac{\partial S}{\partial q}$, the above equation becomes

$$\frac{1}{2m}\left(\frac{\partial S}{\partial q}\right)^2 + \frac{1}{2}m\omega^2\, q^2 + \frac{\partial S}{\partial t} = 0 \qquad (3.27)$$

We note that in the above equation, the only term that involves an explicit dependence of S on t is the term $\dfrac{\partial S}{\partial t}$. We may write the solution of Equation (3.27) in the form

$$S(q, \alpha, t) = W(q, \alpha) - \alpha t \qquad (3.28)$$

where α is a constant of integration.

From Equation (3.28), we get

$$\frac{\partial S}{\partial q} = \frac{\partial W}{\partial q}$$

and

$$\frac{\partial S}{\partial t} = -\alpha$$

Substituting the above in Equation (3.27), we get

$$\frac{1}{2m}\left(\frac{\partial W}{\partial q}\right)^2 + \frac{1}{2}m\omega^2 q^2 = \alpha$$

or

$$\left(\frac{\partial W}{\partial q}\right)^2 = 2m\alpha - 2m \times \frac{1}{2}m\omega^2 q^2$$

or

$$\frac{\partial W}{\partial q} = \sqrt{2m\left(\alpha - \frac{1}{2}m\omega^2 q^2\right)} = m\omega\sqrt{\left(\frac{2\alpha}{m\omega^2} - q^2\right)}$$

Integrating the above gives

$$W = m\omega\int\sqrt{\left(\frac{2\alpha}{m\omega^2} - q^2\right)}\,dq \qquad (3.29)$$

Using Equation (3.29), the solution given by Equation (3.28) takes the form

$$S = m\omega\int\sqrt{\left(\frac{2\alpha}{m\omega^2} - q^2\right)}\,dq - \alpha t \qquad (3.30)$$

We have the transformation equations

$$P = \frac{\partial S}{\partial q} \qquad \text{and} \qquad \beta = \frac{\partial S}{\partial \alpha} \qquad (3.31)$$

Use of Equation (3.30) yields

$$p = m\omega \left(\frac{2\alpha}{m\omega^2} - q^2 \right)^{\frac{1}{2}} \tag{3.32}$$

and

$$\beta = \frac{1}{\omega} \int \frac{dq}{\sqrt{\frac{2\alpha}{m\omega^2} - q^2}} - t$$

or

$$\beta + t = -\frac{1}{\omega} \cos^{-1} q \sqrt{\frac{m\omega^2}{2\alpha}}$$

The above gives

$$q = \sqrt{\frac{2\alpha}{m\omega^2}} \cos \omega(t + \beta) \tag{3.33}$$

The constants α and β in Equation (3.33) may be related to the initial values q_0 and p_0 of coordinate and momentum of the oscillator, respectively. Let at $t = 0$, the particle be at rest at a position displaced by q_0 from the equilibrium position. We then get from Equation (3.27)

$$\left(\frac{\partial S}{\partial q} \right)_o = p_o = \sqrt{2m} \left(\alpha - \frac{m\omega^2 q_o^2}{2} \right)^{\frac{1}{2}}$$

Since $p_0 = 0$, the above result gives

$$\alpha = \frac{m\omega^2 q_o^2}{2} \tag{3.34}$$

From Equation (3.26) we find that α given by Equation (3.34) is the initial total energy of the system. Equation (3.34) gives

$$q_0 = \sqrt{\frac{2\alpha}{m\omega^2}} \tag{3.35}$$

Using Equation (3.35) in Equation (3.33), we get

$$q = q_0 \cos \omega(t + \beta) \tag{3.36}$$

Further, since $q = q_0$ at $t = 0$, according to Equation (3.36), we get

$$\cos \beta = 1 \text{ or } \beta = 0$$

We thus get

$$q = q_0 \cos \omega t \tag{3.37}$$

Conclusion

The Hamilton's principal function S affects a contact transformation which transforms the oscillator with a canonical momentum $\alpha = H =$ Constant total energy and a coordinate β which is zero at $t = 0$. We may write Hamilton's principal function of the oscillator as

$$S = m\omega\int\sqrt{q_o^2 - q^2}\,dq - \frac{m\omega^2 q_o^2 t}{2}$$

[by using Equation (3.34) in Equation (3.30)]

Using the value of q from Equation (3.36), the above gives

$$S = m\omega^2 q_o^2 \int\left(\sin^2\omega t - \frac{1}{2}\right)dt \qquad (3.38)$$

It is now easy to show that S given by Equation (3.38) is the time integral of the Lagrangian (L) of the oscillator. We have

$$L = \text{Kinetic energy} - \text{Potential energy}$$

$$= \frac{1}{2}m\dot{q}^2 - \frac{1}{2}m\omega^2 q^2$$

Substituting from Equation (3.36), the above gives

$$L = \frac{m\omega^2 q_o^2}{2}\left[\sin^2\omega t - \cos^2\omega t\right]$$

$$= \frac{m\omega^2 q_o^2}{2}\left[\sin^2\omega t - 1 + \sin^2\omega t\right]$$

or

$$L = \frac{m\omega^2 q_o^2}{2}\left[2\sin^2\omega t - 1\right]$$

Clearly,

$$\int L\,dt = m\omega^2 q_o^2 \int\left(\sin^2\omega t - \frac{1}{2}\right)dt$$

B. Motion of a Body Falling Freely under Gravity

Consider a body of mass m falling freely under gravity. At some instant of time t, let v be the velocity of the body and z be its height above the ground. The kinetic energy of the body is

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

(p being the linear momentum of the body)

The potential energy (gravitational potential energy) is

$$V = mgz$$

The Hamiltonian which represents the total energy E of the body is thus

$$H = \frac{p^2}{2m} + mgz = E \qquad (3.39)$$

We may write H in the usual notation of S as

$$H = \frac{1}{2m}\left(\frac{\partial S}{\partial z}\right)^2 + mgz \qquad (3.40)$$

where S is the Hamilton's principle function.

Hamilton–Jacobi equation is given by

$$H + \frac{\partial S}{\partial t} = 0$$

Using Equation (3.40), the above gives

$$\frac{1}{2m}\left(\frac{\partial S}{\partial z}\right)^2 + mgz + \frac{\partial S}{\partial t} = 0 \qquad (3.41)$$

The general solution of Equation (3.41) can be written as

$$S(z, \alpha, t) = W(z, \alpha) - \alpha t \qquad (3.42)$$

From the above, we get

$$\left(\begin{array}{c} \dfrac{\partial S}{\partial z} = \dfrac{\partial W}{\partial z} \\ \dfrac{\partial S}{\partial t} = -\alpha \end{array}\right)$$

$$\frac{\partial S}{\partial t} = -\alpha \qquad (3.43)$$

Using Equation (3.43), Equation (3.41) becomes

$$\frac{1}{2m}\left(\frac{\partial W}{\partial z}\right)^2 + mgz - \alpha = 0$$

or

$$\left(\frac{\partial W}{\partial z}\right) = \left[2m(\alpha - mgz)\right]^{\frac{1}{2}}$$

Integrating over the variable z, we obtain

$$W = \int \left[2m(\alpha - mgz)\right]^{\frac{1}{2}} dz + A \qquad (3.44)$$

where A is the constant of integration.

Using the result given by Equation (3.44) in Equation (3.42), we get

$$S = \int \left[2m(\alpha - mgz)\right]^{\frac{1}{2}} dz + C - \alpha t \qquad (3.45)$$

We obtain from the above

$$\frac{\partial S}{\partial \alpha} = \frac{\sqrt{2m}}{2} \int \frac{dz}{\sqrt{\alpha - mgz}} - t = \beta$$

or

$$\beta + t = \frac{\sqrt{2m}}{2}\frac{2\sqrt{\alpha - mgz}}{(-mg)} = -\sqrt{\frac{2}{m}}\frac{(\alpha - mgz)^{\frac{1}{2}}}{g}$$

or

$$(\alpha - mgz)^{\frac{1}{2}} = -g\sqrt{\frac{m}{2}}(\beta + t)$$

or

$$\alpha - mgz = g^2\frac{m}{2}(\beta + t)^2$$

or

$$mgz = \alpha - \frac{mg^2}{2}(\beta + t)^2$$

or

$$z = \frac{\alpha}{mg} - \frac{g}{2}(\beta + t)^2 \qquad (3.46)$$

Let $z = z_0$ and $p = 0$, initially at $t = 0$.

Then we have

$$p = \frac{\partial W}{\partial z} = \sqrt{2m(\alpha - mgz_o)} = 0$$

The above gives

$$\alpha = mgz_0 \qquad (3.47)$$

Using Equation (3.47) in Equation (3.46), we obtain

$$z = \frac{mgz_o}{mg} - \frac{g}{2}(\beta + t)^2$$

or

$$z = z_o - \frac{g}{2}(\beta + t)^2$$

Since at $t = 0$, $z = z_0$ we obtain

$$z_0 = z_o - \frac{g}{2}\beta^2$$

The above gives $\beta = 0$.

Hence, we obtain

$$z = z_o - \frac{1}{2}gt^2$$

This is the equation of motion for the freely falling body.

Illustrative Examples

Example 3.1: A single particle is moving under the Hamiltonian $H = \frac{p^2}{2}$.

(*a*) Solve the Hamilton-Jacobi equation for the generating function $s(q, d, t)$.

(*b*) If β and α are the transformed coordinate and momentum respectively, find the canonical transformation $q = q(\beta, \alpha)$ and $p = p(\beta, \alpha)$

Solution:

(*a*) We have the Hamilton-Jacobi equation as

$$\frac{\partial s}{\partial t} + H(q, p, t) = 0 \qquad (i)$$

where $\qquad\qquad\qquad p = \dfrac{\partial s}{\partial q} \qquad (ii)$

In the problem, $H = \dfrac{p^2}{2}$ and hence using Equation (*ii*) we obtain and according to Equation (*i*)

$$\frac{\partial s}{\partial t} + \frac{1}{2}\left(\frac{\partial s}{\partial q}\right)^2 = 0 \qquad (iii)$$

As H does not depend on q and t, the two terms on the left hand side of Equation (*iii*) can be set equal to, say, $-r$ and $+r$ where r is a function of p.

We then get

$$S = \sqrt{2\gamma}\ q - \gamma t \qquad (iv)$$

Setting $\alpha = \sqrt{2\gamma}$ we get the generating function

$$S = \alpha q - \frac{1}{2}\alpha^2 t \qquad (v)$$

(*b*)   Considering the constant α to be the new momentum P we have the transformation equations

$$p = \frac{\partial s}{\partial q} = \alpha \qquad (vi)$$

$$Q = \frac{\partial s}{\partial P} = \frac{\partial s}{\partial \alpha} = q - \alpha t \qquad (vii)$$

Example 3.2: The Hamiltonian of a system is given by

$$H = \frac{p^2}{2} - \frac{\mu}{q}$$

Solve the corresponding Hamilton-Jacobi equation.

Solution:

Given $\qquad\qquad\qquad H = \dfrac{p^2}{2} - \dfrac{\mu}{q} \qquad (i)$

The Hamilton-Jacobi equation for the system is

$$\frac{\partial s}{\partial t} + H(q, p, t) = 0 \qquad (ii)$$

We have $$p = \frac{\partial s}{\partial q}$$

So that using Equation (*i*) in Equation (*ii*) we get

$$\frac{\partial s}{\partial t} + \frac{1}{2}\left(\frac{\partial s}{\partial q}\right)^2 - \frac{\mu}{q} = 0 \qquad (iii)$$

We may set

$$S = f(t) + g(q) \qquad (iv)$$

where $f(t)$ is a function of only t while $g(q)$ is a function of only q.

Using Equation (*iv*), Equation (*iii*) becomes

$$\frac{df(t)}{dt} + \frac{1}{2}\left(\frac{dg(q)}{dq}\right)^2 - \frac{\mu}{q} = 0$$

or $$\frac{df(t)}{dt} = \frac{\mu}{q} - \frac{1}{2}\left(\frac{dg(q)}{dq}\right)^2 \qquad (v)$$

The L.H.S. of Equation (*v*) depends only on time while the R.H.S. depends only on q. Hence, we may set both sides equal to a constant say $\frac{\mu}{\alpha}$. Thus we get

$$\frac{df(t)}{dt} = \frac{\mu}{\alpha}$$

Integrating, we get

$$f(t) = \frac{\mu}{\alpha}$$

Also $$\frac{\mu}{q} - \frac{1}{2}\left(\frac{dg(q)}{dq}\right)^2 = \frac{\mu}{\alpha}$$

or $$\frac{1}{2}\left(\frac{dg(q)}{dq}\right)^2 = \frac{\mu}{q} - \frac{\mu}{\alpha}$$

or $$\frac{dg(q)}{dq} = \sqrt{2\left(\frac{\mu}{q} - \frac{\mu}{\alpha}\right)}$$

Carrying out the integration we get

$$g(q) = \sqrt{2\mu\alpha}\ \text{arc}\ \sin\left[\sqrt{\frac{q}{\alpha}} + \left(\frac{2\mu q(\alpha - q)}{\alpha}\right)^{1/2}\right]$$

Then the solution of the Hamilton-Jacobi equation is

$$S = \frac{\mu}{\alpha}t + \sqrt{2\mu^2}\ \text{arc}\ \sin\left[\sqrt{\frac{q}{\alpha}} + \left(\frac{2\mu q(\alpha - q)}{\alpha}\right)^{1/2}\right]$$

Example 3.3: Write down the Hamilton-Jacobi equation for a three-dimensional harmonic oscillator and obtain the solution of the equation.

Solution:

The Hamiltonian function of the three-dimensional oscillator can be written as

$$H = \frac{1}{2m}(p_1^2 + p_2^2 + p_3^2) + \frac{1}{2}p_1 q_1^2 + \frac{1}{2}k_2 q_2^2 + \frac{1}{2}k_3 q_3^2 \quad (i)$$

where we have assumed the spring constants k_1, k_2 and k_3 along the three cartesian axes (referred in the above by the suffixes 1, 2, 3) to be different.

Since the oscillator is a conservative system, the Hamiltonian H does not depend on time explicitly and instead it is constant of motion. Thus

$$H = E(\text{say}) \quad (ii)$$

The Hamilton's principal function S is thus

$$S(q_j, P_j, t) = W(q_j, p_j) - Et \quad (iii)$$

We have

$$p_j = \frac{\partial W}{\partial q_j} \quad (iv)$$

Using the above we may write the Hamiltonian given by Equation (i) as

$$H = \sum_{j=1}^{3}\left[\frac{1}{2m}\left(\frac{\partial W}{\partial q_j}\right)^2 + \frac{1}{2}k_j q_j^2\right] = E$$

The above can be alternatively written as

$$\sum_{j=1}^{3}\left[\left(\frac{\partial W}{\partial q_j}\right)^2 + mk_j \cdot q_j^2\right] = 2mE \quad (v)$$

Equation (3.5) is the Hamilton-Jacobi equation for the oscillator. It can be solved using the method of separation of variables according to which we may write

$$W = W_1(q_1) + W_2(q_2) + W_3(q_3) \quad (vi)$$

Substituting Equation (vi) in Equation (v) we get three equations.

$$\left(\frac{\partial W_1}{\partial q_1}\right)^2 + mk_1 q_1^2 = 2m\alpha_1$$

$$\left(\frac{\partial W_2}{\partial q_2}\right)^2 + mk_2 q_2^2 = 2m\alpha_2 \quad (vii)$$

$$\left(\frac{\partial W_3}{\partial q_3}\right)^2 + mk_3 q_3^2 = 2m\alpha_3$$

In the above

$$\alpha_1 + \alpha_2 + \alpha_3 = E \quad (viii)$$

On integrating Equation (*vii*) we obtain

$$W_1 = \int (2m\alpha_1 - mk_1 q_1)^2 \, dq_1$$

$$W_2 = \int (2m\alpha_2 - mk_2 q_2^2) \, dq_2 \qquad (ix)$$

$$W_3 = \int (2m\alpha_3 - mk_3 q_3^2)^2 \, dq_3$$

The constants α_1, α_2, α_3 are designated as the new momenta P_1, P_2, P_3 respectively. The new constant coordinates are given by

$$Q_1 = \frac{\partial W_1}{\partial P_1} = \frac{\partial W_1}{\partial \alpha_1} = \sqrt{\frac{m}{2}} \int \frac{dq_1}{\sqrt{\alpha_1 - \frac{1}{2} k_1 q_1^2}}$$

$$Q_2 = \frac{\partial W_2}{\partial P_2} = \frac{\partial W_2}{\partial \alpha_2} = \sqrt{\frac{m}{2}} \int \frac{dq_2}{\sqrt{\alpha_2 - \frac{1}{2} k_2 q_2^2}} \qquad (x)$$

$$Q_3 = \frac{\partial W_3}{\partial P_3} = \frac{\partial W_3}{\partial \alpha_3} = \sqrt{\frac{m}{2}} \int \frac{dq_3}{\sqrt{\alpha_3 - \frac{1}{2} k_3 q_3^2}}$$

For the conservative system

$$H = K = E = \alpha_1 + \alpha_2 + \alpha_3 \qquad (xi)$$

The equations of motion in the new coordinates are

$$\theta_1 = \frac{\partial k}{\partial P_1} = \frac{\partial E}{\partial \alpha_1} = 1$$

$$\theta_2 = \frac{\partial k}{\partial P_2} = \frac{\partial E}{\partial \alpha_2} = 1 \qquad (xii)$$

$$\theta_2 = \frac{\partial k}{\partial P_3} = \frac{\partial E}{\partial \alpha_3} = 1$$

The above equations on integration give

$$Q_1 = t + \beta_1$$

$$Q_2 = t + \beta_2 \qquad (xiii)$$

$$Q_3 = t + \beta_3$$

From Equation (*x*) and (*xiii*) we obtain

$$t + \beta_1 = \sqrt{\frac{m}{2}} \int \frac{dq_1}{\sqrt{\alpha_1 - \frac{1}{2} k_1 q_1^2}}$$

or

$$\sqrt{\frac{m}{2\alpha_1}} \int \frac{dq_1}{\sqrt{1 - \frac{k_1 q_1^2}{2\alpha_1}}} = t + \beta_1$$

or

$$\sqrt{\frac{m}{k_1}} \sin^{-1} q_1 \sqrt{\frac{k_1}{2\alpha_1}} = t + \beta_1$$

or $$q_1 = \sqrt{\frac{2\alpha_1}{mw_1^2}} \sin w_1(t + \beta_1) \qquad (xiv)$$

Similarly, we obtain

$$q_2 = \sqrt{\frac{2\alpha_2}{mw_2^2}} \sin w_2(t + \beta_2) \qquad (xv)$$

$$q_3 = \sqrt{\frac{2\alpha_3}{mw_3^2}} \sin w_3(t + \beta_3) \qquad (xvi)$$

In the above

$$w_j = \sqrt{\frac{k_j}{m}}, \ j = 1, 2, 3 \qquad (xvii)$$

Example 3.4: Using Hamilton-Jacobi method discuss the motion of a particles of mass m moving in a uniform gravitational field along the z-axis.

Solution:

Let at any instant of time t, p_x, p_y, p_z be the components of momentum of the particle along the x, y and z axes, respectively. We then have the kinetic energy of the particle as.

$$T = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2)$$

If the z-coordinate of the particle at the instant t be z, is

$$V = mgz$$

Thus, the Hamiltonian of the particle is

$$H = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + mgz \qquad (i)$$

The Hamilton-Jacobi equation for the particle is thus

$$\frac{1}{2m}\left[\left(\frac{\partial s_o}{\partial x}\right)^2 + \left(\frac{\partial s_o}{\partial y}\right)^2 + \left(\frac{\partial s_o}{\partial z}\right)^2\right] + mgz = E \qquad (ii)$$

Let us set

$$s_0 = S_1(x) + S_2(y) + S_3(z) \qquad (iii)$$

Substituting Equation (*iii*) in Equation (*ii*) we get

$$\frac{1}{2m}\left[\left(\frac{\partial s_1}{\partial x}\right)^2 + \left(\frac{\partial s_2}{\partial y}\right)^2 + \left(\frac{\partial s_3}{\partial z}\right)^2\right] + mgz = E$$

or $$\frac{1}{2m}\left[\alpha_1^2 + \alpha_2^2 + \left(\frac{\partial s_3}{\partial z}\right)^2\right] + mgz = E \qquad (iv)$$

From (*i*) we find x and y to be the cyclic coordinates and hence

$$\frac{\partial s_1}{\partial x} = \alpha_1 = \text{Constant}$$

$$\frac{\partial s_2}{\partial y} = \alpha_2 = \text{Constant} \tag{v}$$

We may write Equation (*iv*) as

$$\alpha_1^2 + \alpha_2^2 + \left(\frac{\partial s_3}{\partial z}\right)^2 = 2mE - 2m^2gz$$

or

$$\frac{\partial s_z}{\partial z} = \left[2mE - 2m^2gz - \alpha_1^2 - \alpha_2^2\right]^{1/2}$$

Integrating the above we get

$$s_3 = \int \left[2mE - 2m^2 g\, z - \alpha_1^2 - d_2^2\right]^{1/2} dz \tag{vi}$$

Integrating Equation (*v*) we obtain

$$s_1 = \int \alpha_1 dx, \ s_2 = \int \alpha_2 dy \tag{vii}$$

Substituting Equation (*vi*) and (*vii*) in Equation (*iii*) we obtain

$$s_0 = \int \alpha_1 dx + \int \alpha_2 dy + \int [2m(E - mgz) - \alpha_1^2 - \alpha_2^2]^{1/2} dz \tag{viii}$$

At $t = t_0, x_0, y_0, z_0$ be the coordinates of the particle, we get the Hamilton-Jacobi function for the particle as

$$s = s_0 - E\int_{t_0}^{t} dt$$

or

$$S = \int_{x_0}^{x} \alpha_1 dx + \int_{y_0}^{y} \alpha_2 dy + \int_{z_0}^{z} \left[2m(E - mgz) - \alpha_1^2 - d_2^2\right]^{1/2} dz - E(t - t_0) \tag{ix}$$

We now have

$$\beta_1 = \frac{\partial s}{\partial \alpha_1}, \ \beta_2 = \frac{\partial s}{\partial \alpha_2}, \ \beta_3 = \frac{\partial s}{\partial \alpha_3} = \frac{\partial s}{\partial E}$$

We thus get

$$\beta_1 = \frac{\partial}{\partial \alpha_1}\int_{x_0}^{x} \alpha_1 dx + \int_{z_0}^{z} \alpha_1[2m(E - mgz) - \alpha_1^2 - \alpha_2^2]^{-1/2} dz$$

or

$$\beta_1 = x - x_0 - \int_{z_0}^{z} \alpha_1[2m(E - mgz) - \alpha_1^2 - \alpha_2^2]^{-1/2} dz \tag{x}$$

Similarly,

$$\beta_2 = y - y_0 - \int_{z_0}^{z} \alpha_2 [2m(E - mgz) - \alpha_1^2 - \alpha_2^2]^{-1/2} dz \tag{xi}$$

$$\beta_3 = t - t_0 = \int_{z_0}^{z} m[2m(E - mgz) - \alpha_1^2 - \alpha_2^2]^{-1/2} dz \tag{xii}$$

Using the initial values at $t = t_0$ we get

$$\beta_1 = 0 \qquad (xiii)$$

$$\beta_2 = 0 \qquad (xiv)$$

We may hence write

$$x - x_0 = \frac{1}{m}\alpha_1(t - t_0) \qquad (xv)$$

$$y - y_0 = \frac{1}{m}\alpha_2(t - t_0) \qquad (xvi)$$

$$\left[2m(E - mgz_0) - \alpha_1^2 - \alpha_2^2\right]^{1/2} \frac{\overline{z - z_0}\, \overline{t - t_0}}{m} - \frac{g}{2}(t - t_0)^2 \quad (xvii)$$

Equation (*xvii*) can be written as

$$z = -\frac{g}{2}(t - t_0)^2 + \frac{1}{m}\left[2m(E - mgz_0)\alpha_1^2 - \alpha_2^2\right]^{1/2}(t - t_0) + z_0$$

The above shows that the *z*-coordinate of the particle varies with time in a parabolic manner.

Example 3.5: Solve the problem of the motion of particle of mass *m* moving under a central force using Hamilton-Jacobi method.

Solution:

We know that such a motion takes place in a plane and is hence a two-dimensional motion. The convenient generalized coordinates are the polar coordinates *r* and θ in terms of which the Hamiltonian of the particle is given by

$$H = \frac{1}{2m}\left(p_r^2 + \frac{1}{r^2}p_\theta^2\right) + U(r) \qquad (i)$$

Where $U(r)$ is potential corresponding to the central force, p_r and p_q are the momenta congugate to the coordinates *r* and θ respectively. We may write

$$s = s_0 - Et \text{ (since the system is conservative)}(ii)$$

The Hamilton-Jacobi equation for the problem is then

$$\frac{1}{2m}\left(\frac{\partial s_0}{\partial r}\right)^2 + \frac{1}{2mr^2}\left(\frac{\partial s_0}{\partial \theta}\right)^2 + U(r) + U(r) - E = 0 \qquad (iii)$$

Using the method of separation of variables and noting that θ is cyclic, we can obtain s_0 by integrating Equation (*iii*). The s_0 thus obtained when substituted in Equation (*ii*) gives

$$s = p_\theta + \theta + \int\left[2m\{E - U(r)\} - \frac{p_\theta^2}{r^2}\right]^{1/2}d\theta - Et \qquad (iv)$$

For the problem we have

$$\alpha_1 = E \text{ and } \alpha_2 = p_q. \qquad (v)$$

$$\beta_1 = \frac{\partial s}{\partial \alpha_1} = \frac{\partial S}{\partial E} = \int m\left[2m\{E - U(r)\} - \frac{p_\theta^2}{r^2}\right]^{-1/2}dr - t$$

$$\beta_2 = \frac{\partial s}{\partial \alpha_2} = \frac{\partial s}{\partial p_\theta} = \theta - \int \frac{p_\theta}{r^2} \left[2m\{E - U(r)\} - \frac{p_\theta^2}{r^2} \right]^{-\frac{1}{2}} \quad (vi)$$

Both β_1 and β_2 are constants. Hence we may write Equation (v) and (vi) alternatively as

$$t = \int m \left[2m\{E - U(r)\} - \frac{p_\theta^2}{r^2} \right]^{-1/2} + \text{constant} \quad (vii)$$

$$\theta = \int \frac{p_\theta}{r^2} \left[2m\{E - U(r)\} - \frac{p_\theta^2}{r^2} \right]^{-1/2} + \text{constant} \quad (viii)$$

The above equations give the dependence of r on t and θ and hence give the path of motion of the particle.

3.3.1 Separation of Variables in the Hamilton-Jacobi Equation : Action and Angle Variables

Let us consider a periodic system having one degree of freedom. Let q be the generalized coordinate and p the generalized momentum which describe the system. Considering the system to be conservative we have the Hamiltonian of the system given by

$$H = H(q, p) = \alpha(\text{constant}) \qquad (3.48)$$

According to the above, the momentum p for the system is a function of q and the constant α, i.e.,

$$p = p(q, \alpha)$$

A knowledge of the above function gives the trajectory of the system in its two-dimensional phase space.

A practical example of such a conservative one-dimensional system is a linear harmonic oscillator for which p is given by Equation (3.32)

$$p = m\omega \left(\frac{2\alpha}{m\omega^2} - q^2 \right)^{\frac{1}{2}}$$

or
$$p^2 = \frac{m^2\omega^2 2\alpha}{m\omega^2} - m^2\omega^2 q^2$$

or
$$p^2 = 2m\alpha - m^2\omega^2 q^2 \qquad (3.49)$$

The above can be written as

$$\frac{p^2}{2m\alpha} + \frac{m^2\omega^2 q^2}{2m\alpha} = 1$$

or
$$\frac{p^2}{2m\alpha} + \frac{q^2}{\left(\dfrac{2\alpha}{m\omega^2} \right)} = 1 \qquad (3.50)$$

Clearly, the representative point in the phase space of the oscillator traces an ellipse as shown in the Fig. 3.1.
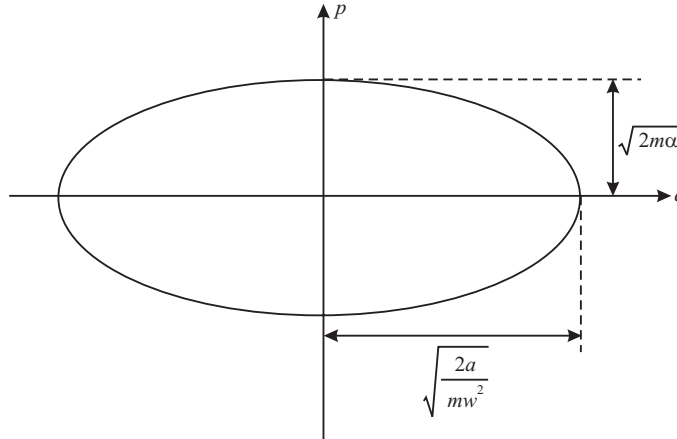
Fig. 3.1

We now introduce a new variable J for the oscillator according to

$$J = \oint p \, dq \tag{3.51}$$

where the integration is taken over one period round the ellipse. J has the dimensions of angular momentum (moment of linear momentum) and is called the phase integral or the action variable for the oscillator.

Substituting for p from Equation (3.49) in Equation (3.51) we get

$$J = \oint \sqrt{2m\alpha - m^2\omega^2 q^2} \, dq$$

The above yields J as a function of α (because q is integrated out) or the Hamiltonian H, i.e.,

$$J = J(\alpha) = J(H) \tag{3.52}$$

Alternatively, we obtain α or H as a function J

$$a = H = H(J) \tag{3.53}$$

The Hamilton's characteristic function S_0 can be written as

$$S_0 = S_0(q, J) \tag{3.54}$$

The generalized coordinate conjugate to J is called the angle variable and is defined by the transformation equation

$$w = \frac{\partial \delta_o}{\partial J} \tag{3.55}$$

The other transformation equation which gives p is

$$p = \frac{\partial \delta_o}{\partial q} \tag{3.56}$$

We thus get the equation of motion for the angle variable to be

$$\dot{w} = \frac{\partial H(J)}{\partial J} = \nu(J) \tag{3.57}$$

where $\nu(J)$ is a constant function of the action variable J only.

We may write the solution of Equation (57) as

$$w = \nu t + \beta \tag{3.58}$$

where β is a constant. We find the angle variable to vary linearly with time.

If Δw be the change in the variable w as q undergoes a change of one cycle, we get

$$\Delta w = \oint \frac{\partial W}{\partial q} dq$$

or

$$\Delta w = \oint \frac{\partial}{\partial q} \frac{\partial \delta_o}{\partial J} dq = \oint \frac{\partial^2 \delta_o}{\partial q \partial J} dq$$

or

$$\Delta w = \frac{d}{dJ} \oint \frac{\partial \delta_o}{\partial q} dq \ (J \text{ being independent of } q) \tag{3.59}$$

Using Equation (3.56) in the above, we obtain

$$\Delta w = \frac{d}{dJ} \oint p dq = \frac{dJ}{dJ} = 1 \ (\text{using Equation(51)}) \tag{3.60}$$

If T be the time required for q to complete one cycle then according to Equation (3.58) we get

$$w = \nu t \tag{3.61}$$

In view of Equation (3.61), Equation (3.60) gives

$$\nu t = 1$$

or

$$T = \frac{1}{\nu} \tag{3.62}$$

It is clear from the above that since T represents the time period for one cycle q, ν must represent the frequency, i.e., ν gives the number of cycles of q in unit time.

Since ν has the dimensions of frequency, according to Equation (3.58), w must have the dimension of angle.

We observe that use of action angle variables in periodic conservative mechanical systems provides simplified method of obtaining time period, or frequency of motion without requiring any detailed treatment.

Application of Action Angle Variable to Obtain the Frequency of a Linear Harmonic Oscillator

Consider the mechanical problem of a one-dimensional harmonic oscillator. The action variable J for the oscillator is given by

$$J = \oint p dq$$

Using p given by Equation (3.32), the above gives

$$J = \oint \sqrt{2m\alpha - m^2\omega^2 q^2} \, dq \qquad (3.63)$$

where $\alpha = H = E =$ Total energy of the system and $\omega = \sqrt{\dfrac{k}{m}}$ being the force constant.

To evaluate the integral in Equation (3.63), let us define a variable θ as

$$q = \sqrt{\dfrac{2\alpha}{m\omega^2}} \sin\theta \qquad (3.64)$$

From Equation (3.64) and (3.63) we get

$$J = \int_0^{2\pi} \sqrt{2m\alpha - m^2\omega^2 \dfrac{2\alpha}{m\omega^2}\sin^2\theta} \times \sqrt{\dfrac{2\alpha}{m\omega^2}}\cos\theta \, d\theta$$

$$J = \int_0^{2\pi} \sqrt{2m\alpha\left(1 - \sin^2\theta\right)} \sqrt{\dfrac{2\alpha}{m\omega^2}}\cos\theta \, d\theta$$

or

$$= \int_0^{2\pi} \sqrt{2m\alpha\cos^2\theta \dfrac{2\alpha}{m\omega^2}}\cos\theta \, d\theta$$

$$= \int_0^{2\pi} \dfrac{2\alpha}{\omega}\cos^2\theta \, d\theta = \int_0^{2\pi} \dfrac{\alpha}{\omega}\left(1 + \cos^2\theta\right) d\theta$$

or

$$J = \dfrac{\alpha}{\omega}\left[2\pi + \left\{\dfrac{\sin^2\theta}{2}\right\}_0^{2\pi}\right]$$

or

$$J = \dfrac{2\pi\alpha}{\omega} = \dfrac{2\pi E}{\omega} \qquad (3.65)$$

The above gives

$$\alpha = H = \dfrac{\omega J}{2\pi} \qquad (3.66)$$

Using Equation (3.57), we get the frequency of the oscillator to be

$$v = \dfrac{\partial H}{\partial J}$$

or

$$v = \dfrac{\partial}{\partial J}\left(\dfrac{\omega J}{2\pi}\right) = \dfrac{\omega}{2\pi} = \dfrac{1}{2\pi}\sqrt{\dfrac{k}{m}}$$

3.3.2 Jacobi's Identity

If f, g and k are any three functions of the coordinates (q's) and the momenta (p's) describing a mechanical system, then the following relation holds

$$[f,[g, k]] + [g, [k, f]] + [k, [f, g]] = 0 \qquad (3.67)$$

The relation expressed in Equation (3.67) is known as **Jacobi's identity**.

Proof: Using the definition and properties of Poisson bracket we get

$$[f, [g, k]] - [g, [f, k]] = \left[f, \sum_j \left(\frac{\partial g}{\partial q_j} \frac{\partial k}{\partial p_j} - \frac{\partial g}{\partial p_j} \frac{\partial k}{\partial q_j} \right) \right]$$

$$- \left[g, \sum_j \left(\frac{\partial f}{\partial q_j} \frac{\partial k}{\partial p_j} - \frac{\partial f}{\partial p_j} \frac{\partial k}{\partial q_j} \right) \right] \tag{3.68}$$

Let us put

$$\sum_j \frac{\partial g}{\partial q_j} \frac{\partial k}{\partial p_j} = a$$

$$\sum_j \frac{\partial g}{\partial p_j} \frac{\partial k}{\partial q_j} = b$$

$$\sum_j \frac{\partial f}{\partial q_j} \frac{\partial k}{\partial p_j} = c \tag{3.69}$$

$$\sum_j \frac{\partial f}{\partial p_j} \frac{\partial k}{\partial q_j} = d$$

Using Equation (3.69), we may write Equation (3.68) as

$$\begin{aligned}
\left[f, [g, k] \right] - \left[g, [f, k] \right] &= [f, (a - b)] - [g, (c - d)] \\
&= [f, a] - [f, b] - [g, c] + [g, d] \\
&= \left[f, \sum_j \left(\frac{\partial g}{\partial q_j} \frac{\partial k}{\partial p_j} \right) \right] - \left[f, \sum_j \left(\frac{\partial g}{\partial p_j} \frac{\partial k}{\partial q_j} \right) \right] \\
&\quad - \left[g, \sum_j \left(\frac{\partial f}{\partial q_j} \frac{\partial k}{\partial p_j} \right) \right] + \left[g, \sum_j \left(\frac{\partial f}{\partial p_j} \frac{\partial k}{\partial q_j} \right) \right] \\
&= \left[f, \sum_j \frac{\partial g}{\partial q_j} \right] \sum_j \frac{\partial k}{\partial p_j} + \left[f, \sum_j \frac{\partial k}{\partial p_j} \right] \sum_j \frac{\partial g}{\partial q_j} \\
&\quad - \left[f, \sum_j \frac{\partial g}{\partial p_j} \right] \sum_j \frac{\partial k}{\partial q_j} - \left[f, \sum_j \frac{\partial k}{\partial q_j} \right] \sum_j \frac{\partial g}{\partial p_j} \\
&\quad - \left[g, \sum_j \frac{\partial f}{\partial q_j} \right] \sum_j \frac{\partial k}{\partial p_j} - \left[g, \sum_j \frac{\partial k}{\partial p_j} \right] \sum_j \frac{\partial f}{\partial q_j} \\
&\quad + \left[g, \sum_j \frac{\partial f}{\partial p_j} \right] \sum_j \frac{\partial k}{\partial q_j} + \left[g, \sum_j \frac{\partial k}{\partial q_j} \right] \sum_j \frac{\partial f}{\partial p_j} \\
&= \sum_i \left\{ \frac{\partial k}{\partial q_j} \left(\left[\frac{\partial f}{\partial p_j}, g \right] + \left[f, \frac{\partial g}{\partial p_j} \right] \right) + \frac{\partial k}{\partial p_j} \left(\left[\frac{\partial f}{\partial q_j}, g \right] + \left[f, \frac{\partial g}{\partial q_j} \right] \right) \right\}
\end{aligned}$$

$$+ \sum_i \left\{ \frac{\partial g}{\partial q_j} \left[f, \frac{\partial k}{\partial p_j} \right] - \frac{\partial g}{\partial p_j} \left[f, \frac{\partial k}{\partial q_j} \right] - \frac{\partial f}{\partial q_j} \left[g, \frac{\partial k}{\partial p_j} \right] + \frac{\partial f}{\partial p_j} \left[g, \frac{\partial k}{\partial q_j} \right] \right\} \quad (3.70)$$

$$\frac{\partial}{\partial x}[f,g] = \left[\frac{\partial f}{\partial x}, g \right] + \left[f, \frac{\partial g}{\partial x} \right] \quad (3.71)$$

Using Equation(3.71) in Equation(3.70), we get

$$[f, [g, k]] - [g, [f, k]] = \sum_j \left\{ -\frac{\partial k}{\partial q_j} \frac{\partial}{\partial p_j}[f,g] + \frac{\partial k}{\partial p_j} \frac{\partial}{\partial q_j}[f,g] \right\}$$

(the last four terms on the RHS of Equation (3.70) on expansion cancel each other out)

or
$$\left[f,[g,k] \right] - \left[g,[f,k] \right] = -\left[k,[f,g] \right]$$

or
$$\left[f,[g,k] \right] - \left[g,[f,k] \right] + \left[k,[f,g] \right] = 0$$

or
$$\left[f,[g,k] \right] + \left[g,[f,k] \right] + \left[k,[f,g] \right] = 0$$

The above is the Jacobi's identity.

Jacobi's identity is helpful in finding the constants or integrals of motion for a system.

3.4 KEPLER'S PROBLEM IN ACTION AND ANGLE VARIABLES

The Hamiltonian in terms of polar coordinate in the orbit's plane can be written as

$$H = \frac{1}{2m} \left(p_r^2 + \frac{\alpha_\theta^2}{r^2} \right) + V(r) \quad (3.72)$$

where α_θ is given as

$$p_\theta^2 + \frac{p_\phi^2}{\sin^2\theta} = \alpha_\theta{}^2 \quad (3.73)$$

$$V(r) = -\frac{k}{r} \quad (3.74)$$

$$W = W_r(r, \alpha) + W_\theta(\theta, \alpha) + W_\phi(\phi, \alpha) \quad (3.75)$$

When E < 0, there will be oscillation in r and θ, along with rotation in ϕ.

$$W_r = \pm \sqrt{2m(\alpha_1 - V(r)) - \frac{\alpha_\theta^2}{r^2}} \, dr \quad (3.76)$$

$$W_\theta = \pm \int \sqrt{\alpha_\theta^2 - \frac{\alpha_\phi^2}{\sin^2\theta}} \, d\theta$$

$$W_\phi = \alpha_\phi \phi \quad (3.77)$$

Now,

$$J_\phi = \oint p_\phi d\phi = \oint \frac{\partial w}{\partial \phi} d\phi = \oint \alpha_\phi d\phi \tag{3.78}$$

For ϕ, the constant p_ϕ is taken to be periodic with period 2π. This refers to the which corresponds to returning of the particle at its original point.

Therefore,

$$J_\phi = 2\pi\alpha_\phi \tag{3.79}$$

where α_ϕ represents the angular momentum along z-axis.

Similarly,

$$J_\theta = \oint p_\theta d\theta = \oint \frac{\partial w}{\partial \theta} d\theta = \oint \pm \sqrt{\alpha_\theta^2 - \frac{\alpha_\phi^2}{\sin^2\theta}} \, d\theta \tag{3.80}$$

Take $\cos\gamma = \frac{\alpha_\phi}{\alpha_\theta}$, , fraction of angular momentum.

From equation (3.80),

$$J_\theta = \alpha_\theta \oint \pm \sqrt{\alpha_\theta^2 - \frac{\cos^2\gamma}{\sin^2\theta}} \, d\theta \tag{3.81}$$

For $\sin\theta_0 = \cos\gamma$, then $p_\theta = 0$ at the turning points, $\theta \in \{\theta_0, \pi - \theta_0\}$.

Equation (3.81) can be written as

$$J_\theta = 4\alpha_\theta \int_{\frac{\pi}{2}}^{\theta_0} \sqrt{\alpha_\theta^2 - \frac{\cos^2\gamma}{\sin^2\theta}} \, d\theta \tag{3.82}$$

Substituting $\cos\theta = \sin\gamma \; \sin\psi$ and $u = \tan\psi$ in equation (3.82) and solving, we get

$$J_\theta = 4\alpha_\theta \int_0^\infty \left(\frac{1}{1+u^2} - \frac{\cos^2\gamma}{1+u^2\cos^2\gamma} \right) du = 2\pi(\alpha_\theta - \alpha_\phi)$$

$$J_\theta + J_\phi = 2\pi\alpha_\theta \tag{3.83}$$

$$J_r = \oint \sqrt{2mE - 2mV(r) - \frac{(J_\theta + J_\phi)^2}{4\pi^2 r^2}} \, dr \tag{3.84}$$

Equation (3.84) implies that $J_r = J_r(E, J_\theta + J_\phi)$ is a two-variable function for $V = V(r)$ and for $E = E(J_r, J_\theta + J_\phi)$

$$\frac{\partial E}{\partial J_\theta} = \frac{\partial E}{\partial J_\phi}$$

This implies $v_\theta = v_\phi$. (3.85)

That shows two frequencies are degenerate for any $V = V(r)$.

When $V(r) = -kr^{(-1)}$, for $E < 0$

$$J_r = -(J_\theta + J_\phi) + \pi k \sqrt{\frac{2m}{-E}}$$

This implies

$$J_r = -(J_\theta + J_\phi) + \pi k \sqrt{\frac{2m}{-E}}$$

$$J_r + J_\theta + J_\phi = \pi k \sqrt{\frac{2m}{-E}}$$

$$E = \frac{-2\pi^2 k^2 m}{(J_r + J_\theta + J_\phi)^2} \qquad (3.86)$$

$$v_\theta = v_\phi = v_r \qquad (3.87)$$

$$v_r = \frac{\partial E}{\partial J_r} = 4\pi^2 k^2 (J_r + J_\theta + J_\phi)^{-3} = \frac{1}{\pi k} \sqrt{\frac{-2E^3}{m}}$$

This gives the orbital frequency in bound Kepler orbit.

Using the relations $\{\alpha_1 = E, \alpha_\theta, \alpha_\phi\}$ and $\{J_r, J_\theta, J_\phi\}$, the Hamilton's characteristic function can be written for this system as below:

$$W = W_r + W_\theta + W_\phi$$

$$W = \frac{\phi J_\phi}{2\pi} \pm \int \sqrt{(J_\theta + J_\phi)^2 - \frac{J_\phi^2}{\sin^2\theta}} \frac{d\theta}{2\pi} \pm \int \sqrt{\frac{-(2\pi mk)^2}{(J_r + J_\theta + J_\phi)^2} + \frac{2mk}{r} - \frac{(J_\theta + J_\phi)^2}{2\pi r}} \, dr$$

The angle variables can be given as

$$\omega_r = \frac{\partial w}{\partial J_r} = \omega_r(r, J_r + J_\theta + J_\phi, J_\theta + J_\phi)$$

$$\omega_\theta = \frac{\partial w}{\partial J_\theta} = \omega_\theta(r, \theta, J_r + J_\theta + J_\phi, J_\theta + J_\phi, J_\phi)$$

$$\omega_\phi = \frac{\partial w}{\partial J_\phi} = \omega_\phi(r, \theta, \phi, J_r + J_\theta + J_\phi, J_\theta + J_\phi, J_\phi)$$

where

$$\dot{\omega}_r = v_r, \quad \dot{\omega}_\theta = v_\theta, \text{ and } \dot{\omega}_\phi = v_\phi \text{ and } v_\theta = v_\phi = v_r.$$

It can be noted that, as per prediction, J_θ and J_ϕ exist in combination as $J_\theta + J_\phi$ only. In addition, the three action variables occur in $J_r + J_\theta + J_\phi$ form only. Hence, all of the frequencies are equal, that is the motion is completely degenerate. This result also agrees with the fact that for negative energies, the orbit is closed with an inverse-square law of force. And the motion in a closed orbit is periodic in nature and hence, degenerate. If the central force has a term with r-3, like is given by first order relativistic corrections, then the orbit is in the form of a precessing ellipse and is no longer closed. In such a case, one of the degeneracies will vanish. However, the motion will still remain singly degenerate, since vθ=vϕ for all central forces.

The degeneracy conditions can be expressed as

$$v_\phi - v_\theta = 0 \text{ and } v_\theta - v_r = 0$$

And the generating function can be written as

$$G = (\omega_\phi - \omega_\theta)J_1 + (\omega_\theta - \omega_r)J_2 + \omega_r J_3$$

The new angle variables are

$$\omega_1 = \omega_\phi - \omega_\theta$$

$$\omega_2 = \omega_\theta - \omega_r$$

$$\omega_3 = \omega_r$$

As two among the new frequencies, v_1 and v_2 are 0, the new action variables can be obtained from the transformation equations

$$J_\phi = J_1$$

$$J_\theta = J_2 - J_1$$

$$J_r = J_3 - J_2$$

Thus,

$$J_1 = J_\phi$$

$$J_2 = J_\theta + J_\phi$$

$$J_3 = J_r + J_\theta + J_\phi$$

And the Hamiltonian can be written as

$$H = -\frac{2\pi^2 m k^2}{J_3^2}$$

These constants can also be derived from angular momentum L, energy E, and Laplace-Runge-Lenz vector A.)

The action-angle variable treatment of the Kepler problem also leads to five algebraic constants of the motion. Among these five constants, two constants represent orbit plane (xy-plane), which are the inclination angle i and the longitude of the ascending node Ω. There are three constants

specifying the form of the ellipse, which are the semi-major axis a, eccentricity ε, and angle ω. These are all shown in Figure given below.
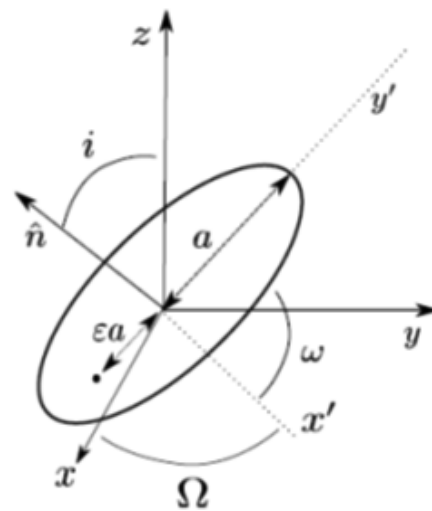
Fig. 3.2 *Schematic diagram of the orbit with the five parameters providing the specifications.*

The relation among these constants can be given as follows:

$$\cos i = \frac{J_1}{J_2}$$

$$a = -\frac{k}{2E} = \frac{J_3^2}{4\pi^2 mk}$$

$$\epsilon = \sqrt{1 - \left(\frac{J_2}{J_3}\right)^2}$$

$$\Omega = 2\pi\omega_1$$

$$\omega = 2\pi\omega_2$$

Which provide a proper physical interpretation to $(J_1, J_2, J_3, \omega_1, \omega_2)$. Also $J_2 = 2\pi\alpha_\theta = 2\pi l$

When any additional forces (such as some planet, some relativistic corrections, etc.) perturb the orbit, the modified orbit can be described by these action-angle variables.

1. What is the Hamiltonian for the system which involves time explicitly?
2. What are the Hamilton's canonical equations?
3. Which equation is referred to as the Hamilton-Jacobi equation?
4. What will be the Hamiltonian function of the one-dimensional harmonic oscillator?
5. Write Hamilton's principal function of the oscillator.
6. Write the equation of motion for the freely falling body.
7. Define Jacobi's identity.

3.5 ANSWERS TO 'CHECK YOUR PROGRESS'

1. Consider a mechanical system of s degrees of freedom described by the generalized coordinates q1, ..., qs and generalized momenta p1, ...,ps. The Hamiltonian for the system which involves time explicitly is H = H(q, p, t) = H(q1, ..., qs, p1, ..., ps, t)

2. The Hamilton's canonical equations are

$$\dot{q}_k = \frac{\partial H}{\partial p_k}; \quad \text{and} \quad \dot{q}_k = -\frac{\partial H}{\partial q_k}$$

3. $$H\left(q_1, \dots, q_s, \frac{\partial F}{\partial q_1}, \dots, \frac{\partial F}{\partial q_s}, t\right) + \frac{\partial F}{\partial t} = 0$$

4. Consider a one-dimensional harmonic oscillator of mass m. Let q and p be respectively the coordinate and momentum of the oscillator at any instant of time t.

   The Hamiltonian function H of the oscillator is

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2$$

5. We may write Hamilton's principal function of the oscillator as

$$S = m\omega \int \sqrt{q_o^2 - q^2} \, dq - \frac{m\omega^2 q_o^2 t}{2}$$

6. $$z = z_o - \frac{1}{2}gt^2$$

7. If f, g and k are any three functions of the coordinates (q's) and the momenta (p's) describing a mechanical system, then the following relation holds

[f,[g, k]] + [g, [k, f]] + [k, [f, g]] = 0

The relation expressed above is known as Jacobi's identity.

3.6 SUMMARY

- The Hamilton's canonical equations are

$$\dot{q_k} = \frac{\partial H}{\partial p_k}; \quad \text{and} \quad \dot{q_k} = -\frac{\partial H}{\partial q_k}$$

- The Hamilton–Jacobi equation is a first-order partial differential equation in (s + 1) variables, namely, q1, ..., qs, t.

- We write the general solution of the Hamilton–Jacobi equation as

$$S = S(q_1, ..., q_s, \alpha_1, ..., \alpha_s, t) + \alpha$$

- The Hamiltonian of the system then has no explicit dependence on time and is a constant of motion representing the total energy E of the system.

- We may write Hamilton's principal function of the oscillator as

$$S = m\omega \int \sqrt{q_o^2 - q^2}\, dq - \frac{m\omega^2 q_o^2 t}{2}$$

- The action-angle variable treatment of the Kepler problem also leads to five algebraic constants of the motion. Among these five constants, two constants represent orbit plane (xy-plane), which are the inclination angle i and the longitude of the ascending node Ω. There are three constants specifying the form of the ellipse, which are the semi-major axis a, eccentricity ε, and angle ω.

3.7 KEY TERMS

- **Jacobi's identity:** It is a characteristic of a binary operation which explains how the order of evaluation influences the result of the operation.

- **Canonical coordinates:** The sets of coordinates on phase space which can be used to explain a physical system at any given point in time are called Canonical coordinates.

- **Action-angle coordinates:** They are a set of canonical coordinates useful in solving many integrable systems. The method of action-angles is useful for obtaining the frequencies of oscillatory motion without solving the equations of motion.

- **Poisson bracket:** It is an important binary operation in Hamiltonian mechanics which governs the time evolution of a Hamiltonian dynamical system.

3.8 SELF-ASSESSMENT QUESTIONS AND EXERCISES

Short Answer Questions

1. State the method for solving a mechanical problem using Hamilton-Jacobi method.
2. Write in brief about the Hamilton-Jacobi equation.
3. Write a short note on Hamilton's characteristic function.
4. Mention the application of action angle variable to obtain the frequency of a linear harmonic oscillator.

Long Answer Questions

1. Describe Lagrange's equations for conservative systems.
2. Discuss Hamilton's principal function.
3. Give solution of one dimensional harmonic oscillator problem using Hamilton-Jacobi method.
4. Explain motion of a body falling freely under gravity.
5. Give a detailed account of action and angle variables.

3.9 FURTHER READING

Rao, K. Sankara. 2009. Classical Mechanics. New Delhi: PHI Learning Private Limited.

Upadhyaya, J.C. 2010. Classical Mechanics, 2nd Edition. New Delhi: Himalaya Publishing House.

Goldstein, Herbert. 2011. Classical Mechanics, 3rd Edition. New Delhi: Pearson Education India.

Gupta, S.L. 1970. Classical Mechanics. New Delhi: Meenakshi Prakashan.

Takwala, R.G. and P.S. Puranik. 1980. New Delhi: Tata McGraw Hill Publishing.

UNIT 4   CLASSICAL STATISTICAL MECHANICS

Structure

4.0   INTRODUCTION

From the earlier discussion, we can conclude that classical mechanics is an approximation of quantum mechanics. Classical mechanics cannot explain the stability of an atom. According to classical idea, an orbital electron continuously revolves and keeps nucleons in the centre (nucleus having positively charged) with a strong attractive force and emits energy in the form of electromagnetic radiation (Rutherford's model). Due to this continuous emission of energy, the orbital electrons should come closer to the nucleus, once they lose all their energy ultimately to collapse into the nucleus. So, an atom should be unstable but in reality we do not see that happening. So, what accords stability to the atom? The answer is provided by the quantum idea.

In the introduction, we have talked about the discrete energy emission from gaseous atom, which cannot be described and interpreted by the classical ideas. It is obvious to introduce quantum ideas to explain the several behaviours of atoms and subatomic particles (microscopic size particles).

4.1   OBJECTIVES

After going through this unit, you will be able to:

- Understand phase space, configuration space, hodograph, and Liouville's theorem

- Discuss the Maxwell Boltzmann Distribution Law of Velocity

- Describe various types of ensembles and thermodynamic functions in different ensembles along with canonical partition function

4.2   A PRIORI PROBABILITY

The first postulate of statistical mechanics tells us the frequency of each of the possible states or microstates occurring in the ensemble. This postulate is often called the principle of equal a priori probabilities. It says that if the microstates have the same energy, volume, and number of particles, then they occur with equal frequency in the ensemble. An isolated system, with N, V, E, has equal probability to be in any of the W(N,V,E) quantum states or Each and every one of the W(N,V,E) quantum states is represented with equally probability.

By priori, we mean something which exists in our mind prior to its actual occurrence and independent of experience. Let us consider an example of a box having two compartments of equal size by means of a partition.
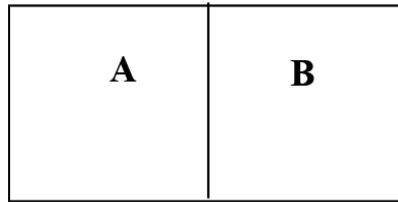
If we throw a coin from a certain distance, the probability of the coin to enter into compartment A is equal to the probability of the coin to enter into compartment B.

So **the principal of assuming equal probabilities of equal events is known as principal of equal a priori probability**.

Classically a priori probability is defined as the ratio of cases favoring the event to the total number of cases in the sample space.

It can be given by the relation:

$$P = \frac{\text{cases favoring the event}}{\text{total number of cases in the sample space}} = \frac{n}{N}$$

Example 4.1: A priori probability of rolling the number 1, 3, or 5, in a dice roll.

Total number of cases in the sample space are 6 and the cases favoring the event are 3 (1, 3, or 5). Therefore, a priori probability is 3/6 = 0.5.

Example 4.2: A priori probability of drawing a king of hearts in a deck of cards.

Total number of cases in the sample space are 52 and the cases favoring the event are 1 (a king of hearts). Therefore, a priori probability is 1/52

4.3   PHASE SPACE AND LIOUVILLE'S THEOREM

In the Lagrangian formulation of mechanics, for describing the motion of a system having s degrees of freedom, the system at any instant of time t

is represented by a point in an abstract *s*-dimensional mathematical space called the **configuration space** of the system. The point is called the **system point** at the instant *t*.

As time passes, the system point moves in the configuration space and it traces out, in general, a curve that gives the trajectory or path of the system.

In the Hamiltonian formulation, *s* generalized coordinates and *s* generalized momenta are independent variables for the system. An abstract *s* dimensional mathematical space, any point of which gives the *s* momenta of the system, is called the **momentum space** of the system. Clearly, any point in the momentum space represents the state of motion of the system at some instant of time. With progress of time, the point representing the state of motion moves in the momentum space. The curve traced out by the point is called **hodograph**.

To describe a function such as the Hamiltonian function $H(q, p, t)$ for the system we need a combination of the configuration space (coordinate space) and the momentum space for the system. Clearly, such a space is $2s$ dimensional. Such an abstract $2s$ dimensional mathematical space, any point of which represents the *s* coordinates $(q_1, \ldots, qs)$ and *s* momenta $(p_1, \ldots, ps)$ of the system at some instant of time, is called **phase space** of the system. Any point of the phase space describes not only the position of the system as a whole but also the state of motion of the system at some instant of time. As time passes, the point representing the configuration and the state of motion of the system in the phase space traces out a trajectory called the **phase trajectory**.

The concept of phase space and phase trajectory can be understood from the following example: consider a linear harmonic oscillator of mass m and oscillating along the X-axis.

The total energy of the oscillator when the displacement is *x* from the equilibrium position is given by

$$E = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\omega^2 x^2 = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 \qquad (4.1)$$

where $p = m\dot{x}$ is the momentum of the oscillator and $\omega = \sqrt{\dfrac{k}{m}}$ is the angular frequency of the oscillator.

We may express Equation (1) as

$$\frac{p^2}{2mE} + \frac{x^2}{\dfrac{2E}{m\omega^2}} = 1 \qquad (4.2)$$

In view of Equation (4.2), we find that the oscillator traverses an elliptical path in a two-dimensional space with *x* and *p* as the axes and having the semi-major axis equal to $\sqrt{\dfrac{2E}{m\omega^2}}$ and semi-minor axis equal to $\sqrt{2mE}$. In

Figure 4.1 are shown different paths of the oscillator corresponding to different energies. We observe that the phase space of the oscillator is two-dimensional and phase trajectories are ellipses of different semi-axes corresponding to different total energies.
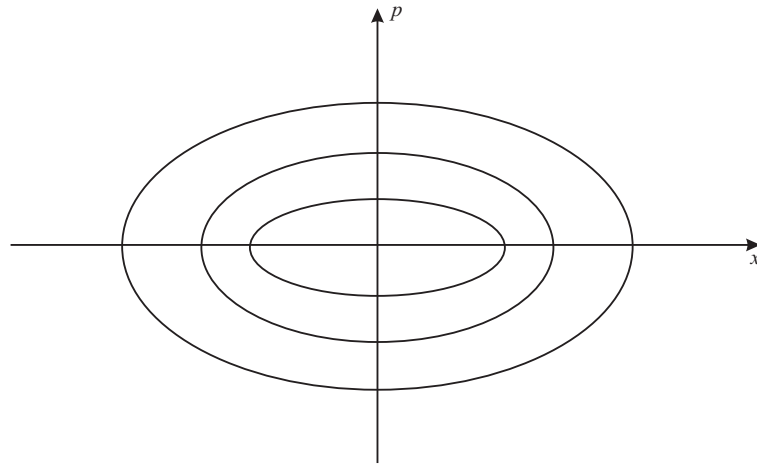
Fig. 4.1 *Paths of Oscillation*

One important feature of phase trajectory is that no two phase trajectories can intersect with each other. This can be seen as follows: Let two trajectories cross at the phase point (*xi,pi*). If we consider this point to represent the position and the momentum at $t = 0$ then there will be two possible momenta along which the motion could start. This is not possible because the solutions of the oscillator equation

$$\ddot{x} + \omega^2 x \;=\; 0$$

at any instant for $\dot{x}$ and hence for p is unique.

4.3.1 Liouville's Theorem

The dynamical state of a mechanical system at any instant of time is represented by a point in the phase space of the system which is an imaginary mathematical space of 2s dimensions if 's' be the degrees of freedom of the system. As the system develops with time, the point representing the dynamical state called the representative point traces a path or trajectory determined by the Hamilton's canonical equations given by

$$\hat{q}_i = \frac{\partial H}{\partial p_i}, \; \hat{p}_i = \frac{-\partial H}{\partial p_i} \quad i = 1, \cdots, s$$

Where H is Hamiltonian of the system given as

$$H = \;(q_i, ..., q_s, p_i, ..., p_s, t)$$

As a result of motion, the density p (number of phase point per unit volume of the phase space at a given time in the phase space) changes with time. Our interest lies in the determination of the rate at which the density changes with time at a given point in the phase space. To obtain $\frac{dp}{dt}$ we use the theorem given by Liouville. The theorem consists of two parts:

(i) The density in phase space is a conserved quantity i.e., $\dfrac{dp}{dt} = 0$.

(ii) Extension in phase space is conserved i.e., $\dfrac{d}{dt}(\delta\Gamma) = 0$. This means that the volume available to a particular number of phase points is conserved throughout the phase space.

Proof of Theorem (1)

Consider the volume of phase space located between

q_1 and $q_1 + \delta q_1$, q_2 and $q_2 + \delta q_2$,..., q_s, and $q_s + \delta q_s$,

p_1 and $p_1 + \delta p_1$ ... , p_s and $p_s + \delta p_s$

The number of phase points located in this volume i.e., in the volume $(\delta q, x \dots x \ \delta q_s) \times (\delta p_1 \dots x \ \delta p_s)$ changes as the coordinates q and momenta p vary with time.

In a time dt, the change in the number of phase points in the above volume is equal to $\left(\dfrac{\partial p}{\partial t}\right) dt \, (\delta q_1 \, x \dots x \, \delta p_s)$.

The above change is due to the number of phase points entering and leaving the volume in the time dt.

Finding out the net increase δN in the number of phase in the above volume in time dt, we get the rate of increasing of δN to be given by

$$\frac{d}{dt}(\delta N) \;=\; -\sum_{i=1}^{s}\left\{\rho\left(\frac{\partial \dot{q}_i}{\partial q_i}+\frac{\partial \dot{p}_i}{\partial p_i}\right)\left(\frac{\partial p}{\partial q_i}\dot{q}_i+\frac{\partial p}{\partial p_i}\dot{p}_i\right)\right\} dt$$
$$\underset{\times\, \delta q_1\, x\, \dots\, x\, \delta p_s}{}$$

But
$$\frac{d}{dt}(\delta N) \;=\; \frac{\partial p}{\partial t}dt \;\; \delta q_1 x \dots \delta p_s$$

equating the above two relations we obtain

$$\frac{\partial p}{\partial t}dt\, \delta q_i\, x \cdots x\, \delta p_s \;=\;$$

$$-\sum_{i=1}^{s}\left\{\rho\left(\frac{\partial \dot{q}_i}{\partial q_i}+\frac{\partial \dot{p}_i}{\partial p_i}\right)+\left(\frac{\partial p}{\partial q_i}\dot{q}_i .+\frac{\partial p}{\partial p_i}\dot{p}_i\right)\right\}\times dt$$
$$\times\, \delta q_1\, x\, \dots\, x\, \delta p_s$$

or
$$\frac{\partial p}{\partial t} \;=\; -\sum_{i=1}^{s}\left\{\rho\left(\frac{\partial \dot{q}_i}{\partial q_i}+\frac{\partial \dot{p}_i}{\partial p_i}\right)+\left(\frac{\partial p}{\partial q_i}\dot{q}_i .+\frac{\partial p}{\partial p_i}\dot{p}_i\right)\right\}$$

We have the Hamilton's equations

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i . = \frac{-\partial H}{\partial q_i}$$

Now
$$\frac{\partial \dot{q}_i}{\partial q_i} = \frac{-\partial \dot{p}_i}{\partial p_i} = \frac{\partial^2 H}{\partial q_i \, \partial p_i}$$

Since the order of differential is immaterial (coordinates and momenta being independent variables we get

$$\sum_{i-1}^{s}\left(\frac{\partial p_i}{\partial q_i}+\frac{\partial \dot{p}_i}{\partial p_i}\right)=0$$

Substituting the above in the expression for $\dfrac{\partial p}{\partial t}$ we get

$$\left(\frac{\partial p}{\partial t}\right)_{q,p} + \sum_{i=1}^{s} \frac{\partial p}{\partial q_i} \frac{dq_i}{dt} + \sum_{i=1}^{s} \frac{\partial p}{\partial p_i} \frac{dp_i}{dt} = 0$$

The above equation is identical with this equation of continuity in hydrodynamics. Considering $\rho = \rho(q, p, t)$ we get

$$\frac{dp}{dt} = \frac{\partial p}{\partial t} + \sum_{i=1}^{s} \frac{\partial p}{\partial q_i} \dot{q}_i + \sum_{i=1}^{s} \frac{\partial p}{\partial p_i} \dot{p}_i$$

Comparing the above two equations we get

$$\frac{dp}{dt} = 0$$

We thus find that the density in phase space is conserved.

Proof of Theorem (2)

We know that

$$\delta N = \rho \delta V$$

Taking total time derivative of the above equation we get

$$\frac{d}{dt}(\delta N) = \frac{dp}{dt} \delta V + \rho \frac{d}{dt}(\delta V)$$

Since the number of phase points in a given region of the phase space is invariant i.e., phase points are neither created nor destroyed, we get

$$\frac{d}{dt}(\delta N) = 0$$

Hence we get

$$\frac{dp}{dt} \delta V + \rho \frac{d}{dt}(\delta V) = 0$$

But $\quad \dfrac{dp}{dt} = 0$ (Liouville's theorem (1)) and hence

$$\rho \frac{d}{dt}(\delta V) = 0$$

Since $\rho \neq 0$ we finally get

$$\frac{d}{dt}(\delta V) = 0$$

The above result represents conservation of extension in phase space.

4.4 STATISTICAL EQUILIBRIUM

In some situations, the resultant of several concurrent forces may be zero. In such a case, the net effect of all the forces acting on a particle will be zero

and the particle is said to be in equilibrium.

In other words, a body is said to be in equilibrium, when the resultant of the force system acting on it is zero. If a body is in equilibrium, it will continue to remain in a state of rest or of uniform motion.

Equilibrant

According to Newton's second law of motion, a body starts moving with uniform acceleration if it is acted upon by a force. When a body is subjected to a number of concurrent forces, it moves in the direction of resultant force with uniform acceleration. However, if another force which is equal in magnitude of the resultant but opposite in direction is applied to the body, the body comes to rest. Hence equilibrant of a system of forces is a single force which acts along with the other forces to keep the body in equilibrium.

(a) Given system of forces   (b) Resultant (R)   (c) Equilibrant (E)
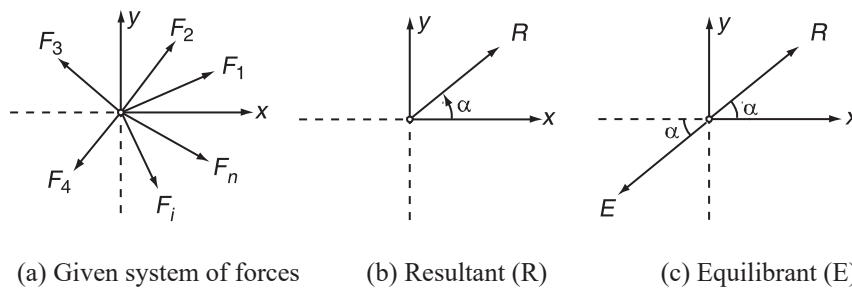
Fig. 4.2

Example 4.3: Three forces act on a particle 'O' as shown in Fig. 4.3(a). Determine the value of 'F' such that the resultant of these three forces is horizontal. Find the magnitude and direction of the fourth force which when acting along with the given three forces will keep 'O' in equilibrium.
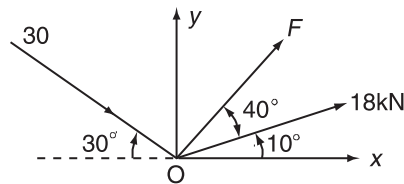
Fig. 4.3(a)

Solution: Angle between 'F' and Ox is 40º + 10º = 50º and 30 kN is acting towards the particle. Since the resultant is horizontal; $\Sigma F_y = 0$.

$\Sigma F_y = 0$ givess,

$$-30 \sin 30° + 18 \sin 10° + F \sin 50° = 0$$

$F = 15.5$.

$\Sigma F_x = 30 \cos 30° + F = \cos 50° + 18 \cos 10° = 53.67$

Equilibrant is -53.67 kN, as shown in Fig. 4.3(b)
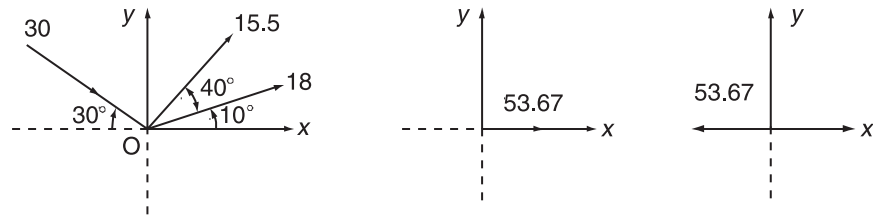
Fig. 4.3(b)

Equation and Equilibrium

The resultant of given system of forces is expressed as,

$$\vec{R} = \Sigma \vec{F}_i \qquad ...(4.)$$

$$= (\Sigma F_x)\vec{i} + (\Sigma F_y)\vec{j} \qquad ...(4.)$$

When a particle is in equilibrium, the resultant force is zero. In other words,

$$\vec{R} = 0 \qquad ...(4.)$$

$$\Sigma F_x = 0 \text{ and} \qquad ...(4.)$$

$$\Sigma F_y = 0 \qquad ...(4.)$$

The x and y axes can be chosen arbitrarily through the point of concurrency. However, for problems involving bodies on inclined planes, the axes are selected along tangential and normal directions to the plane. It is also observed that in the case of problems involving equilibrium of particle, since total number of equilibrium equations are two, the number of unknowns can be only two. Determination of a force with magnitude and direction or magnitudes of two forces, whose directions are known also can be solved. These concepts are illustrated in numerical examples.

Check Your Progress

1. What do you mean by priori probability?

2. Define configuration space.

3. What is momentum space?

4. What do you understand by hodograph?

5. What is phase trajectory?

6. When is a body said to be in equilibrium?

4.5 MAXWELL BOLTZMANN DISTRIBUTION LAW OF VELOCITY

The basic postulates of MB statistics are:

 (i) The associated particles are distinguishable.

 (ii) Each energy state can contain any number of praticles.

 (iii) Total number of particles in the entire system is constant.

 (iv) Total energy of all the particles in the entire system is constant.

(v) Particles are spinless.

Examples: Gas molecules at high temperature and low pressure.

Maxwell-Boltzmann Distribution Law

Consider a system composed of N distinguishable, non-interacting particles. Let out of these N particles N_1, N_2, ..., N_i particles are to be distributed in energy levels E_1, E_2, ..., E_i respectively and these energy levels have g_1, g_2, ..., g_i, number of quantum states correspondingly. Since the total energy E and total number of particles N are constant for the system, we can write

$$\sum_i N_i E_i = E \qquad \qquad ...(4.3)$$

$$\sum_i N_i = N \qquad \qquad ...(4.4)$$

The number of ways in which the groups of particles N_1, N_2, ..., N_i could be chosen from N particles is

$$W_1 = \frac{N!}{N_1! N_2! ... N_i!} = \frac{N!}{\prod\limits_i^n N_i!} \qquad \qquad ...(4.5)$$

where Π denotes the product.

Now, N_i particles can be distributed in g_i state in $(g_i)N_i$ ways. Considering all the values of i, total number of arrangement would be

$$W_2 = \prod\limits_i^n (g_i)N_i \qquad \qquad ...(4.6)$$

Therefore, the total number of ways W by which all the N particles could be distributed among the quantum states is

$$W = W_1 W_2 = \frac{N!}{\prod\limits_i^n N_i!} \prod\limits_i^n (g_i)N_i \qquad \qquad ...(4.7)$$

which is the Maxwell-Boltzmann distribution law for n distinguishable particles.

Now taking the natural logarithm on both sides of eqn. (4.7), we get,

$$\ln W = \ln N! + \sum_i N_i \ln g_i - \sum_i \ln N_i! \qquad \qquad ...(4.8)$$

Applying Stirling approximation (i.e., $\ln x! = x \ln x - x$, where x is very large), we get from eqn. (4.8)

$$\ln W = N \ln N - N + \sum_i N_i \ln g_i - \left[\sum_i (N_i \ln N_i - N_i) \right]$$

$$= N \ln N - \sum_i N_i \ln \frac{N_i}{g_i} \qquad \qquad ...(4.9)$$

Now differentiating both sides, we get,

$$d(\ln W) = d(N \ln N) - \sum_i dN_i \ln \frac{N_i}{g_i} - \sum_i N_i \, d\left(\ln \frac{N_i}{g_i} \right)$$

$$= -\sum_i dN_i \ln \frac{N_i}{g_i} - \sum_i N_i \frac{dN_i}{N_i}$$

$$= -\sum_i dN_i \ln \frac{N_i}{g_i} - \sum_i dN_i$$

$$= -\sum_i dN_i \ln \frac{N_i}{g_i} \qquad \qquad ...(4.10)$$

$$\left[\because \sum_i N_i = \text{constant}, \therefore \sum_i dN_i = 0 \right]$$

For the most probable distribution, $d(\text{in } W)_{max} = 0$

Therefore $\qquad \sum_i dN_i \ln \frac{N_i}{g_i} = 0 \qquad \qquad ...(4.11)$

Since the system is in equilibrium, total number of particle and the total energy of the system are constant. So,

$$\sum_i dN_i = 0 \qquad \qquad ...(4.12)$$

and $\qquad \sum_i E_i \, dN_i = 0 \qquad \qquad ...(4.13)$

Multiplying eqn. (4.12) by α and eqn. (4.13) by b and then adding to eqn. (4.11), we get,

$$\sum_i \left[\ln \left(\frac{N_i}{g_i} \right) + \alpha + \beta E_i \right] dN_i = 0 \qquad \qquad ...(4.14)$$

Since dN_i's are independent of one another, the above equation hold only if,

$$\ln \left(\frac{N_i}{g_i} \right) + \alpha + \beta E_i = 0 \text{ , or, } N_i = g_i e^{-(\alpha + \beta E_i)}$$

Now the Maxwell-Boltzmann distribution function is given by

$$f(E_i) = \frac{N_i}{g_i} = \frac{1}{e^{(\alpha + \beta E_i)}} \qquad \qquad ...(4.15)$$

which physically gives the probability of a particle to occupy the energy state E_i.

4.6 EQUATION OF ENERGY, TYPES OF ENSEMBLES AND THERMODYNAMIC FUNCTIONS IN DIFFERENT ENSEMBLES

In this section, we will discuss the three ensembles of statistical mechanics, the microcanonical ensemble, the canonical ensemble and the grand canonical ensemble. Here canonical means simply standard or acceptable and the canonical ensemble therefore holds the central place in statistical mechanics. Logically the canonical ensemble should be introduced first. The microcanonical and grand canonical ensembles then follow as a special case (all systems having identical energies) and an extension (to systems having variable number of particles) of the canonical ensemble, respectively.

4.6.1 Micro Canonical Ensemble

A system (a solid, liquid or a gas) which is completely isolated from its surroundings has constant energy U and a constant number of particles N. We will suppose it is also contained at constant volume V. If the system has a number of different types of particles (a number of components), the number of each component, $N1; N2 ; : : :$ is also constant. A micro-canonical ensemble is an assembly of mental copies of this isolated system. Since the energy of each copy is the same, $E = U$, the need to consider an ensemble is really superfluous.

One system will do the work.

Since all the copies have the same energy, U, the probability of observing the system with energy E_S is

$$P_S = \text{constant} \times \ \delta \left(E_S - U\right).$$

$$(4.16)$$

If the quantum energy state U of the system is degenerate with degeneracy W, then there are W states having energy U or equivalently W ways of forming the observed system. Using the equal a priori probability hypothesis and

$$\sum_{S=1}^{W} P_S = 1,$$

$$(4.17)$$

we must have

$$P_S = \frac{1}{W} .$$

This statistical probability and thermodynamics are related through the famous Boltzmann relation,

$$S = k \log W$$

$$(4.18)$$

This could be introduced as a hypothesis but it follows from the relation between S and PS developed in the next section for the central canonical ensemble.

$$S = -k \sum_{S} P_S \ \log \ P_S$$

$$= -k \, W \left(\frac{1}{W} \ \log \ \frac{1}{W}\right)$$

$$= k \ \log \ W \ (U, \ V, \ N)$$

Thus, as expected, the appropriate thermodynamic function is $S(U; V; N)$ which is a maximum at equilibrium for an isolated system. This also demonstrates that Boltzmann's relation (4.18) applies to an isolated system.

Since isolated systems are difficult to realize in practice, the microcanonical ensemble is not often used.

4.6.2 The Canonical Ensemble

We consider now a system (a solid, liquid or a gas) in contact with a heat bath at constant temperature. A canonical ensemble is an assembly of mental copies of this system. Since energy can be transferred between the system and the heat bath, the energy of the systems in the assembly differ. To represent the possible energy states of the system adequately in the ensemble, we must have enough copies that each state is represented at least once in the ensemble.

All copies have the same temperature T, the temperature of the heat bath. We could imagine constructing the canonical ensemble from the single system in the microcanonical ensemble in the following way. We divide the isolated system into many (say 10^9) parts. If the original system contained 1025 particles each part is still large enough to represent the macroscopic properties of the original system. We take one part as the system itself and regard the remaining parts as making up the mental copies in the canonical ensemble. We may also regard the one part as making up the system and the remaining parts as making up a heat bath. Since the new system is in good thermal contact with the remaining (surrounding) copies, there is exchange of heat between system and the heat bath. The exchange of heat serves to keep the system at constant T.

This canonical ensemble is the Gibbs canonical ensemble. Using the Boltzmann combinatorial method to find the probability of observing the system $P_S = \dfrac{1}{Z} e^{-\beta E_S}$ in this ensemble we found

$$(4.19)$$

where $Z = \sum_S e^{-\beta E_S}$ is the canonical partition function.

The canonical ensemble is generally the most useful in practice since we most often deal with systems in thermal equilibrium (constant T) with their surroundings. The energy states fluctuate and the probability of observing the system in a given energy state at constant T is given by (4.19). We also see that the microcanonical ensemble is a special case of the canonical ensemble in which all the systems have the same energy.

4.6.3 The Grand Canonical Ensemble and Canonical Partition Function

Here we consider an open system in contact with a heat and particle bath. The system is open meaning that particles as well as heat can be exchanged between the bath and the system. An example is a solid or liquid (the system) in contact with its vapor (the bath) in equilibrium so that particles are freely exchanged between the liquid and the vapor. Since the vapor and liquid are in equilibrium, the chemical potential, μ, is the same in each (and particles are exchanged to maintain μ the same in each phase) The grand canonical ensemble is an assembly of mental copies of this open system. In the ensemble all possible states of the system are represented; that is, all possible values of N and all possible energy states ES (N) for a given N.

We sub-divide the closed system having N particles in Fig. 6.1 into two parts; one small part having Na particles in volume Va and the second having Nb

particles in the remaining volume, Vb. We choose Na << Nb = N − Na . The small sub-system is our open system and particles can be freely exchanged between it and the remaining N −Na particles. The ensemble is then an assembly of mental copies of this open sub-system.

To develop the statistics for the open system we recall that the probability of observing the closed system of N particles in energy state ES is from (4.19)

$$P_S = Z^{-1} e^{-\beta E_S} = e^{\beta(F - E_S)}$$

Assuming weak mechanical contact between the sub-system (N_a particles) and the remainder of the system (N_b particles) we can write

$$E_S = E_{Sa}(N_a) + E_{Sb}(N_b)$$

and

$$F(T,\ V,\ N) = F_a(T,\ V_a,\ N_a) + F_b(T,\ V_b,\ N_b)$$

so that

$$P_S = e^{-\beta E_{Sa}(N_a)} \times e^{-\beta E_{Sb}(N_b)} \times e^{\beta F(T,\ V,\ N)}$$

The probability $P_{Sa}(N_a)$ of observing the sub-system in energy state $E_{Sa}(N_a)$ for any given state of the system b is

$$P_{Sa}(N_a) = \sum_{Sb} P_S(N) = e^{-\beta E_{Sa}(N_a)} e^{\beta(F - F_b)}$$

where as usual

$$F_b = -kT \ \log \sum_{Sb} e^{-\beta E_{Sb}(N_b)}$$

Since $N_a \ll N - N_b$ we may write

$$F_a = F - F_b = \left(\frac{\partial F}{\partial N}\right) N_a + \left(\frac{\partial F}{\partial V}\right) V_a$$

$$= \mu N_a - p_a V_a$$

Thus

$$P_{Sa}(N_a) = e^{-\beta p_a V_a} \ e^{-\beta(E_{Sa}(N_a) - \mu N_a)} \tag{4.20}$$

We have the central result that the probability of observing the open sub-system having N_a particles and in energy state $E_{Sa}(N_a)$ is proportional to exp $(-\beta(E_S(N_a) - \mu N_a))$. We now focus our attention entirely on the sub-system and drop the subscript a and write this probability as

$$P_S(N) \propto e^{-\beta(E_S(N) - \mu N)}$$

We also define the grand canonical partition function by

$$\mathcal{Z} = \sum_{S,N} e^{-\beta(E_S(N) - \mu N)} \tag{4.21}$$

Since we want the probability to be normalized so that

$$\sum_{S,N} P_S(N) = 1 \tag{4.22}$$

we have, from (4.20) and (4.21),

$$\sum_{S,N} P_S(N) = 1 = e^{-\beta PV} \mathcal{Z} \tag{4.23}$$

This gives

$$P_S(N) = \mathcal{Z}^{-1} \, e^{-\beta(E_S(N) - \mu N)} \tag{4.24}$$

and our relation between thermodynamics and statistics for open systems as

$$pV = kT \, \log \, \mathcal{Z} \, (T, \, V, \, \mu) \, . \tag{4.25}$$

The thermo-dynamic potential is given as,

$$\Omega \, (T, \, V, \, \mu) = F \, (T, \, V, \, N) - \mu \overline{N} \tag{4.26}$$

In this relation μ is

$$\mu = \left(\frac{\partial F}{\partial N} \right)_{T, V} = \left(\frac{\partial G}{\partial N} \right)_{T, P}$$

Also μ is the Gibbs Free energy per particle, $G = \mu \overline{N}$ The thermodynamic potential is the appropriate function to describe an open system since T, V and μ are constant so that Ω will be constant while F, for example, can fluctuate as N fluctuates. The Ω can be directly related to $\mathcal{Z}$ using the thermodynamic relation, $G = F + pV$ so that

$$\Omega = G - pV - \mu \overline{N} = -pV$$

and

$$\Omega = -kT \, \log \, \mathcal{Z} \, (T, \, V, \, \mu) \tag{4.27a}$$

This can be used as the fundamental relation between statistics and thermody-namics for open systems. It is interesting to note that in going from a closed to an open system (at constant μ) we transferred from F to Ω using. (4.26) This gave us a function (T, V, μ) of variables T , V , and μ rather than a function Ω (T, V, N) of variables T , V , and N. In introducing the grand canonical ensemble we have effectively made the same transformation of variables by summing over N. That is

$$\mathcal{Z} \, (T, \, V, \, \mu) = \sum_N \sum_S e^{-\beta(E_S(N) - \mu N)}$$

$$= \sum_N e^{-\beta \mu N} \, Z(T, \, V, \, N) \tag{4.27b}$$

This serves to eliminate the dependence on N and introduce μ.

Using the thermodynamic potential we can obtain some useful relations for $P_S(N)$, $\overline{N}$, U and S. From (4.26)

$$\Omega = U - TS - \mu \overline{N} = \sum_{S,N} P_S(N) \, E_S(N) - TS - \mu \overline{N}$$

Hence

$$P_S(N) = \frac{\partial \Omega}{\partial E_S(N)} = -\beta^{-1} \, \frac{\partial}{\partial E_S(N)} \, \log \, \mathcal{Z} \tag{4.28}$$

and

$$\overline{N} = - \left(\frac{\partial \Omega}{\partial \mu} \right)_{T,V} = \beta^{-1} \, \frac{\partial}{\partial \mu} \, (\log \, \mathcal{Z})_{T,V} \tag{4.29}$$

Also,

$$\frac{\partial}{\partial \beta} \log \mathcal{Z} = -\mathcal{Z}^{-1} \sum_{S,N} (E_S(N) - \mu N)\, e^{-\beta(E_S(N) - \mu N)}$$

$$= -\sum_{S,N} P_S(N)\, E_S(N) + \mu \overline{N}$$

so that

$$U = -\frac{\partial}{\partial \beta}\, (\log \mathcal{Z})_{V,\mu} + \mu \overline{N} \qquad (4.30)$$

The entropy is related to the grand canonical probability function $P_S(N)$ in the given form as that is

$$S = -kT \sum_{S,N} P_S(N)\, \log\, P_S(N) = -k\, \langle \log\, P_S(N)\rangle \qquad (4.31)$$

This follows from eqs. (4.24) and (4.27) since

$$\log\, P_S(N) = -\,(E_S(N) - \mu N) - \log\, \mathcal{Z}$$

and

$$-\beta^{-1} \sum_{S,N} P_S(N)\, \log\, P_S(N) = \sum_{S,N} P_S(N)\, \big[\, E_S(N) - \mu N \,\big] - \Omega$$

$$= U - \mu \overline{N} - \Omega = TS\,.$$

Check Your Progress

7. Name the three ensembles of statistical mechanics.

8. What is considered in the grand canonical ensemble?

4.7 ANSWERS TO 'CHECK YOUR PROGRESS'

1. The first postulate of statistical mechanics tells us the frequency of each of the possible states or microstates occurring in the ensemble. This postulate is often called the principle of equal a priori probabilities. It says that if the microstates have the same energy, volume, and number of particles, then they occur with equal frequency in the ensemble.

2. In the Lagrangian formulation of mechanics, for describing the motion of a system having s degrees of freedom, the system at any instant of time t is represented by a point in an abstract s-dimensional mathematical space called the configuration space of the system. The point is called the system point at the instant t.

3. In the Hamiltonian formulation, s generalized coordinates and s generalized momenta are independent variables for the system. An abstract s dimensional mathematical space, any point of which gives the s momenta of the system, is called the momentum space of the system.

4. Any point representing the state of motion moves in the momentum space. The curve traced out by the point is called hodograph.

5. As time passes, the point representing the configuration and the state of motion of the system in the phase space traces out a trajectory called the phase trajectory.

6. A body is said to be in equilibrium when the resultant of the force system acting on it is zero.

7. The three ensembles of statistical mechanics are, the microcanonical ensemble, the canonical ensemble and the grand canonical ensemble,

8. An open system in contact with a heat and particle bath is considered in the grand canonical ensemble.

4.8 SUMMARY

- To describe a function such as the Hamiltonian function $H(q, p, t)$ for the system we need a combination of the configuration space (coordinate space) and the momentum space for the system.

- One important feature of phase trajectory is that no two phase trajectories can intersect with each other.

- As the system develops with time, the point representing the dynamical state called the representative point traces a path or trajectory determined by the Hamilton's canonical equations given by

$$\hat{q}_i = \frac{\partial H}{\partial p_i}, \hat{p}_i = \frac{-\partial H}{\partial p_i} \quad i = 1, \cdots, s$$

- The first postulate of statistical mechanics tells us the frequency of each of the possible states or microstates occurring in the ensemble. This postulate is often called the principle of equal a priori probabilities. It says that if the microstates have the same energy, volume, and number of particles, then they occur with equal frequency in the ensemble.

- The various methods of statistical mechanics are applied to discuss some average or most probable properties of large assemblies of electrons, atoms, molecules, etc. Before the advent of quantum mechanics, Maxwell, Boltzmann, Gibbs etc., applied statistical methods with the help of classical physics. These methods are collectively known as Classical Statistics or Maxwell-Boltzmann (MB) Statistics.

- Canonical means simply standard or acceptable and the canonical ensemble therefore holds the central place in statistical mechanics.

- A system (a solid, liquid or a gas) which is completely isolated from its surroundings has constant energy U and a constant number of particles N. We will suppose it is also contained at constant volume V. If the system has a number of different types of particles (a number of components), the number of each component, $N1; N2 ; : : :$ is

also constant. A micro-canonical ensemble is an assembly of mental copies of this isolated system.

- If the quantum energy state U of the system is degenerate with degeneracy W, then there are W states having energy U or equivalently W ways of forming the observed system.

- The canonical ensemble is generally the most useful in practice since we most often deal with systems in thermal equilibrium (constant T) with their surroundings. The energy states fluctuate

4.9 KEY TERMS

- **Configuration space:** The vector space explained by generalized coordinates is known as the configuration space of the physical system.

- **Hodograph:** The curve traced out by any point in the momentum space representing the state of motion is called hodograph.

- **Ensemble:** In physics, specifically statistical mechanics, an ensemble is an idealization consisting of a large number of virtual copies of a system, considered all at once, each of which represents a possible state that the real system might be in.

4.10 SELF-ASSESSMENT QUESTIONS AND EXERCISES

Short Answer Questions

1. Write in brief about priori probability.
2. What do you mean by statistical equilibrium?
3. State the basic postulates of MB statistics.

Long Answer Questions

1. Discuss Liouville's theorem.
2. Explain Maxwell-Boltzmann Distribution Law.
3. Describe the different types of ensembles of statistical mechanics.

4.11 FURTHER READING

Rao, K. Sankara. 2009. *Classical Mechanics*. New Delhi: PHI Learning Private Limited.

Upadhyaya, J.C. 2010. *Classical Mechanics,* 2nd Edition. New Delhi: Himalaya Publishing House.

Goldstein, Herbert. 2011. *Classical Mechanics,* 3rd Edition. New Delhi: Pearson Education India.

Gupta, S.L. 1970. *Classical Mechanics.* New Delhi: Meenakshi Prakashan.

Takwala, R.G. and P.S. Puranik. 1980. New Delhi: Tata McGraw Hill Publishing.

UNIT 5   QUANTUM STATISTICAL MECHANICS

Structure

5.0   INTRODUCTION

Quantum Mechanics (QM); also known as quantum physics, quantum theory, the wave mechanical model, or matrix mechanics, quantum field theory; is a fundamental theory in physics which describes nature at the smallest scales of energy levels of atoms and subatomic particles. In the mathematically rigorous formulation of quantum mechanics developed by Paul Dirac, David Hilbert, John von Neumann, and Hermann Weyl, the possible states of a quantum mechanical system are symbolized as unit vectors (called state vectors).

In order to understand the origin of quantum physics and the subsequent development of an altogether new and conceptually different mathematical theory of quantum mechanics, it is first of all necessary to understand the phenomena at micro-level that what was happening at the atomic and subatomic levels. The new aspects of nature and phenomena that were revealed at these levels are referred to as quantum phenomena, the word 'quantum' referring to peculiar aspects of nature that go against common sense. The study of quantum phenomena has come to be known as quantum physics.

The Bose–Einstein (B–E) statistics explain one of two methods in which a group of non-interacting, indistinguishable particles might inhabit

a set of discrete energy levels at thermodynamic equilibrium. The cohesive streaming of laser light and the frictionless crawling of superfluid helium are explained by the aggregation of particles in the same state, which is a property of particles following Bose–Einstein statistics. Satyendra Nath Bose created the theory of this behaviour (1924–25), recognising that a group of similar and indistinguishable particles can be distributed in this way. Albert Einstein, in partnership with Bose, later embraced and expanded on the concept. The Bose–Einstein statistics only apply to particles that are not restricted to single occupancy of the same state, i.e., particles that do not obey the Pauli exclusion principle. These particles are known as bosons and have integer spin values.

An ideal Bose gas is a quantum-mechanical phase of matter, analogous to a classical ideal gas. It is composed of bosons, which have an integer value of spin, and obey Bose–Einstein statistics. The statistical mechanics of bosons were developed by Satyendra Nath Bose for a photon gas, and extended to massive particles by Albert Einstein who realized that an ideal gas of bosons would form a condensate at a low enough temperature, unlike a classical ideal gas. This condensate is known as a Bose–Einstein condensate. Einstein condensation is the limit of high particle densities and low temperatures (quantum limit), where one finds important qualitative differences between bosons, fermions and classical particles. Liquid helium is a physical state of helium at very low temperatures at standard atmospheric pressures. Liquid helium may show superfluidity.

A free particle is a particle that, in some sense, is not bound by an external force, or equivalently not in a region where its potential energy varies. In classical physics, this means the particle is present in a 'Field-Free' space. In quantum mechanics, it means a region of uniform potential, usually set to zero in the region of interest since potential can be arbitrarily set to zero at any point (or surface in three dimensions) in space. By a free particle we mean a particle which moves freely in space without the influence of any force. Hence, for a free particle the potential energy is zero. Landau theory in physics is a theory that Lev Landau introduced in an attempt to formulate a general theory of continuous (i.e., second-order) phase transitions. It can also be adapted to systems under externally-applied fields, and used as a quantitative model for discontinuous (i.e., first-order) transitions. Critical exponents describe the behavior of physical quantities near continuous phase transitions. It is believed, though not proven, that they are universal, i.e., they do not depend on the details of the physical system, but only on some of its general features. In physics, mathematics and statistics, scale invariance is a feature of objects or laws that do not change if scales of length, energy, or other variables, are multiplied by a common factor, and thus represent a universality. The technical term for this transformation is a dilatation (also known as dilation), and the dilatations can also form part of a larger conformal symmetry. Dimensional analysis is a branch of mathematics which deals with dimensions of quantities. Each physical phenomenon can be expressed by an equation that represents the relationship between the variables governing the phenomenon.

In this unit, you will learn about the general concept of Quantum

Statistical (QS) mechanics, Bose-Einstein statistics, black-body radiation, pressure of an ideal Bose gas, Einstein condensation, theory of liquid helium, Fermi Dirac statistics, free electron theory of solids, Landau theory of phase transition, critical indices, dimensional analysis and density and energy function with electron spin in hydrogen like atom.

5.1 OBJECTIVES

After going through this unit, you will be able to:

- Understand the general concept of Quantum Statistical (QS) mechanics
- Analyse the Bose-Einstein statistics
- Explain the black-body radiation
- Discuss the concept of Einstein condensation
- Discuss the theory of liquid helium
- Elaborate on the Fermi Dirac statistics
- State the free electron theory of solids
- Illustrate the Landau theory of phase transition
- Discuss the concept of dimensional analysis

5.2 GENERAL CONCEPT OF QUANTUM STATISTICAL MECHANICS

In order to understand the origin of quantum physics and the subsequent development of an altogether new and conceptually different mathematical theory of quantum mechanics, it is first of all necessary to understand the background of the crisis in physics which was witnessed in the beginning of the 20th century.

Towards the end of the 19th century and the beginning of the 20th century many new discoveries took place. The discovery of X-rays in 1895, the laws of radioactivity in 1896, electron in 1897, dependence of electron's mass on its velocity, the laws of photoelectric effect, the laws of Compton effect are a few in a very imposing list of discoveries. Many new experiments, such as Franck and Hertz experiment, Davisson-Germer experiment, Thomson's experiment were performed during the period. Many new aspects of nature were encountered while dealing with physical problems in the domain of small particles, namely atoms and subatomic particles. What was astonishing was that the new discoveries, the results of new experiments and the phenomena at atomic and subatomic levels could not be understood in terms of the then existing laws of classical physics. The phenomena at microlevel were found to be quite strange and one had to lose one's common sense in order to perceive what was happening at the atomic and subatomic levels. The new aspects of nature and phenomena that were revealed at these levels are referred to as *quantum phenomena*, the word 'quantum' referring to peculiar aspects of nature that go against common sense.

The study of quantum phenomena has come to be known as quantum physics.

Like classical physics, quantum physics also has been provided with a mathematical apparatus. The entirely new conceptual structure for dynamics in particular and physics in general, has been evolved during the last century. The currently accepted structure developed by Schrödinger, Heisenberg, Max Born, Jordan, Dirac and many others to deal with problems in the microdomain, i.e., at atomic and subatomic levels is termed as quantum mechanics.

Scope of Quantum Mechanics

The laws of quantum physics that govern the elementary particles are, however, not unconcerned with the macroscopic world and instead represent generalization of classical laws including them as special cases. The laws of quantum physics have been found to be the most general laws of nature discovered so far.

We may note that just as theory of relativity extends the range of application of physical laws to the region of very high velocities and just as the universal constant of fundamental significance 'c' (speed of light in vacuum) characterizes relativity, so a universal constant of fundamental significance 'h' (Planck's constant) characterizes quantum physics which includes classical physics as a special case.

It is often said that 'Revolution' was brought about through the discovery of quantum mechanics. The word revolution suggests that something has been overturned completely. We may note that the discovery of quantum mechanics has not overturned the laws of classical physics in any way. The motion of a simple pendulum is described in the same way even today as it was done prior to the discovery of quantum mechanics. Classical ideas embodied in the laws of classical physics have their own limits of applicability. The classical theories of physics do not find universal validity in the sense that they are only good phenomenological laws and are unable to tell us everything even about macroscopic bodies. There exists no comprehensive classical theory of matter. Classical physics does not provide answers to:

- Why the densities of materials are what they are?
- Why the elastic constants have the values they have?
- Why a rod breaks if the tension in the rod exceeds a certain limit?
- Why copper melts at 1083°C?
- Why sodium vapour emits yellow light?
- Why copper conducts electricity but sulphur does not?
- Why uranium atom disintegrates spontaneously releasing energy?, etc.

We find a host of observation for which classical physics has to tell us very little or nothing at all. Besides, the facts of chemistry are not understood in terms of classical laws.

With the advent of quantum mechanics our knowledge has expanded enormously about the laws of physics in the realm of small particles which has consequently enabled us to build, if not comprehensive, at least a good theory of matter.

The theory of quantum mechanics has explained all kinds of details, such as why an oxygen atom combines with two hydrogen atoms to make one molecule of water, and so on. Quantum mechanics thus supplies the theory behind chemistry. It has been realized that fundamental theoretical chemistry is based on the theory of quantum mechanics.

5.2.1 Wave Nature of Microparticles: de-Broglie's Hypothesis

Around 1923, Louis de-Broglie suggested that the idea of duality should be extended not only to radiation but also to all microparticles. He hypothesized that just as a quantum of radiation has a wave associated with it which governs its motion in space, so also a quantity of matter has a corresponding wave (which may be called matter wave) that governs its motion in space.

The universe is essentially composed of only two entities namely matter and radiation. de-Broglie agreed that since one of the entities, namely radiation, has dual nature, the other entity matter must also exhibit dual character. His hypothesis is consistent with the symmetry principle of nature.

De-Broglie proposed to associate, with every microparticle, corpuscular characteristics namely energy E and momentum p on the one hand, and wave characteristics namely frequency ν and wavelength λ on the other hand. According to de-Broglie, the mutual dependence between the characteristics of the two kinds was accomplished, through the Planck's constant h as

$$E = h\nu \quad \text{and} \quad p = \frac{h\nu}{c} = \frac{h}{\lambda} \qquad \qquad ...(5.1)$$

This relation is known as de-Broglie's equation.

The wavelength λ of matter wave associated with a microparticle is called de-Broglie wavelength of the particle. De-Broglie's hypothesis had profound importance from the fact that relation in Equation (5.1) was assumed to be satisfied not only for photons (zero rest mass), but for all microparticles, particularly for those which possess rest mass and which were associated with corpuscles.

Confirmation of de-Broglie's Hypothesis

Walter Elsasser, for the first time in 1926, pointed out that the wave nature of matter could be tested by allowing a beam of electrons of appropriate energy to be incident on a crystalline solid in which periodic arrangement of atoms might serve as a three-dimensional array of diffracting centres for the electron wave (if it at all exists), when diffraction peaks in characteristic directions might he observed.

The above idea was confirmed experimentally by Clinton Davisson and Lester Germer in the United States and George Thomson in Scotland.

Davisson and Germer's Experiment

The experimental arrangement used by Davisson and Germer is schematically shown in the Figure 5.1.
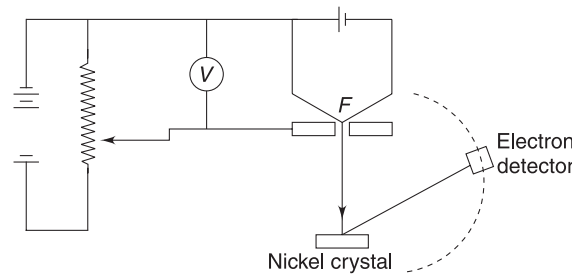
Fig. 5.1 *Davisson and Germer Experiment*

F is a filament which emits electrons when heated electrically.

The emitted electrons are accelerated through a potential V whose value can be adjusted as required by means of a potential divider arrangement. The accelerated electrons having kinetic energy E are then allowed to pass through a system of narrow slits so as to obtain a thin collimated beam of electrons. The beam of electrons thus obtained is then allowed to be incident normally on a single crystal C of nickel enclosed in a vacuum chamber. The crystal can be rotated about the incident beam as the axis. D is an electron detector which detects only elastically scattered electrons. The detector can be moved along an arc of a circle about the crystal so as to measure the intensity of elastic scattering in different directions in front of the crystal.

The intensities of the different beam at different angles λ and for different values of the accelerating potential were determined. The results obtained are shown in the Figure 5.2 and Figure 5.3. A peak in the intensity was observed at $\phi = 50°$ for $V = 54$ volts. Such an observation does not find explanation on the basis of particle motion. However, it finds explanation in terms of interference phenomenon which is characteristic of wave only.

The wavelength of electrons impinging the crystal are given by $\lambda = \dfrac{h}{p}$, according to de-Broglie's equation. We may assume Bragg reflections for electron wave to occur from certain families of atomic planes as in the case of X-ray diffraction from crystals. Bragg reflection obeying Bragg's equation is shown in the Figure 5.4. Bragg's equation is given by:

$$2d \sin \theta = m\lambda; \quad m = 1, 2, 3 \qquad \qquad ...(5.2)$$
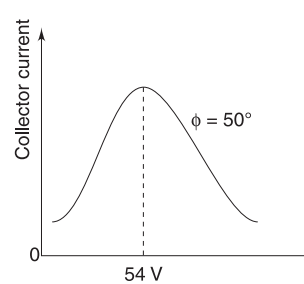
Fig. 5.2
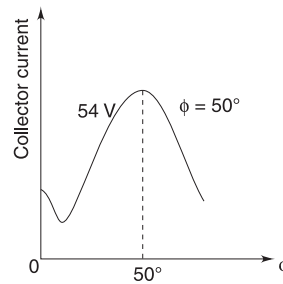
Fig. 5.3
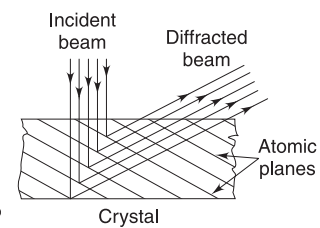
Fig. 5.4

Using X-ray analysis on the crystal it is found that at $\phi = 50°$, a Bragg reflection occurs from atomic plane having interplanar spacing $d = 0.91$ Å and the corresponding Bragg angle of reflection or glancing angle is 65° (as indicated in the figure). Considering $m = 1$ we obtain

$$2 \times (0.91 \text{ Å}) \sin 65° = \lambda$$

or $$\lambda = 1.65 \text{ Å} \qquad \qquad ...(5.3)$$

For electrons having kinetic energy $E = 54$ eV, the de-Broglie wavelength is,

$$\lambda = \frac{h}{\sqrt{2mE}} \qquad \qquad …(5.4)$$

Substituting h, m and E we obtain,

$$\lambda = 1.65 \text{ Å}. \qquad \qquad …(5.5)$$

The existence of electron wave and the validity of de-Broglie equation are thus established.

We may note that in the above calculations the value $m = 1$ is used. If $m = 2$ or more, then there should occur intensity peaks for different values of ϕ. However, no such peaks are observed experimentally.

GP Thomson's Experiment

Thomson's experiment is analogous to Debye-Scherrer X-ray diffraction method.

The experimental arrangement consisted of a glass envelope in which electrons were emitted from a heated filament. The emitted electrons were suitably accelerated and collimated to give a uni-directional, thin, monoenergetic beam of electrons. The beam thus obtained was allowed to fall normally on a polycrystalline material as shown in Figure 5.5. The scattered (diffracted) electrons were recorded on a photographic film placed perpendicular to the incident beam.
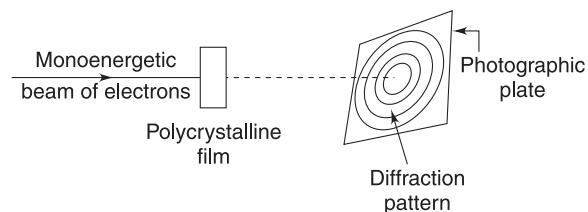
Fig. 5.5 *GP Thomson Experiment*

On the photographic plate a set of concentric circles were observed. The pattern of circles obtained was found to be a characteristic of the crystal used.

On replacing the electron beam by a monochromatic X-ray beam a similar circular pattern was observed on the photographic plate.

From the knowledge of the wavelength of the electron beam

$$\left(\lambda = \frac{h}{p} = \frac{h}{\sqrt{2mE}}\right)$$ it was possible to determine the geometry of the crystal lattice which was found to be in complete agreement with that obtained using X-ray diffraction analysis of the crystal. It is thus clear that electron beam is diffracted by a crystal in the same way as X-rays.

It is important to note in the experiment of Davisson and Germer and of Thomson the following:

- In the process of acceleration, an electron behaves like a particle of charge – e and mass m.

- During the process of diffraction, the same electron behaves like a wave of wavelength l = $\lambda = \frac{h}{p}$.

Thus, the electron which shows wave-like property in one part of the experiment exhibits particle-like properties in two other parts of the same experiment. Clearly, for a complete description both the particle aspect as well as the wave aspect become necessary.

Conclusion

The experiments of Davisson and Germer and Thomson give clear evidence of the existence of wave properties of electrons. Besides, the experiments confirm the validity of de-Broglie equation at least for the electron.

Experiments on diffraction of molecular beam of hydrogen and atomic beam of helium by the lithium fluoride crystal were performed by Estermann, Stern and Frisch. Hydrogen molecule and helium atom being very much different from each other as well as from electron, their successful experiments led to the *universality of matter waves*.

Fermi, Marshall and Zinn performed interference and diffraction experiments with slow neutrons and obtained results confirming de-Broglie's hypothesis. It is important to note that *neutron diffraction* is nowadays an important technique in crystal structure studies as a complement to X-ray and electron diffraction techniques.

5.3 BOSE EINSTEIN STATISTICS

The basic postulates of BE statistics are:

(i) The associated particles are identical and indistinguishable.

(ii) Each energy state can contain any number of particles.

(iii) Total energy and total number of particles of the entire system are constant.

(iv) The particles have zero or integral spin, i.e., $0\hbar$, $1\hbar$, $5\hbar$, $50\hbar$, etc., where $\hbar$ is the unit of spin.

(v) The wave function of the system is symmetric under the positional exchange of any two particles.

Examples: Photon, phonon, all mesons (π, κ, η) etc., these are known as Bosons.

[**Note: Symmetric and Anti-symmetric wave function**

Suppose the allowed wave function for n-particles system is $\psi(1, 2, 3, ..., r, s, ... n)$, where the integers within the argument of y represent the coordinates of the n-particles relative to some fixed orgion. Now, if we interchange the position of any two particles, say, r and s, the resulting wave function becomes $\psi(1, 2, 3, ..., s, r, ...n)$. The wave function y is said to be symmetric when

$$\psi(1, 2, 3, ..., r, s, ... n) = \psi(1, 2, 3, ..., s, r, ...n)$$

and anti-symmetric when

$$\psi(1, 2, 3, ..., r, s, ... n) = -\psi(1, 2, 3, ..., s, r, ...n)]$$

Bose-Einstein Distribution Law

Let Ni number of identical, indistinguishable, non-interacting particles are to be distributed among gi quantum states each having energy E_i. So, in the ith energy level, there are $(N_i + g_i)$ total objects. Keeping the first quantum state fixed, the remaining $(N_i + g_i - 1)$ objects can permuted in $(N_i + g_i - 1)!$ possible ways. But since the particles and the quantum states are indistinguishable, we have to deduct Ni! ways and $(g_i - 1)!$ ways from the all possible ways to get effective number of arrangements. Thus, total number of possible ways of arrangement for the *i*th state is

$$W_i = \frac{(N_i + g_i - 1)!}{N_i!(g_i - 1)!} \qquad ...(5.6)$$

Hence the total number of ways of the entire distribution of N particles in n number of energy levels of the system is

$$W = \prod_i^n \frac{(N_i + g_i - 1)!}{N_i!(g_i - 1)!} \qquad ...(5.7)$$

where P denotes the product symbol.

If we assume that N_i and g_i are very large, eqn. (5.7) reduces to

$$W = \prod_i^n \frac{(N_i + g_i)!}{N_i!\, g_i!} \qquad ...(5.8)$$

Taking natural logarithm of both the sides of eqn. (5.8) we get,

$$\ln W = \sum_i \left[\ln(N_i + g_i)! - \ln N_i! - \ln g_i!\right]$$

$$= \sum_i \left[(N_i + g_i)\ln(N_i + g_i) - N_i \ln N_i - g_i \ln g_i\right] \qquad ...(5.9)$$

(using Stirling approximation)

Now, differentiating equation (5.9) to obtain the probable distribution, we get,

$$d(\ln W)_{max} \sum_i [\ln(N_i + g_i) - \ln N_i]dN_i \qquad ...(5.10)$$

Also we have other two conditions given byz

$$\sum_i d N_i = 0 \text{ (conservation of total no. of particles)} \qquad ...(5.11)$$

$$\sum_i E_i\, d N_i = 0 \text{ (conservation of total energy)} \qquad ...(5.12)$$

Multiplying eqn. (5.11) by $(-\alpha)$ and eqn. (5.12) by $(-\beta)$ and then adding with eqn. (5.10) we get,

$$\sum_i [\ln(N_i + g_i) - \ln N_i - \alpha - \beta E_i]dN_i = 0 \qquad ...(5.13)$$

Since dN_i's are independent of one another, the above equation holds only if,

$$\sum_i [\ln(N_i + g_i) - \ln N_i - \alpha - \beta E_i] = 0$$

or, $\qquad N_i = \dfrac{g_i}{e^{(\alpha + \beta E_i)} - 1}$ ...(5.14)

Now the Bose-Einstein distribution function is given by

$$f(E_i) = \frac{N_i}{g_i} = \frac{1}{e^{(\alpha + \beta E_i)} - 1}$$ ...(5.15)

which represents the probability of finding a boson with energy E_i.

Application of Bose-Einstein statistics

Let us discuss the application of Bose-Einstein statistics.

Planck's Law of Black-body Radiation

In quantum mechanics we have already established Planck's law of black-body radiation which exactly accounts for the observed energy density in case of a black-body radiation. Here shall re-derive the same Planck's law by using Bose-Einstein statistics. In order to derive this, let us consider a black-body chamber of volume V kept at a constant temperature T and filled with radiant energy that can be considered as an assembly of photons. If the number of photons in the chamber is very large, the spacing between two successive energy levels becomes very small making the energy lavels almost continuous. Thus if the energy of the photons ranges from E to $E + dE$, the number of degnerate states gi should be replaced by $g(E)dE$ and the total number of photons N_i in these states should be replaced by $N(E)dE$ in the distribution function given by equation (5.14). Thus the total number of photons in the chamber having energy between E to $E + dE$ can be written by the following eqn. (5.14).

The number of photons having energy range between E to $E + dE$ can be written by using eqn. (5.14) as given by

$$N(E)dE = \frac{g(E)}{e^{(\alpha + \beta E)} - 1} dE$$ ...(5.16)

where $g(E)dE$ is the number of states of photons having energy between E to $E + dE$. In case of black-body radiation, the total number of particles are not conserved because, photons are absorbed and re-emitted frequently by the walls of the chamber, i.e.,

$\sum_i dN_i$ 0, which implies $\alpha = 0$. Also it is experimentally established

that $\beta = \dfrac{1}{kT}$, where, $k = 1.38 \times 10^{-23}$ Joule/Kelvin is known as the Boltzmann constant. Thus eqn. (7.48) reduces to

$$N(E)dE = \frac{g(E)dE}{e^{E/kT} - 1}$$ ...(5.17)

Now the number of quantum states corresponding to the momentum range from p to $p + dp$ is shown in this book

$$g(E)dE = g_s \frac{4\pi V p^2 dp}{h^3}$$ ...(5.18)

where, g_s is the spin-degeneracy or duplicity of a quantum state. Since a photon has two spin orientations in trensverse direction, $g_s = 2$ for photons.

$$\therefore \qquad g(p)dp \;=\; \frac{8\pi V p^2 dp}{h^3} \qquad\qquad ...(5.19)$$

The energy of a photon of frequency v is $E = hv$ and so its momentum is $p = hv/c$, where c is the speed of a photon in free space.

$$\therefore \qquad dp = \frac{h}{c}\,dv \qquad\qquad ...(5.20)$$

Substituting the value of p and dp in eqn. (5.19), we get the number of quantum states having frequency range between v and $v + dv$ as

$$g(v)dv \;=\; \frac{8\pi V}{c^3}\; v^2 dv \qquad\qquad ...(5.21)$$

Hence, expressing eqn. (5.16) in terms of v we get,

$$\begin{aligned}
N(v)dv \;&=\; \frac{g(v)dv}{e^{hv/kT}-1}\\
&=\; \frac{8\pi V}{c^3}\cdot\frac{v^2 dv}{e^{hv/kT}-1} \qquad\qquad ...(5.22)
\end{aligned}$$

which represents the number of photons having frequency range between v to $v + dv$ kept in a chamber of volume V at temperature T.

Therefore, the energy density (total energy per unit volume) of the photons within frequency range dv is given by

$$\begin{aligned}
u(v)dv \;&=\; \frac{hv}{V}N(v)dv\\[4pt]
&=\; \frac{hv}{V}\cdot\frac{8\pi V}{c^3}\cdot\frac{v^3 dv}{e^{hv/kT}-1}
\end{aligned}$$

$$\text{or,} \qquad u(v)dv \;=\; \frac{8\pi h}{c^3}\cdot\frac{v^3 dv}{e^{hv/kT}-1} \qquad\qquad ...(5.23)$$

which is the Planck's law of black-body radiation that we have deduced earlier in quantum mechanics.

5.4 BLACK-BODY RADIATION

In our world of material, all the objects emit thermal radiation from its surface at any temperature. The radiation characteristics depend on the temperature and properties of the surface. As the surface temperature of the object increases, wavelength of the radiation goes from infrared to white region.

The thermally agitated accelerating particles emits the radiation and distributed the energies in continuous manner, as the result continuous spectrum of electromagnetic radiation emitted by the object. The classical theory of thermal radiation was inadequate, when one consider the distribution of wavelengths in the thermal radiation emitted by a blackbody.

A blackbody is defined as it absorbs all the thermal radiations, whatever may be the wavelength incident on it. It neither reflects nor transmits any radiation, and appears black.

Fig. 5.6 *Absorption of radiations in a blackbody*

An isothermal enclosure behaves same as a blackbody. Let us consider a blackbody, placed in an isothermal enclosure. In thermal equilibrium, the body will emit the full radiations independent of the nature of the substance. The thermal radiations in an isothermal enclosure are, therefore, termed as 'Blackbody Radiations.'

Stefan's Law

The frequency distribution (or wavelength distribution) of radiation from a blackbody cavity or an isothermal enclosure was studied in the late nineteenth century. Figure 5.7 shows the intensity versus wavelength curve in different temperature of a blackbody radiation.

According to Stefan's law, the emissive power of the surface of a blackbody is given by

$$P = \sigma A e T^4 \hspace{4cm} ...(5.24)$$

where σ is the Stefan-Boltzmann constant, and its value is 5.670×10^{-8} W/m^2.K^4, A is the surface area of the blackbody, e is the emissivity of the surface and T is the temperature in Kelvin. For a perfectly blackbody $e = 1$.

The intensity formulae is given by

$$I = \sigma T^4 \hspace{4cm} ...(5.25)$$

where $\qquad I = \dfrac{P}{A}$

The above equations are called Stefan's laws in blackbody radiations.
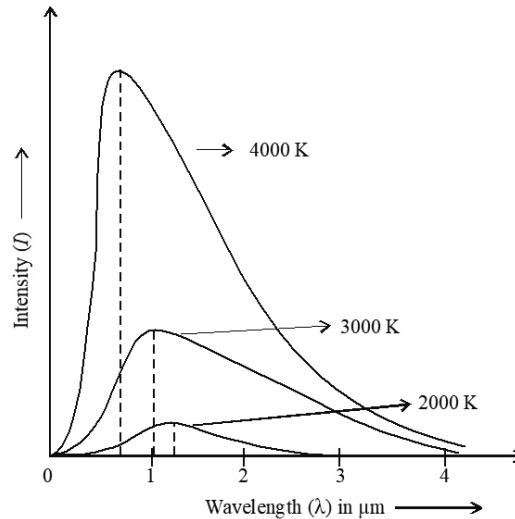
Fig. 5.7 *Intensity-wavelength distribution of a blackbody radiation*

Wein's Radiation Law

To explain the observed spectral distribution (given in figure 5.7), Wein first found the form of E_λ (emissive energy) as the function of wavelength (λ) and temperature (T), i.e.,

$$E_\lambda = E_\lambda\ (\lambda,\ \text{T}\ ).$$

Wein's consider an isothermal cavity (sphere), increasing gradually. According to Wein, the energy density (E_λ) of the blackbody radiation of wavelengths between λ and $\lambda + d\lambda$ from the isothermal cavity of temperature T is given by

$$E_\lambda d\lambda = \frac{A}{\lambda^5} e^{-B/\lambda T} d\lambda \qquad \qquad ...(5.26)$$

where A and B are two constants.

The above equation is termed as "Wein's Distribution Law" in blackbody radiation.
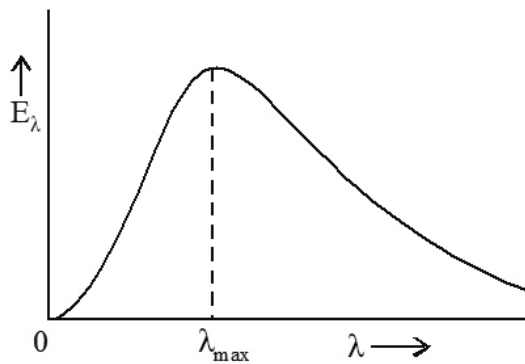
Fig. 5.8 *Variation of E_λ with λ (Wein's distribution raw)*

Wein's also showed that

$$T\lambda_{max} = \text{constant} \qquad \qquad ...(5.27)$$

where lmax is the wavelength at the maximum energy of a blackbody radiation.

The above equation is termed as "Wein's Displacement Law" in blackbody radiation.

Numerically

$$T\lambda_{max} = 2.898 \times 10^{-3} \text{ m.K}$$

Rayleigh-Jeans's Distribution Law

Rayleigh-Jeans's distribution law in blackbody radiation based on the principle of equipartition of energy in classical limits. They consider that the energy is distributed in all possible modes of free vibration which might be assigned to radiation of energy. Thus, they considered the average energy (energy per mode in each degree of freedom) of an oscillator as $\bar{\in} = KT$ (classical approach).

According to Rayleigh-Jeans's law

$$E_v dv = \frac{8\pi v^2 KT}{c^3} dv \qquad \qquad ...(5.28)$$

where the energy of the radiation varies from v to $v + dv$ and K is the Boltzman's constant. The above equation is known as "Rayleigh-Jeans's distribution law" in blackbody radiation.

In terms of wavelength. the "Rayleigh-Jeans's distribution law" takes the form

$$E_v \, dv = \frac{8\pi KT}{\lambda^4} d\lambda \qquad \qquad ...(5.29)$$

where $c = n\lambda$ and the energy of the radiation varies from λ to $\lambda + d\lambda$.

So, "Rayleigh-Jeans's distribution law" states that the energy density (E_λ) of a blackbody radiation is inversely proportional to the fourth power of its wavelength (λ) at temperature T,

i.e., $\qquad \qquad E_\lambda \, \alpha \, \dfrac{1}{\lambda^4}$ $\qquad \qquad ...(5.30)$

Ultraviolet Catastrophe

The exact explanation of blackbody spectra was first given by Rayleigh and Jeans. The spectral distribution of a blackbody radiation is given in Fig. 5.9.

According to Rayleigh-Jeans's formulae $E_\lambda \, d\lambda = \dfrac{8\pi \, KT}{\lambda^4} d\lambda$. From this equation, we can say, the energy density increases with decreasing wavelength. But the actual spectra (experimental results) shows that the energy density increases with increasing wavelength (for low wavelengths).

The discrepancy in theoretical drawback of Rayleigh-Jeans's law with experimental results of blackbody radiation (for low wavelengths) is termed as "ultraviolet catastrophe".

[Note: The words "ultraviolet" for short wavelength and "catastrophe" mean infinite energy occurs as the wavelength tends to zero; i.e., when $l \to 0$, $E_\lambda \to \infty$]
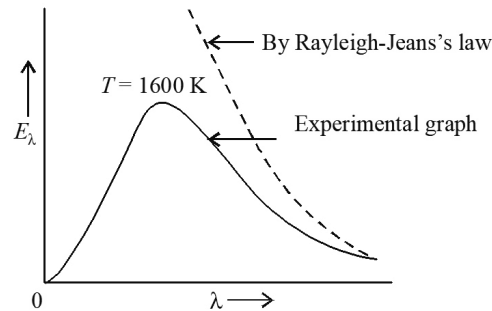
Fig. 5.9 Comparison of experimental spectra with Rayleigh-Jeans's law

Planck's Distribution Law

In 1900, Max Planck developed the complete explanations of blackbody radiation that fit into experimental results. Planck developed his ideas on the basis of "quantum" aspects.

To explain the blackbody radiation Planck assumed that

(a) The isothermal cavity (blackbody) oscillates like an atomic oscillator and energy radiate.

(b) The cavity is vibrated only in frequency v and its full multiple; i.e., the energy will be hv and its full multiple upto nhv, where n is an integer. h is called "Planck's Constant".

So, we can say the energy will be discrete and quantited. Each quantited discrete energy value leads to different "quantum state", represented

by the quantum number n, i.e., the discrete energy of different quantum state will be $h\nu$, $2h\nu$, $3h\nu$, ... $\nu h\nu$ ($n = 1, 2, 3, ... \nu$). The difference of energy between two consecutive energy state will be $h\nu$.

(c) The oscillators emit or absorb energy when transition takes place, i.e., the oscillators emit or absorb energy when it changes the quantum states.
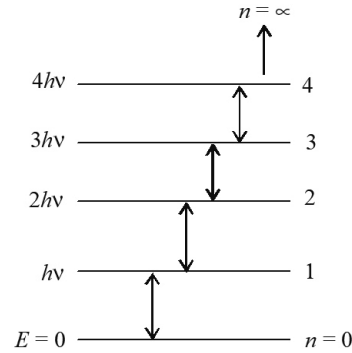
Fig. 5.10 *Allowed energy levels of an oscillator*

Planck consider that the energy is proportional to $e^{-nh\nu/KT}$.

Let us consider there are N number of oscillators of total energy E_n. The energy per oscillator is given by

$$\bar{\varepsilon} = \frac{E_n}{N} \qquad ...(5.31)$$

where $\qquad N = N_0 + N_1 + N_2 + ... + N_n + ... \qquad ...(5.32)$

and $\qquad E_n = 0.N_0 + \varepsilon N_1 + 2\varepsilon N_2 + ... + n\varepsilon N_n + ... \qquad ...(5.33)$

$N_0, N_1, N_2, ... N_n + ...$ are the number of oscillators correspond to energy $0, \varepsilon, 2\varepsilon, ... n\varepsilon...$ respectively and $\varepsilon = h\nu$.

According to Maxwell's distribution law, the total number of oscillators having energy nhv is given by

$$N_n = \sum_{n=0}^{\alpha} N_0 e^{-nh\nu/KT} \qquad ...(5.34)$$

where K is the Boltzmann's constant.

According to Planck's consideration (energy is proportional to $e^{-nh\nu/KT}$), the total energy of the oscillators is given by

$$E_n = \sum_{n=0}^{\alpha} N_0 e^{-nh\nu/KT} nh\nu \qquad ...(5.35)$$

So, the average energy $\bar{\varepsilon}$ (energy per oscillator) is given by

$$\bar{\varepsilon} = \frac{\sum_{n=0}^{\alpha} N_0 e^{-nh\nu/KT} nh\nu}{\sum_{n=0}^{\alpha} N_0 e^{-nh\nu/KT}}$$

[from equations (5.31), (5.34) and (5.35)].

$$\Rightarrow \quad \bar{\varepsilon} = -\frac{d}{d\beta} \ln \sum_{n=0}^{\alpha} e^{-(nh\nu\beta)} \text{, where } \beta = \frac{1}{KT}$$

$$= -\frac{d}{d\beta}\ln[1 + e^{-hv\beta} + e^{-2hv\beta} + \dots + e^{-nhv\beta} + \dots]$$

$$= -\frac{d}{d\beta}\ln\left[\frac{1}{1 - e^{-hv\beta}}\right] = -\frac{d}{d\beta}\ln[1 - e^{-hv\beta}]$$

$$= \frac{hv.e^{-hv\beta}}{1 - e^{hv\beta}} = \frac{hv}{e^{hv\beta} - 1} = \frac{hv}{e^{hv/KT} - 1} \qquad \dots(5.36)$$

As the number of oscillators per unit volume within the frequency range v and $v + dv$ is given by

$$dN \qquad = \frac{8\pi v^2}{c^3} dv,$$

the emissive energy density (energy per unit volume) within the frequency range v and $v + dv$ is given by

$$E_v \, dv = dN\overline{\varepsilon} \ = \ \frac{8\pi v^2}{c^3} dv \times \frac{hv}{e^{hv/KT} - 1}$$

$$= \frac{8\pi h v^3}{c^3} \frac{1}{e^{hv/KT} - 1} dv \qquad \dots(5.37)$$

The above equation is known as Planck's distribution law in blackbody radiation.

In terms of wavelength, the Planck's distribution law having the form

$$E_\lambda d\lambda = \frac{8\pi hc}{\lambda^5} \frac{1}{e^{hc/\lambda KT} - 1} d\lambda \qquad \dots(5.38)$$

where $v = \dfrac{c}{\lambda}$ and h is called Planck's constant.

Graphical Representation of Wein's, Rayleigh-Jeans's and Planck's Distribution Law
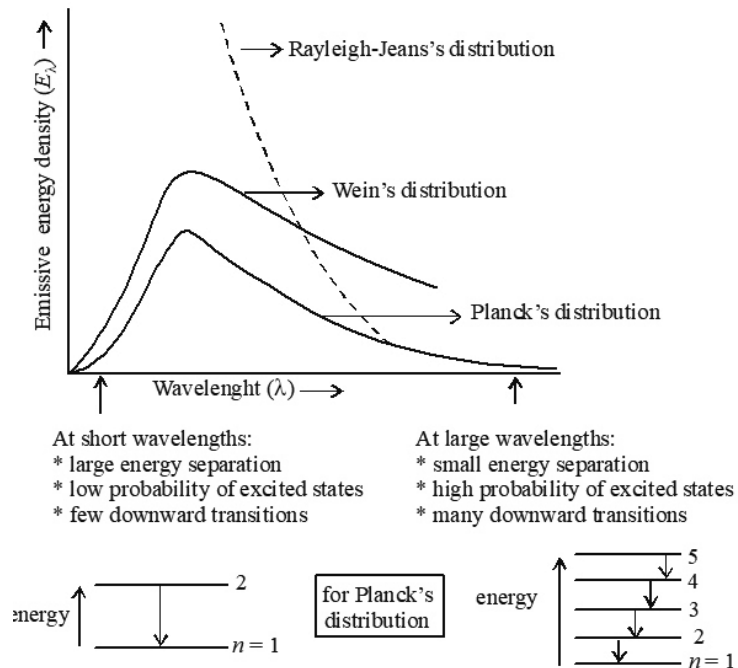
Fig. 5.11 *Emissive energy density-wavelength distribution of Wein's, Rayleigh-Jeans's and Planck's law in blackbody radiation*

Corollary 1. Wein's Distribution Law from Planck's Law: When the temperature (T) or the wavelength (λ) is very low of equation (5.38), then the term $\dfrac{hc}{\lambda KT}$ is very much greater than unity $\left(\dfrac{hc}{\lambda KT} \gg 1\right)$. Hence, the form of equation (5.38) will be

$$E_\lambda \, d\lambda = \frac{8\pi hc}{\lambda^5} e^{-hc/\lambda KT} d\lambda$$

$$= \frac{A}{\lambda^5} e^{-B/\lambda T} d\lambda \qquad\qquad ...(5.39)$$

which is Wein's distribution law [Refer equation (5.26)] where $A = 8\pi hc$ (constant) and $B = \dfrac{hc}{K}$ (constant).

Corollary 2. Rayleigh–Jeans Distribution Law from Planck's Law: When the temperature (T) or the wavelength (λ) is very much large, then $\dfrac{hc}{\lambda KT}$ in equation (5.36) is very small. Then the term $e^{hc/\lambda KT}$ can be expanded as

$$e^{hc/\lambda KT} = \left[1 + \frac{hc}{\lambda KT} + \frac{1}{2}\left(\frac{hc}{\lambda KT}\right)^2 + ...\right]$$

$$= 1 + \frac{hc}{\lambda KT} \qquad\qquad \text{[neglecting higher term for lower values]}$$

So, the equation (4.15) having the form

$$E_\lambda d\lambda = \frac{8\pi hc}{\lambda^5} \cdot \frac{1}{\left[\left(1 + \dfrac{hc}{\lambda KT}\right) - 1\right]} d\lambda$$

$$= \frac{8\pi hc}{\lambda^5} \times \frac{\lambda KT}{hc} d\lambda$$

$$= \frac{8\pi KT}{\lambda^4} d\lambda \qquad\qquad ...(5.40)$$

which is Rayleigh–Jeans's distribution law (Refer equation (5.29)).

Corollary 3. Stefan's Law from Planck's Law: From equation (5.37), the Planck's distribution law in terms of frequency is given by

$$E\nu \, d\nu = \frac{8\pi h\nu^3}{c^3} \frac{1}{e^{h\nu/KT} - 1} d\nu$$

By integrating the above equation from 0 to ∞, we can obtain the total radiation energy. So, the total energy is given by

$$E = \int_0^\infty E_\nu d\nu = \frac{8\pi h}{c^3} \int_0^\infty \frac{\nu^3}{e^{h\nu/KT} - 1} d\nu$$

$$= \frac{8\pi(KT)^4}{c^3 h^3} \int_0^\infty \frac{x^3}{e^x - 1} dx \quad \left[\text{Putting } x = \frac{h\nu}{KT}, \ dx = \frac{h}{KT} d\nu\right]$$

$$= \frac{8\pi KT^4}{c^3 h^3} \times \frac{\pi^4}{15} \left[\int_0^\infty \frac{x^3}{e^x - 1} dx = \frac{\pi^4}{15}\right] \qquad\qquad ...(5.41)$$

The emissive intensity of any blackbody radiation is defined as

$$I = E\left(\frac{c}{4}\right) = \frac{2\pi^5 K^4 T^4}{15 c^2 h^3} = \sigma T^4 \qquad \qquad ...(5.42)$$

where $\sigma = \dfrac{2\pi^5 K^4}{15 c^2 h^3}$ $(= 5.67 \times 10^{-8} \text{ W/m}^2/\text{K}^4)$ is called Stefan's constant.

Equation (5.42) is Stefan's law in blackbody radiation

Check Your Progress

1. What is quantum physics?

2. How does the theory of relativity characterize quantum physics?

3. State the Louis de-Broglie idea of duality.

4. What is mutual dependence according to de-Broglie?

5. What is the wavelength of electrons impinging the crystal according to de-Broglie's equation?

6. What are basic postulates of BE statistics?

7. Two particles are to be distributed in three cells. Find out the total number of possible ways both from formula and in tabular form in BE statistics particles.

8. How will you define the Planck's law of black-body radiation?

9. Define the term blackbody.

5.5 PRESSURE OF AN IDEAL BOSE GAS

In this section, we shall study the thermodynamic properties of a gas of non-interacting bosons. We will show that the symmetrization of the wavefunction due to the indis-tinguishability of particles has important consequences on the behavior of the system. The most important consequence of the quantum mechanical symmetrization is the Bose-Einstein condensation, which is in this sense a special phase transition as it occurs in a system of non-interacting particles. We shall consider, as an example, a gas of photons and a gas of phonons.

Equation of state

We consider a gas of non-interacting bosons in a volume V at temperature T and chemical potential μ. The system is allowed to interchange particles and energy with the surround-ings. The appropriate ensemble to treat this many-body system is the grand canonical ensemble.

Non-relativistic Bosons. Our bosons are *non-relativistic* particles with spin s, whose one-particle energies $\epsilon(\mathbf{k})$

$$\epsilon(\mathbf{k}) = \epsilon(k) = \frac{\hbar^2 k^2}{2m}, \qquad \epsilon_0 = \epsilon(0) = 0$$

include only the kinetic energy term.

Negative chemical potential. The chemical potential obeys

$$-\infty < \mu < \epsilon_0, \qquad \epsilon_0 = 0 .$$

A chemical potential larger than the lowest energy state would lead to nonphysical level occupation.

$$n(\epsilon_r) = \langle \hat{n}_r \rangle = \frac{1}{e^{\beta(\epsilon_r - \mu)} - 1} .$$

Approaching the thermodynamic limit: We consider a situation when the gas is in a box with volume $V = L_x L_y L_z$ and subject to periodic boundary conditions

In the thermodynamic limit ($N \to \infty$, $V \to \infty$, with $n = N/V = $ const), the sums over the wavevector $\vec{k}$ can be replaced by integrals as in the case of the Fermi gas.

However, here we have to be careful when μ happens to approach the value 0. In order to see what kind of trouble we then might get into, let us calculate the ground state occupation.

Occupation of the lowest energy state. We consider the expectation value of the ground state for μ approaching zero from below, viz when $-\beta\mu \ll 1$:

$$n(\epsilon_0) = \frac{1}{e^{-\beta\mu} - 1} = \frac{1}{(1 - \beta\mu + \ldots) - 1} \approx -\frac{1}{\beta\mu}, \qquad \epsilon_0 = 0 ,$$

which means that $n(\epsilon_0) = \langle \hat{n}_{r=0} \rangle$ diverges. The lowest energy state may hence by occupied *macroscopically*. This is the case when

$$\frac{1}{|\mu|\beta} \sim N, \qquad |\mu| \sim \frac{k_B T}{N}, \qquad 1 - z \sim \frac{1}{N} . \tag{5.43}$$

Density of states: The density of states $D(E)$,

$$D(E) \sim \sqrt{E}, \qquad \lim_{E \to 0} D(E) = 0,$$

vanishes for $E \to 0$. This is where we are going to encounter a problem: if we replace

$$\frac{1}{V} \sum_r \quad \to \quad \int dE \ D(E)$$

we will get that the ground state has zero weight even though, as we have just shown that it can be macroscopically occupied. Fermionic systems do not encounter this problem due to the Pauli principle, which imposes that $\langle \hat{n}_r \rangle \in [0, 1]$.

Special treatment for the ground state. The problem with the potentially macroscopic occupation of the ground state can be solved by giving it via

$$\begin{aligned}
\beta\Omega(T, V, z) &= \sum_r \ln\left[1 - e^{-\beta(\epsilon_r - \mu)}\right] &\tag{5.44}\\
&= (2s+1)\frac{V}{(2\pi)^3} 4\pi \int_0^\infty dk \ k^2 \ \ln\left(1 - z e^{-\beta\epsilon(k)}\right) &\\
&+ \underbrace{(2s+1)\ln(1 - z)}_{\substack{\text{occupation of the}\\ \text{ground state } \epsilon(0) = 0}} . &\tag{5.45}
\end{aligned}$$

a special treatment. For the bosonic grand canonical
potential $\Omega(T, V, \mu)$ we have split the $\sum_r$ into an integral over all states, and into the ϵ_0 contribution (the last term).

'Irrelevance' of condensate. The ground state contribution Equation (5.45) to the grand canon-ical potential Ω is formally irrelevant in the thermodynamic limit as a consquence of the scaling Equation (5.43) of the chemical potential:

$$\lim_{V \to 0} \frac{\ln(1-z)}{V} \approx \lim_{V \to 0} \frac{-\ln(N)}{V} \to 0 .$$

We note, however, that the size of the *condensate*, that is the number of particles occupying the ground state, determines how many particles occupy energiers $E > E_0$, viz the density of the *normal fluid*.

Dimensionless variables. With the dimensionless variable x and the thermal de Broglie wavelength λ,

$$x = \hbar k \sqrt{\frac{\beta}{2m}}, \qquad \lambda = \sqrt{\frac{2\pi\beta\hbar^2}{m}},$$

we write (14.3) as

$$\beta\Omega(T, V, z) = \frac{2s+1}{\lambda^3} \frac{4V}{\pi} \int_0^\infty dx\, x^2 \ln\left(1 - ze^{-x^2}\right), \tag{5.46}$$

all in parallel to the transformations performed for the Fermi gas.

Taylor expansion. We recall that the Taylor series expansion

$$\ln(1-y) = -\sum_{n=1}^\infty \frac{y^n}{n}, \qquad |y| < 1$$

may be used, to express the integral

$$\int_0^\infty dx\, x^2 \ln\left(1 - ze^{-x^2}\right) = -\frac{\sqrt{\pi}}{4} \sum_{n=1}^\infty \frac{z^n}{n^{5/2}}$$

in terms of

$$g_{5/2}(z) = -\frac{4}{\sqrt{\pi}} \int_0^\infty dx\, x^2 \ln\left(1 - ze^{-x^2}\right) = \sum_{n=1}^\infty \frac{z^n}{n^{5/2}} . \tag{5.47}$$

Note that $g_{5/2}(z)$ and $f_{5/2}(z)$, differ by a sign $(-1)^{n+1}$ in the summand. For later uses we also define $g_{3/2}(z)$ as

$$g_{3/2}(z) = z\frac{d}{dz} g_{5/2}(z) = \sum_{n=1}^\infty \frac{z^n}{n^{3/2}} . \tag{5.48}$$

Note that $f_{3/2}(z)$ was defined analogously in as $zd(f_{5/2})/dz$.

Bosonic grand canonical potential. With (5.47), the grand canonical potential (5.46) takes the form

$$\boxed{\beta\Omega(T, V, z) = -\frac{2s+1}{\lambda^3} V g_{5/2}(z)} . \tag{5.49}$$

Except for the extra term on the right-hand side, and for an exchange $g_{5/2} \leftrightarrow f_{5/2}$,

Pressure. From $\Omega = -PV$ we get

$$\boxed{\beta P = \frac{2s+1}{\lambda^3} g_{5/2}(z)} , \tag{5.50}$$

in analogy to above equation.

Particle density. For the particle density N/V we derived in this book relation

$$\frac{\langle N \rangle}{V} = \frac{1}{\beta V}\left[\frac{\partial}{\partial \mu} \ln \mathcal{Z}\right] = \frac{z}{V}\left[\frac{\partial}{\partial z} \ln \mathcal{Z}\right] = -\frac{\beta z}{V}\left[\frac{\partial}{\partial z} \Omega\right],$$

namely that $\beta\Omega = -\ln \mathcal{Z}$. In this case the condensate term (5.45) contributes. We then find

$$n = \frac{2s+1}{\lambda^3} g_{3/2}(z) + \frac{2s+1}{V} \frac{z}{1-z} \qquad (5.51)$$

for the density of particles $n = \langle \hat{N} \rangle / V$ in terms of the fugacity z, where we have used (5. 49) and (5. 48).

Ground state occupation. The term
$$n_0 = \frac{2s+1}{V} \frac{z}{1-z} \qquad (5.52)$$

in (5.51) describes the contribution of the ground state to the particle density n. When n_0 becomes macroscopically large on speaks of a *Bose-Einstein condensation.*

Internal energy. The internal energy U is given by
$$U = -\left(\frac{\partial}{\partial \beta} \ln \mathcal{Z}(T, z, V) \right)_{z,V} = \left(\frac{\partial}{\partial \beta} \beta \Omega(T, V, z) \right)_{z,V}. \qquad (5.53)$$

Note that the fugactity z is kept constant in above expression. Our result (in equation 5.49) states that
$$\beta \Omega(T, V, z) \sim -\lambda^{-3} \sim -\beta^{-3/2},$$
which then leads with (5.53) to

$$\frac{U}{V} = \frac{3k_B T}{2} \frac{2s+1}{\lambda^3} g_{5/2}(z). \qquad (5.54)$$

This expression is the same as the one for the Fermi gas when $f_{5/2}(z)$ is substituted by $g_{5/2}(z)$. The reason is that the ground state energy ϵ_0 vanishes, $\epsilon_0 = 0$. It does hence not matter how many particle occupy the lowest energy level.

Caloric equation of state. Combining Equation (5.54) with (5.50) one can derive the caloric equation of state:

$$U = \frac{3}{2} PV,$$

which is identical to the one obtained for the ideal Fermi gas.

Classical Limit

The classical limit (non-degenerate Bose gas) corresponds to low particle densities and high temperatures. The fugacity is then small,

$$z = e^{\beta \mu} \ll 1,$$

with the Bose-Einstein distribution

$$\langle \hat{n}_r \rangle = \frac{1}{z^{-1} e^{\beta \epsilon_r} - 1} = \frac{z e^{-\beta \epsilon_r}}{1 - z e^{-\beta \epsilon_r}} \approx z e^{-\beta \epsilon_r} \ll 1$$

reducing to the Maxwell-Boltzmann distribution, just as for a fermionic system. The differences between Bose-, Fermi- and Boltzmann statistics are in next order of the order $1/z$ and hence small.

Expansion in the fugacity. As $z \ll 1$, it is sufficient to retain only the first two terms of the series for $g_{5/2}(z)$ and $g_{3/2}(z)$:

$$g_{5/2}(z) \approx z + \frac{z^2}{2^{5/2}}, \qquad g_{3/2}(z) \approx z + \frac{z^2}{2^{3/2}} \; .$$

With that, the particle density (5.51) takes the following form:

$$n \approx \frac{2s+1}{\lambda^3} z \left(1 + \frac{z}{2^{3/2}}\right) + \frac{2s+1}{V} \frac{z}{1-z} \tag{5.55}$$

Irrelevance of the ground-state contribution. We note that the ground-state contribution (5.52) vanishes generically in the thermodynamic limit, being proportional to $1/V$. The number of particles occupying the ground state is finite only for $z \to 1$. For the case of small fugacities considered here we can neglect it generically, obtaining

$$n \approx \frac{2s+1}{\lambda^3} z \left(1 + \frac{z}{2^{3/2}}\right) \; . \tag{5.56}$$

Convergence radius. That ground state contributions can be generically neglected at elevated temperatures follows also from the following consideration.

- An expansion in a physical parameter, like z or β, converges only as long as one one remains within the same phase, here the high-temperature quasi-classical gas `
- An expansion *diverges* once a phase boundary is encountered, in our case the transition to a phase with macroscopically occupied ground-state level. A condensed low-temperature state can therefore not be described within a high-temperature expansion.

Classical limit. In the strict classical limit we retain only the the terms $\sim z$ on the right-hand side of (5.56) We thus have that

$$n\lambda^3 \approx (2s+1)z^{(0)}, \qquad \boxed{z^{(0)} \approx \frac{n\lambda^3}{2s+1}} \; . \tag{5.57}$$

in the zeroth approximation.

Classical equation of state. We note that (5.57) is identical to corresponding **given** expression **in this book** for Fermions. The expression (5.60) for the pressure reduces then with $g_{5/2} \to z^{(0)}$ to the equation of state for classical particles:

$$\beta P = \frac{2s+1}{\lambda^3} \frac{n\lambda^3}{2s+1}, \qquad VP = \langle \hat{N} \rangle k_B T, \qquad n = \langle \hat{N} \rangle / V \; .$$

First order correction. We solve (5.57),

$$\frac{n\lambda^3}{2s+1} \approx z^{(0)} \approx z^{(1)} \left(1 + \frac{z^{(1)}}{2^{3/2}}\right) \approx z^{(1)} \left(1 + \frac{z^{(0)}}{2^{3/2}}\right)$$

for $z^{(1)}$, obtaining

$$z^{(1)} \approx \frac{z^{(0)}}{1 + \frac{z^{(0)}}{2^{3/2}}} \approx z^{(0)} \left(1 - \frac{z^{(0)}}{2^{3/2}}\right) \; , \tag{5.58}$$

which may be substituted in the pressure equation (5.50) for the ideal Bose gas:

$$\beta P \approx \frac{2s+1}{\lambda^3} g_{5/2}(z) \approx \frac{2s+1}{\lambda^3} z^{(1)} \; . \tag{5.59}$$

Quantum correction. Taking all together **Equation** (5.59), (5.58) and (5.57), gives us

$$\boxed{PV = \langle \hat{N} \rangle k_B T \left[1 - \frac{n\lambda^3}{4\sqrt{2}(2s+1)}\right]} \; . \tag{5.60}$$

The last term in this expression are the *quantum corrections.*

– Equation of states for a real gas, like the van der Waals equation in this book, posses "similar" additive corrections with respect to the ideal case, which are however due to the interaction between particles. The additive terms present in (5.60) originate on the other side from the indistinguishability principle and *not* from the interaction among particles.

– The correcting term for the ideal Fermi gas quasi-classical equation of state is positive, contributing as a "repulsion" among particles. For the Bose gas, the additive term is negative and therefore contributes as an "attraction" among particles.

Quantitatively, the quantum corrections are much smaller than terms coming from the interaction among particles.

5.6   EINSTEIN CONDENSATION

We consider now the limit of *high* particle densities and low temperatures (quantum limit), where one finds important qualitative differences between bosons, fermions and classical particles.

Particle density. We rewrite the the particle density (5.51) as

$$n = \int_0^\infty \frac{D(\epsilon)d\epsilon}{e^{\beta(\epsilon-\mu)} - 1} + \frac{2s+1}{V}\frac{z}{1-z} , \tag{5.61}$$

where the the density of states $D(\epsilon)$

$$D(\epsilon) = A\sqrt{\epsilon}, \qquad A = \frac{2s+1}{(2\pi)^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} .$$

Dimensionless variables. The regular contribution to the particle density in (5.61) can be evaluated for $\mu = 0$ as

$$\lim_{\mu\to 0}\int_0^\infty \frac{A\sqrt{\epsilon}d\epsilon}{e^{\beta(\epsilon-\mu)} - 1} = \frac{A}{\beta^{3/2}}\int_0^\infty \frac{\sqrt{x}dx}{e^x - 1} = \frac{A}{\beta^{3/2}}\cdot 2.61 ,$$

where we have used the dimensionless variable $x = \beta\epsilon$. The original expression (5.62) for n then becomes

$$\boxed{n = 2.61A(k_BT)^{3/2} + \frac{2s+1}{V}\frac{1}{e^{-\beta\mu} - 1}} , \qquad A = \frac{2s+1}{(2\pi)^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} . \tag{5.62}$$

This is a mixed representation where we have taken the limes $\mu \to 0$ for the regular contribution, but not for the occupation of the ground state.

Bose-Einstein condensation. It is evident from (5.62) that there is a critical temperature T_c,

$$n = 2.61A(k_BT_c)^{3/2}, \qquad \boxed{n = 2.61\frac{2}{\sqrt{\pi}}\frac{2s+1}{\lambda_c^3}} , \qquad \lambda_c = \sqrt{\frac{h^2}{2\pi mk_BT_c}} \tag{5.63}$$

for which the regular contribution would fall below the desired particle density n. We have used that $(4\pi)^{3/2}/4\pi^2 = 2/\sqrt{\pi}$.

– T_c is the *Bose-Einstein transition temperature.*

– A non-vanishing negative chemical potential $\mu < 0$ would lead to an even small regular term in (5.61). There is therefore no way that the regular term could account for all particle for $T < T_c$.

– The transition takes place when $n\lambda_c^3/(2s + 1) = 2.61\cdot 2/\sqrt{\pi} \approx 2.9$, viz when the thermal wavelength λ_c is of the order of the inter-particle distance.

Scaling of the chemical potential. Rewriting (5.62) for small $|\mu|$ as

$$n - n_c(T) \sim \frac{2s+1}{V} \frac{k_B T}{-\mu}, \qquad \boxed{(-\mu) \sim \frac{2s+1}{V} \frac{k_B T}{n - n_c(T)}}, \qquad (5.64)$$

where $n_c(T) = 2.61(2s+1)/\lambda^3$. The chemical potential scales therefore like $1/V$, viz it strictly vanishes only in the thermodynamic limit $V \to \infty$.

First excited state. The energy level are quantized for a particle in a box,

$$\epsilon(\mathbf{k}) = \frac{k_x^2 + k_y^2 + k_z^2}{2m}, \qquad k_\alpha = \frac{2\pi}{L_\alpha} n_\alpha, \qquad \alpha = x, y, z ,$$

The volume is with $V = L_x L_y L_z$ the product of the linear dimensions.

Diverging occupation of the first excited state. The energy ϵ_1 of one of the first excited states, corresponding e.g. to $(n_x, n_y, n_z) = (1, 0, 0)$, then scales as

$$\epsilon_1 \sim \frac{1}{L_x^2} \sim V^{-2/3}, \qquad \epsilon_1 \gg |\mu| \sim V^{-1} .$$

The occupation n_1 of the first excited state,

$$n_1 = \frac{1}{e^{\beta(\epsilon_1 - \mu)} - 1} \approx \frac{1}{e^{\beta\epsilon_1} - 1} \approx \frac{1}{1 + \beta\epsilon_1 - 1} \sim V^{2/3},$$

therefore diverges in the thermodynamic limit $V \to \infty$. The ground-state occupation $n_0 \sim V$ diverges in contract to a *macroscopic* value.

> **Note:** The Bose-Einstein condensation is characterized by divergences in occupation numbers. The ground state is however the only state with a *macroscopic* occupation number.

Experimental verification. The Bose-Einstein condensation was predicted by Satyendra Bose and Albert Einstein in 1924-1925. It took almost 70 years to have an experimental corroboration of this phenomenon with the ultracold gas systems. Previous experiments had been done with ^{4}He as well as with hydrogen.

5.7 THEORY OF LIQUID HELIUM

Liquid helium is a physical state of helium at very low temperatures at standard atmospheric pressures. Liquid helium may show superfluidity.

At standard pressure, the chemical element helium exists in a liquid form only at the extremely low temperature of −269 °C (−452.20 °F; 4.15 K). Its boiling point and critical point depend on which isotope of helium is present: the common isotope helium-4 or the rare isotope helium-3. These are the only two stable isotopes of helium. See the table below for the values of these physical quantities. The density of liquid helium-4 at its boiling point and a pressure of one atmosphere (101.3 kilopascals) is about 125 g/L (0.125 g/ml), or about one-eighth the density of liquid water.

Liquefaction: Helium was first liquefied on July 10, 1908, by the Dutch physicist Heike Kamerlingh Onnes at the University of Leiden in the Netherlands. At that time, helium-3 was unknown because the mass spectrometer had not yet been invented. In more recent decades, liquid helium has been used as a cryogenic refrigerant (which is used in cryocoolers), and liquid helium is produced commercially for use in superconducting magnets

such as those used in Magnetic Resonance Imaging (MRI), Nuclear Magnetic Resonance (NMR), Magnetoencephalography (MEG), and experiments in physics, such as low temperature Mössbauer spectroscopy.

Liquefied helium-3: A helium-3 atom is a fermion and at very low temperatures, they form two-atom Cooper pairs which are bosonic and condense into a superfluid. These Cooper pairs are substantially larger than the interatomic separation.

Characteristics of Critical Indices

The temperature required to produce liquid helium is low because of the weakness of the attractions between the helium atoms. These interatomic forces in helium are weak to begin with because helium is a noble gas, but the interatomic attractions are reduced even more by the effects of quantum mechanics. These are significant in helium because of its low atomic mass of about four atomic mass units. The zero point energy of liquid helium is less if its atoms are less confined by their neighbors. Hence in liquid helium, its ground state energy can decrease by a naturally occurring increase in its average interatomic distance. However at greater distances, the effects of the interatomic forces in helium are even weaker.

Because of the very weak interatomic forces in helium, the element remains a liquid at atmospheric pressure all the way from its liquefaction point down to absolute zero. Liquid helium solidifies only under very low temperatures and great pressures. At temperatures below their liquefaction points, both helium-4 and helium-3 undergo transitions to superfluids. (Refer the Table 5.1 below.)

Liquid helium-4 and the rare helium-3 are not completely miscible. Below 0.9 kelvin at their saturated vapor pressure, a mixture of the two isotopes undergoes a phase separation into a normal fluid (mostly helium-3) that floats on a denser superfluid consisting mostly of helium-4. [Citation needed] This phase separation happens because the overall mass of liquid helium can reduce its thermodynamic enthalpy by separating. At extremely low temperatures, the superfluid phase, rich in helium-4, can contain up to 6% of helium-3 in solution. This makes the small-scale use of the dilution refrigerator possible, which is capable of reaching temperatures of a few millikelvins.

Superfluid helium-4 has substantially different properties from ordinary liquid helium.
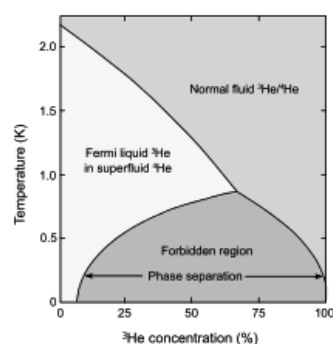
Fig. 5.12 *Liquid Helium 3 and 4 Isotopes in Phase Diagram, Showing the Demixing zone*

History of Critical Indices: In 1908 the Dutch physicist Kamerlingh-Onnes succeeded in liquifying a small quantity of helium. In 1923 he provided advice to the Canadian physicist John Cunningham McLennan who was the first to produce quantities of liquid helium almost on demand. Important early work on the characteristics of liquid helium was done by the Soviet physicist Lev Landau, later extended by the American physicist Richard Feynman.

Data of Liquid Helium: Superfluidity in liquid helium is possible. Only at the extraordinarily low temperature of −269 °C (−452.20 °F; 4.15 K) does the chemical element helium exist in a liquid form at standard pressure.

Table 5.1 *Data of Liquid Helium*

Properties of liquid helium	Helium-4	Helium-3
Critical temperature[3]	5.2 K (−267.95 °C)	3.3 K (−269.85 °C)
Boiling point at one atmosphere[3]	4.2 K (−268.95 °C)	3.2 K (−269.95 °C)
Minimum melting pressure[7]	25 bar (360 psi)	29 bar (420 psi) at 0.3 K (−272.850 °C)
Superfluid transition temperature at saturated vapor pressure	2.17 K (−270.98 °C)[8]	1 mK in the absence of a magnetic field[9]

5.8 FERMI DIRAC STATISTICS

The basic postulates of FD statistics are:

(i) Particles are identical and indistinguishable.

(ii) Total energy and total number of particles of the entire system is constant.

(iii) Particles have half-integral spin, i.e., $\frac{1}{2}\hbar, \frac{3}{2}\hbar, \frac{37}{2}\hbar$, etc.

(iv) Particles obey Pauli's exclusion principle, i.e., no two particles in a single system can have the same value for each of the four quantum numbers. In other words, a single energy state can contain at best a single particle with appropriate spin.

(v) The wave function of the system is anti-symmetric under the positional exchange of any two particles.

Example: Electron, proton, neutron, all hyperons (Λ, Σ, Ξ, Ω) etc., these are known as *Fremions*.

Fermi-Dirac Distribution Law

Consider a system of N indistinguishable, non-interacting particles obeying Pauli's exclusion principle. Let N_1, N_2, N_3, ..., N_i, ..., N_n particles in the system have energies E_1, E_2, E_3, ..., E_i, ..., E_n, respectively and let g_i is the number of degenerate quantum state in the energy level E_i. According to Pauli's exclusion principle, a single quantum state can be occupied by at best one particle. Since N_i particles are to be distributed among g_i degenerate states ($g_i \geq N_i$) having the same energy E_i, N_i states will be filled up and ($g_i - N_i$) states will remain vacant. Now g_i states can be arranged in gi! possible ways. But since the particles and the quantum states are indistinguishable, we have to deduct N_i! ways and ($g_i - N_i$)! ways from the all possible ways to

get effective number of arrangements. Thus, total number of possible ways of arrangement for the ith state is

$$W_i = \frac{(g_i)!}{N_i!(g_i - N_i)!} \qquad ...(5.65)$$

Hence the total number of ways for the entire distribution of N particles in n number of energy levels of the system is

$$W = \prod_i^n \frac{(g_i)!}{N_i!(g_i - N_i)!} \qquad ...(5.66)$$

where Π denotes the product symbol.

Now taking natural logarithm on both sides of eqn. (5.66) and applying Stirling approximation, we get,

$$\ln W = \sum_i [\ln g_i - \ln N_i - \ln(g_i - N_i)]$$

$$= \sum_i [g_i \ln g_i - N_i \ln N_i - (g_i - N_i)\ln(g_i - N_i)] \qquad ...(5.67)$$

Now, differentiating eqn. (5.67) to obtain the most probable distribution, we get,

$$d(\ln W)_{max} = \sum_i [-\ln N_i + \ln(g_i - N_i)]d N_i = 0 \qquad ...(5.68)$$

Considering the conservation of total energy and total number of particles, we can write

$$\sum_i dN_i = 0 \text{ (conservation of total no. of particles)} \qquad ...(5.69)$$

$$\sum_i E_i dN_i = 0 \text{ (conservation of total energy)} \qquad ...(5.70)$$

Multiplying eqn. (5.69) by a and eqn. (5.70) by b and then adding to eqn. (5.69) we get,

$$\sum_i [-\ln N_i + \ln(g_i - N_i) - \alpha - \beta E_i]dN_i = 0 \qquad ...(5.71)$$

Since dN_i's are independent of one another, the expression in the bracket in the eqn. (6.71) is zero for each N_i. Thus

$$N_i = \frac{g_i}{e^{(\alpha+\beta E_i)} + 1} \qquad ...(5.72)$$

Hence the Fermi-Dirac distribution function is given by

$$f(E_i) = \frac{N_i}{g_i} = \frac{1}{e^{(\alpha+\beta E_i)} + 1} \qquad ...(5.73)$$

which represents the probability of finding a fermion with energy Ei.

5.8.1 Energy and Pressure of Ideal fermi Gas

Fermi Distribution at Zero and Non-zero Temperatures

The Fremi-Dirac distribution function is

$$f(E_i) = \frac{N_i}{g_i} = \frac{1}{e^{(\alpha+\beta E_i)} + 1}$$

Now for the fermions in statistical equilibrium at absolute temperature

T, it is found that $\alpha = \dfrac{-E_F}{kT}$ and $\beta = \dfrac{1}{kT}$

where, E_F is the *Fremi energy* of the system (i.e., the maximum energy at absolute zero) k is the Boltzmann constant = 1.38×10^{-23} Joule/Kelvin.

Here $\dfrac{N_i}{g_i}$ is known as the occupation index.

Thus, the *F-D* distribution function reduces to

$$(E_i) = \dfrac{1}{e^{(E_i - E_F)/kT} + 1} \qquad \qquad ...(5.74)$$

Now, let us discuss two distinct cases.

Case – 1 ($T = 0$ K)

$$f(E_i) = 1, \text{ when } E_i < E_F$$
$$= 0, \text{ when } E_i > E^F$$

Thus at $T = 0\,K, f(E_i)$ is a step function (see Fig. 7.3) which implies all the energy states up to $E_i = E_F$ are filled up and the states above E_F are empty.
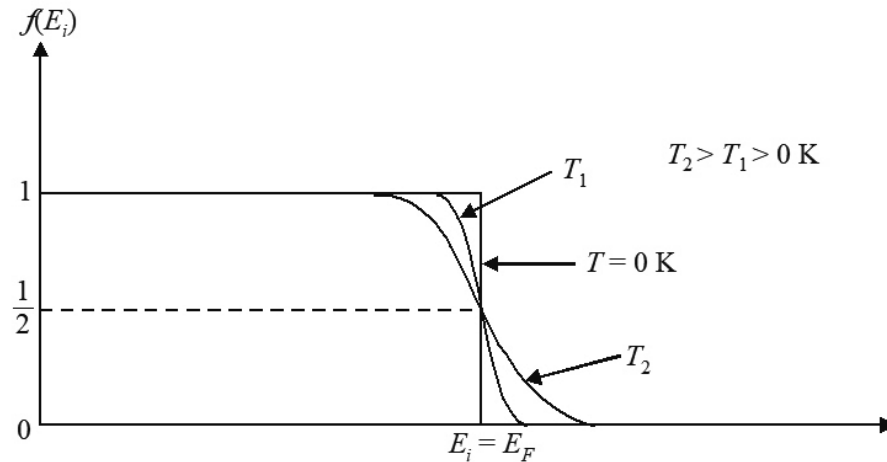
Fig. 5.13 *Plot of Fermi-Dirac distribution function as a function of Ei*

Case – 2 ($T > 0$ K)

As the temperature increases beyond $T = 0$ K, some of the electrons in the level just below the Fermi level go to levels just above EF resulting a gradual change in the occupation index $f(E_i)$. If $E_i = E_F$, $(E_i) = \dfrac{1}{e^0 + 1} = \dfrac{1}{2}$, i.e., at a finite temperature Fermi energy of a system is that energy for which 50% energy states are filled and 50% states are vacant.

Fermi Energy for Electron Gas in Metals

Metals are characterised by the presence of a good number of free electrons. These electrons move about at random within the metals. While moving the free electrons collide among themselves also encounter with the fixed ion cores. Such behaviour of the free electrons is similar to that of molecules of a gas. Thus metals can be treated as electron gas. Since the electrons have half-integral spin angular momenta and they obey Pauli's exclusion principle, they follow F-D statistics.

Let us consider an electron gas consisting of N electrons occupying volume V. If N is very large, the spacing between two successive energy levels becomes very small making the energy levels almost continuous. Then if the energy of the electrons ranges between E to $E + dE$, the number of degenerate states gi should be replaced by $g(E)\, dE$ and the total number of electrons Ni in these states should be replaced by $N(E)\, dE$ in the distribution function given by equation (5.72). Thus the total number of electrons in the electron gas having energy between E to $E + dE$ can be written eqn.(5.72) and eqn. (5.74) as

$$N(E)dE = \frac{g(E)\,dE}{\exp\left(\dfrac{E_i - E_F}{kT}\right) + 1} \qquad ...(5.75)$$

Since the electrons have two allowed values of spin quantum number $\left(m_s = \pm\dfrac{1}{2}\right)$, the total number of their allowed states between energy for the energy range E to $E + dE$ can be written as

$$g(E)dE = 2 \times \frac{1}{h^3}(4\pi V \sqrt{2E}\, m^{3/2})\, dE \qquad ...(5.76)$$

Hence the density of states g(E) for a Fermionic gas is given by

$$g(E) = \frac{1}{h^3}(8\pi V \sqrt{2mE}\, m) \qquad ...(5.77)$$

which shows that $g(E)$ depends only on E for a single type of fermionic gas kept in a fixed volume V.

From Fig. (5.13) it is obvious that at $T = 0$ K, all the single-particle state up to energy E_F are filled up. Thus at $T = 0$ K,

$$N(E)dE = \frac{g(E)\,dE}{e^{-\infty} + 1} = g(E)dE \quad [e^{-\infty} = 0] \qquad ...(5.78)$$

i.e., total number of electrons is equal to the total number of single-particle energy states, hence

$$N = \int_0^{E_F} N(E)\, dE$$

$$= \int_0^{E_F} g(E)\, dE \qquad\qquad \text{[using eqn. (5.78)]}$$

$$= \frac{1}{h^3}(8\sqrt{2}\pi V m^{3/2})\int_0^{E_F} E^{1/2} dE \qquad\qquad \text{[using eqn. (5.77)]}$$

or, $\qquad N = \dfrac{16\sqrt{2}\pi V m^{3/2}}{3h^3} E_F^{3/2} \qquad ...(5.79)$

which is the expression for the total number of electrons in the metal at

$T = 0$ K.

Therefore,

$$E_F = \frac{h^2}{8m}\left(\frac{3N}{\pi V}\right)^{2/3} \qquad ...(5.80)$$

which is the expression for Fermi energy of the electrons in the metal.

If $n = N/V$ denotes the concentration of the electrons in the metal, the Fermi energy of the electrons in the metal is

$$E_F = \frac{h^2}{8m}\left(\frac{3n}{\pi}\right)^{2/3} \qquad \qquad ...(5.81)$$

which showns that Fermi energy of the electrons depends solely on their concentration.

Total Energy at Absolute Zero Temperature

The total energy of electrons at absolute zero ($T = 0$ K) is given by

$$(E_0)_{\text{tot}} = \int_0^{E_F} E.N(E)\, dE$$

$$= \int_0^{E_F} E.g(E)\, dE \qquad \qquad \text{[using eqn. (5.81)]}$$

$$= \frac{1}{h^3}(8\sqrt{2}\pi V m^{3/2})\int_0^{E_F} E^{3/2} dE \qquad \qquad \text{[using eqn. (5.77)]}$$

$$= \frac{16\sqrt{2}\pi V m^{3/2}}{5h^3} E_F^{5/2}$$

$$= \frac{3}{5} N E_F \qquad \qquad \text{[using eqn. (5.79)]} \qquad ...(5.82)$$

which shows quantum effect because, classically the total energy of a system at absolute zero is nothing but zero.

The average energy per electron at $T = 0$ K can be written as

$$<E_0> = \frac{(E_0)_{\text{tot}}}{N} = \frac{3}{5} E_F \qquad \qquad ...(5.83)$$

which shows that at absolute zero temperature, the average energy per electron is equal to $\frac{3}{5}$ times the Fermi energy.

The ground-state pressure of the system is defined by

$$P_0 = \frac{2E_0}{3V} = \frac{2}{5} n E_F \qquad \qquad \text{[using eqn. (5.82)]}$$

Substituting for EF, the foregoing expression takes the form

$$P_0 = \frac{2}{5}\frac{h^2}{8m}\left(\frac{3}{\pi}\right)^{2/3} n^{5/3} \propto n^{5/3}$$

Thus the ground-state pressure of an electron gas at absolute zero solely depends on its concentration.

Fermi Temperature (T^F)

Fermi temperature (T_F) is the temperature equivalent of Fermi energy (F_F) and it is defined as $T_F = \dfrac{E_F}{k}$, where k (= 1.38×10^{-23} Joule/Kelvin) is the Boltzmann constant.

5.9 FREE ELECTRON THEORY OF SOLIDS

By a free particle we mean a particle which moves freely in space without the influence of any force. Hence, for a free particle the potential energy is zero.

Restricting our discussion to motion in one-dimension, say, along the x-axis, we have $V(x) = 0$ for all values of x so that the wavefunction $\psi(x)$ describing the state of the particle of mass m and a total energy E satisfies the Schrödinger equation,

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

or $\qquad \dfrac{d^2\psi(x)}{dx^2} + \dfrac{2m}{\hbar^2}\,E\psi(x) = 0$

or $\qquad \dfrac{d^2\psi(x)}{dx^2} + k^2\,\psi(x) = 0 \hspace{3cm}$ ...(5.84)

Where $\qquad k^2 = \dfrac{2m}{\hbar^2}\,E \hspace{3.5cm}$ ...(5.85)

The most general solution of Equation (5.84) is a combination of two linearly independent plane wave solutions e^{ikx} and e^{-ikx}

$$\psi_k(x) = Ae^{ikx} + Be^{-ikx} \hspace{2.5cm} ...(5.86)$$

where A and B are arbitrary constants. The complete wavefunction is given by,

$$\psi_k(x,\,t) = Ae^{\,i(kx-\omega t)} + Be^{\,-i(kx+\omega t)} \hspace{2cm} ...(5.87)$$

Where $\qquad \omega = \dfrac{E}{\hbar} = \dfrac{\hbar k^2}{2m} \hspace{3cm}$ ...(5.88)

The first term in Equation (5.86) $\psi_+(x, t) = Ae^{i(kx-\omega t)}$ is a wave travelling along the positive x-axis while the second term $\psi_-(x, t) = Be^{-i(kx+\omega t)}$ represents a wave travelling along the negative x-axis. Both the waves $\psi_+(x, t)$ and $\psi_-(x, t)$ travelling along opposite directions are associated with the motion of the free partical having well defined momentum and energy. The momentum associated with $\psi_+(x, t)$ is $p_+ = \hbar k$ while that with $\psi_-(x, t)$ is $p_- = -\hbar k$. Both $\psi_+(x, t)$ and $\psi_-(x, t)$ belong to the energy $\dfrac{\hbar^2 k^2}{2m}$. Since for free particle motion there are no boundary conditions, there exist no restrictions on the values of k and E. Clearly the states of the free particle are continuous or unbound.

It is important to discuss some of the physical subtleties present in the free particle motion:

(1) The probability density corresponding to the solution $\psi_+(x, t)$ is,

$$P_+(x,t) = |\psi_+(x,t)|^2 = |A|^2 = \text{Constant independent of } x \text{ and } t$$

The probability density corresponding to the solution $\psi_-(x, t)$ is,

$$P_-(x, t) = |\psi_-(x, t)|^2 = |B|^2 = \text{Constant independent of } x \text{ and } t.$$

The above result is a purely quantum mechanical result having no explanation according to classical mechanics. Since the particle represented by the waves $\psi_+(x, t)$ and $\psi_-(x, t)$ have well defined momenta and energy we have the uncertainty in momentum $\Delta p = 0$ and uncertainty in energy $\Delta E = 0$. According to Heisenberg's uncertainty principle we get the uncertainty in the position $\Delta x \to \infty$ and the uncertainty in the time $\Delta t \to \infty$. Thus there is complete loss of information about the position and time for any state of the particle.

(2) The speed of the plane waves $\psi_+(x, t)$ and $\psi_-(x, t)$ is given by

$$v_{\text{wave}} = \frac{\omega}{k} = \frac{E}{\hbar k} = \frac{\hbar^2 k^2 / 2m}{\hbar k} = \frac{\hbar k}{2m} \qquad ...(5.89)$$

The speed of the particle according to classical mechanics is given by,

$$v_{\text{particle}} = \frac{p}{m} = \frac{\hbar k}{m} \qquad ...(5.90)$$

We thus observe,

$$v_{\text{particle}} = 2v_{\text{wave}} \qquad ...(5.91)$$

The above means that the particle travels with a speed which is double the speed of the waves representing the particle.

(3) The wavefunction representing the particle is not normalizable. This is because,

$$\int_{-\infty}^{+\infty} \psi_+^*(x, t)\,\psi_+(x, t)\, dx = |A|^2 \int_{-\infty}^{+\infty} dx = \infty \qquad ...(5.92)$$

And

$$\int_{-\infty}^{+\infty} \psi_-^*(x, t)\,\psi_-(x, t)\, dx = |B|^2 \int_{-\infty}^{+\infty} dx = \infty \qquad ...(5.93)$$

We may conclude from the above result that the solutions of the Schrödinger equation $\psi_+(x, t)$ and $\psi_-(x, t)$ do not represent physical situation because wavefunction representing the state of any system must be quadratically integrable. We may make a formal conclusion that a free particle described by the laws of quantum mechanics cannot have sharply defined momentum and energy. We may further conclude that a free partical cannot be represented by single (monochromatic) plane wave. Physically acceptable representation of a free particle is a wave packet. We may further conclude that solutions of the Schrödinger equation which are physically acceptable cannot be plane waves.

The Potential Step

Consider a particle of mass m moving in a one-dimensional potential specified shown in the Figure (5.14). Mathematically, the potential function $V(x)$ is of the form,

$$V(x) = 0 \qquad x < 0 \qquad \text{(Region I)}$$
$$\quad = V_0 \qquad x > 0 \qquad \text{(Region II)}$$

The particle moving freely in Region I encounters the potential V_0 at $x = 0$.
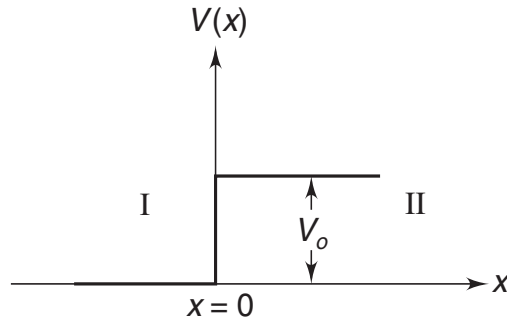
Fig. 5.14 *The Potential Step*

The wavefunction $\psi(x)$ describing the state of the particle in general satisfies the Schrödinger equation,

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

or
$$\frac{d^2\psi(x)}{dx^2} + \frac{2m}{\hbar^2}[E - V(x)]\,\psi(x) = 0 \qquad \ldots(5.94)$$

In the above, E is the total energy of the particle.

In the Region *I*, if $\psi_1(x)$ is the wavefunction, Equation (5.94) takes the form

$$\ldots(5.95)$$
$$\frac{d^2\psi_1(x)}{dx^2} + \frac{2m}{\hbar^2}\,E\,\psi_1(x) = 0$$

or
$$\frac{d^2\psi_1(x)}{dx^2} + k^2\,\psi_1(x) = 0$$

where
$$k^2 = \frac{2m}{\hbar^2}\,E \qquad\qquad \ldots(5.96)$$

If $\psi_2(x)$ be the wavefunction of the particle in Region II, Equation (5.94) gives,

$$\frac{d^2\psi_2(x)}{dx^2} + \frac{2m}{\hbar^2}[E - V_o]\,\psi_2(x) = 0$$

or
$$\frac{d^2\psi_2(x)}{dx^2} + \alpha^2\psi_2(x) = 0 \qquad \ldots(5.97)$$

Where
$$\alpha^2 = \frac{2m}{\hbar^2}[E - V_o] \qquad \ldots(5.98)$$

Most general solutions of Equations (5.95) and (5.96) can be written as,

$$\psi_1(x) = Ae^{ikx} + Be^{-ikx} \qquad ...(5.99)$$

$$\psi_2(x) = Ce^{i\alpha x} + De^{-i\alpha x} \qquad ...(5.100)$$

In the above A, B, C and D are constants which may be determined using the boundary conditions on the wavefunctions. The first term in Equation (5.99), $\psi_{1+}(x) = Ae^{ikx}$ represents a plane wave travelling along the positive x-axis in the Region I and can be considered as an *incident wave*, while the second term $\psi_{1-}(x) = Be^{-ikx}$ representing a plane wave in Region I travelling along the negative x-axis can be considered as the wave *reflected* at the potential step at $x = 0$. The first term in Equation (5.100) $\psi_{2+}(x) = Ce^{i\alpha x}$ represents a plane wave travelling in Region II along the positive x-axis and can be considered as the wave *transmitted* in Region II from the potential step at $x = 0$, while the second term $\psi_{2-}(x) = De^{-i\alpha x}$ represents a plane wave in Region II travelling along the negative x-axis. Since throughtout the Region II there exists no potential boundary from which reflection can occur, $\psi_{2-}(x)$ must vanish which requires D to be equal to zero so that Equation (5.100) reduces to,

$$\psi_2(x) = C e^{i\alpha x} \qquad ...(5.101)$$

We have the following boundary conditions in view of single valuedness and continuity of wavefunction at a potential boundary:

(i) $\qquad\qquad \psi_1(x) = \psi_2(x) \quad$ at $x = 0 \qquad ...(5.102)$

Using the above we obtain from Equations (5.99)

$$A + B = C \qquad ...(5.103)$$

(ii) $\qquad \dfrac{d\psi_1(x)}{dx} = \dfrac{d\psi_2(x)}{dx} \quad$ at $x = 0 \qquad ...(5.104)$

From Equation (5.99) we have

$$\frac{d\psi_1(x)}{dx} = ikAe^{ikx} - ikBe^{-ikx} \qquad ...(5.105)$$

From Equation (4.18) we have

$$\frac{d\psi_2(x)}{dx} = i\alpha C e^{i\alpha x} \qquad ...(5.106)$$

Using Equations (5.105) and (5.106) we get using Equation (5.104)

$$k(A - B) = \alpha C \qquad ...(5.107)$$

Solving Equations (5.103) and (5.107) we obtain,

$$C = \frac{2k}{k + \alpha} A \qquad ...(5.108)$$

$$B = \frac{k - \alpha}{k + \alpha} A \qquad ...(5.109)$$

If we consider the constant A as the amplitude of the incident wave, constants B and C can respectively be interpreted as the reflected and the transmitted amplitudes. We now consider the results on

reflection and transmission in two cases, namely when $E > V_0$ and when $E < V_0$.

Case 1: $E > V_0$

The wavefunction in Region I is,

$$\psi_1(x) = Ae^{ikx} + Be^{-ikx} \qquad \qquad ...(5.110)$$

We get on differentiating Equation (5.110) with respect to x

$$\frac{d\psi_1(x)}{dx} = ik[Ae^{ikx} - Be^{-ikx}] \qquad \qquad ...(5.111)$$

Taking complex conjugate, Equation (5.110) becomes

$$\psi_1^*(x) = A^*e^{-ikx} + B^*e^{ikx} \qquad \qquad ...(5.112)$$

Taking complex conjugate, Equation (5.111) we get

$$\frac{d\psi_1^*(x)}{dx} = -ik[A^*e^{-ikx} - B^*e^{ikx}] \qquad \qquad ...(5.113)$$

The general expression for probability current density is given by

$$J = \frac{-i\hbar}{2m}[\psi^*\nabla\psi - \psi\,\nabla\psi^*] \qquad \qquad ...(5.114)$$

Since we are considering one-dimensional motion we get from Equation (5.114) the probability current density in Region I to be

$$J_1 = \frac{-i\hbar}{2m}\left[\psi_1^*(x)\frac{d\psi_1(x)}{dx} - \psi(x)\frac{d\psi_1^*}{dx}\right]$$

Substituting from Equations (5.110), (5.111) (5.112), and (5.113) in the above we obtain

$$J_1 = \frac{\hbar k}{m}(|A|^2 - |B|^2) \qquad \qquad ...(5.115)$$

The first term on the right hand side of Equation (5.115) gives the probability current density of the incident wave / beam

$$(J_1)_{\text{incident}} = \frac{\hbar k}{m}|A|^2 \qquad \qquad ...(5.116)$$

while the second term gives the probability current density of the reflected wave / beam

$$(J_1)_{\text{reflected}} = \frac{\hbar k}{m}|B|^2 \qquad \qquad ...(5.117)$$

Let us now consider Region II in which the wavefunction is given by

$$\psi_2(x) = Ce^{i\alpha x} \qquad \qquad ...(5.118)$$

The above gives

$$\frac{d\psi_2(x)}{dx} = i\alpha\,Ce^{idx} \qquad \qquad ...(5.119)$$

Taking complex conjugate of Equation (5.119) we get

$$\psi_2^*(x) = C^*\,e^{-i\alpha x} \qquad \qquad ...(5.120)$$

and taking complex conjugate of Equation (5119) we get

$$\frac{d\psi_2^*(x)}{dx} = -i\alpha\, C^* e^{-i\alpha x} \qquad \qquad ...(5.121)$$

The probability current density in Region II by definition is given by

$$J_2 = \frac{-i\hbar}{2m}\left[\psi_2^*(x)\frac{d\psi_2}{dx} - \psi_2 \frac{d\psi_2^*}{dx}\right] \qquad \qquad ...(5.122)$$

Substituting from Equations (5.118), (5.119, (5.120) and (5.121) in Equation (5.122) we obtain

$$J_2 = \frac{\hbar\alpha}{m}|C|^2 \qquad \qquad ...(5.123)$$

Since in Region II, there exists only the transmitted wave we get the probability current density of the transmitted wave/beam,

$$(J_2)_{\text{transmitted}} = \frac{\hbar\alpha}{m}|C|^2 \qquad \qquad ...(5.124)$$

The *reflectance* or the *reflection coefficient* is, by definition, given by

$$R = \frac{\text{Probability current density for reflected beam}}{\text{Probability current density for incident beam}}$$

$$= \frac{(J_1)_{\text{reflected}}}{(J_1)_{\text{incident}}}$$

Using Equations (5.116) and (5.117) we obtain

$$R = \frac{\frac{\hbar k}{m}|B|^2}{\frac{\hbar k}{m}|A|^2} = \frac{|B|^2}{|A|^2}$$

Using Equation (5.109) in the above we obtain

$$R = \left(\frac{k-\alpha}{k+\alpha}\right)^2 \qquad \qquad ...(5.125)$$

Similarly, the transmittance or transmission coefficient is

$$T = \frac{(J_2)_{\text{transmitted}}}{(J_1)_{\text{incident}}}$$

Using Equations (5.116) and (5.124), the above gives

$$T = \frac{\frac{\hbar\alpha}{m}|C|^2}{\frac{\hbar k}{m}|A|^2} = \frac{\alpha}{k}\frac{|C|^2}{|A|^2}$$

Using Equation (5.108) in the above we get

$$T = \frac{\alpha}{k}\left(\frac{2k}{k+\alpha}\right)^2$$

or $\qquad T = \dfrac{4k\alpha}{(k + \alpha)^2}$ $\qquad\qquad$ ...(5.126)

We note the following:

1. We have $k = \sqrt{\dfrac{2m}{\hbar^2} E}$, a real positive quantity

   $\alpha = \sqrt{\dfrac{2m}{\hbar} (E - V_0)}$, a real positive quantity under the condition $E > V_0$

   Equation (5.125) then shows that R is a real positive quantity, meaning that a certain fraction of the incident particles gets reflected on encountering the potential step at $x = 0$. This result is in contrast to classical mechanics, according to which a particle going over a potential step, under the condition $E > V_0$, would slow down in order to conserve energy but would never be reflected. The observed result is a consequence of the wave properties of the particle. In other words, we can say that reflection under the condition $E > V_0$ is a *quantum mechanical effect*.

2. For $E \gg V_0$, that is for $\alpha \to k$ from below, the ratio of the reflected flux to the incident flux, that is, $|R|^2$ approaches zero. This agrees with intution which tells us that at very high incident energies, the presence of the step is but a small perturbation on the propagation of the wave.

Case 2: $\quad E < V_0$

In this case, α given by $\alpha = \sqrt{\dfrac{2m}{\hbar^2} (E - V_0)}$ becomes imaginary. We may write

$$\alpha = \sqrt{-\dfrac{2m}{\hbar^2} (V_0 - E)} = i\sqrt{\dfrac{2m}{\hbar^2} (V_0 - E)} = i\beta \qquad\qquad ...(5.127)$$

hence $\quad \beta = \sqrt{\dfrac{2m}{\hbar} (V_0 - E)}$ is real positive. $\qquad\qquad$ ...(5.128)

The solution of the Schrödinger equation in Region II is now given by

$$\psi_2(x) = C\, e^{ii\beta x} = C\, e^{-\beta x} \qquad\qquad ...(5.129)$$

We find that $\psi_2(x)$ does not blow up at $x = +\infty$

The reflection coefficient given by Equation (5.125), in this case becomes

$$R = \left(\dfrac{k - i\beta}{k + i\beta}\right)\left(\dfrac{k - i\beta}{k + i\beta}\right)^* = \left(\dfrac{k - i\beta}{k + i\beta}\right)\left(\dfrac{k + i\beta}{k - i\beta}\right) \qquad\qquad ...(5.130)$$

or $\qquad\qquad R = 1$

Thus, when $E < V_0$, as in classical mechanics there is total reflection.

It can, however, be seen that the transmission coefficient given by

Equation (5.126) does not vanish. Clearly, a part of the incident wave penetrates into the classically forbidden region, *Such penetration phenomenon again is characteristic of waves permitting a 'tunneling' through barriers that would totally block particles in classical description.*

Asymmetric Square Well

Consider a particle of mass m moving in a one-dimensional infinitely deep asymmetric potential well as shown in the Figure (5.15), the potential function $V(x)$ being of the form,

$$V(x) = +\infty$$
for $x < 0$ [Region I]

$$= 0$$
for $0 \leq x \leq a$ [Region II]

$$= +\infty$$
for $x > a$ [Region III]
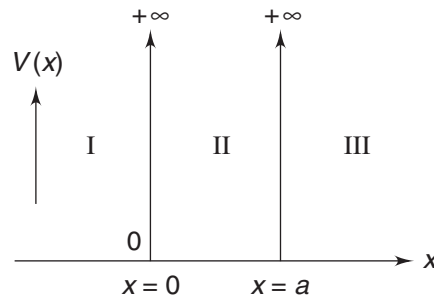
Fig. 5.15 *Asymmetric Square Well*

Classically, the particle remains confined within the well and moves with constant momentum back and forth as a result of repeated reflections from the walls of the well at $x = 0$ and at $x = a$.

Since $V(x) = +\infty$ for $x < 0$ (i.e., in Region I) as well as for $x > a$ (i.e., in Region III), the wavefunctions of the particle in these two regions are zero, i.e.,

$$\psi(x = 0) = 0 = \psi(x = a) \qquad \qquad ...(5.131)$$

If $\psi(x)$ represents the wavefunction for the particle inside the well ($0 \leq x \leq a$), we have the Schrödinger equation

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

or $$\frac{d^2\psi(x)}{dx^2} + k^2\psi(x) = 0 \qquad \qquad ...(5.132)$$

Where $$k = \sqrt{\frac{2m}{\hbar^2} E} \qquad \qquad ...(5.133)$$

The general solutions of Equation (5.132) are

$$\psi(x) = C\,e^{ikx} + De^{-ikx}$$

or $\qquad \psi(x) = A \sin kx + B \cos kx \qquad \qquad \ldots(5.134)$

where A and B are constants.

Using the boundary condition given by Equation (5.131), namely $\psi(0) = 0$ in Equation (5.134) we get

$$B = 0$$

so that the solution becomes

$$\psi(x) = A \sin kx \qquad \qquad \ldots(5.135)$$

Further, applying the other boundary condition namely $\psi(a) = 0$, we get from Equation (5.135)

$$A \sin ka = 0$$

The above gives either $A = 0$ or $\sin ka = 0$. However, $A = 0$ leads to $\psi(x) = 0$ everywhere which is not possible. Hence, we obtain

$$\sin ka = 0$$

The above gives

$$ka = n\pi; \quad n = \text{A positive integer}$$

$$= 1, 2, 3, \ldots \qquad \qquad \ldots(5.136)$$

We may note that n cannot be 0 because that would make $k = 0$ so that wavefunction would vanish everywhere.

From Equation (5.136) we thus get

$$k = \frac{n\pi}{a} \qquad \qquad \ldots(5.137)$$

Using Equation (5.137) in Equation (5.135) we get the energy eigenfunctions of the particle to be given by

$$\psi_n(x) = A \sin\left(\frac{n\pi}{a} x\right) \quad ; n = 1, 2, \ldots \qquad \ldots(5.138)$$

Constant A can be determined from the requirement that the eigenfunctions are normalized, i.e.,

$$\int_0^a \psi_n^*(x)\,\psi_n(x)\,dx = 1$$

The above gives,

$$A^2 \int_0^a \sin^2\left(\frac{n\pi}{a} x\right) dx = 1$$

or $\qquad A^2 \dfrac{a}{2} = 1$

or $\qquad A = \sqrt{\dfrac{2}{a}} \qquad \qquad \ldots(5.139)$

The energy eigenfunctions are thus

$$\psi_n(x) = \sqrt{\frac{2}{a}}\,\sin\left(\frac{n\pi}{a}x\right); \quad n = 1, 2, \dots \qquad \dots(5.140)$$

Using Equation (5.133) in Equation (5.137) we get the energy eigenvalues of the particle to be given by

$$\sqrt{\frac{2m}{\hbar^2}E_n} = \frac{n\pi}{a}$$

or

$$E_n = \frac{\pi^2\hbar^2}{2ma^2}n^2, \quad n = 1, 2, \dots \qquad \dots(5.141)$$

We find the energy to be quantized, only certain values of energy are permitted. This is as expected because the states of a particle which are confined within a limited region of space are bound states and the energy eigenvalue spectrum is discrete. This result is in sharp contrast to the result in classical physics in which the energy of the particle given by $E = \dfrac{p^2}{2m}$ (p being the momentum of the particle) can assume any value continuously from a minimum to a maximum.

From Equation (5.141) we get

$$E_{n+1} - E_n = \frac{\hbar^2\pi^2}{2ma^2}(2n+1) \qquad \dots(5.142)$$

Clearly, the energy levels are not equispaced.

We have

$$\frac{E_{n+1} - E_n}{E_n} = \frac{2n+1}{n^2} = \frac{2}{n} + \frac{1}{n^2}$$

Clearly, in the classical limit, the above gives,

$$\underset{n\to\infty}{\mathrm{Lt}}\left(\frac{E_{n+1} - E_n}{E_n}\right) = \underset{n\to\infty}{\mathrm{Lt}}\left(\frac{2}{n} + \frac{1}{n^2}\right) = 0 \qquad \dots(5.143)$$

meaning that the levels become so close together that they become practically indistinguishable forming a continuum.

The lowest energy state or the ground state corresponds to $n = 1$. The ground state energy is given by,

$$E_1 = \frac{\hbar^2\pi^2}{2ma^2} \qquad \dots(5.144)$$

And the ground state wavefunction is given by,

$$\psi_1(x) = \sqrt{\frac{2}{a}}\,\sin\left(\frac{\pi}{a}x\right) \qquad \dots(5.145)$$

Energy given by Equation (5.144) is called the zero point energy because there exists no state with zero energy.

The plot of some of the eigenfunctions with x are shown in Figure (5.16). We observe from the plots that the eigenfunction $\psi_n(x)$ has $(n-1)$ nodes.
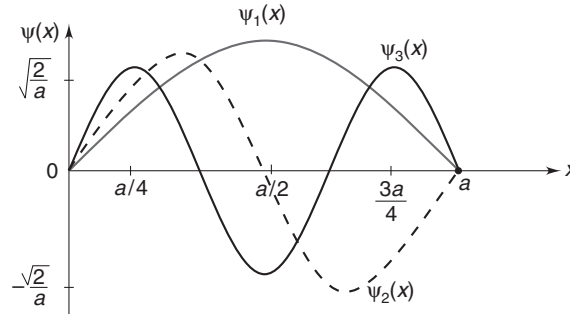
Fig. 5.16 *Plot of Engenfunctions*

Discussion on Zero Point Energy

If the particle inside the well has zero energy then it will come to rest and will be localized within the limited region defining the well. Heisenberg's uncertainty relation then will require the particle to acquire a finite momentum and hence a minimum kinetic energy. Since the particle is confined in the region $0 < x < a$, it has a maximum position uncertainty $\Delta x = a$ and hence a minimum momentum uncertainty $\Delta p \sim \dfrac{\hbar}{a}$ which in turn corresponds to a minimum kinetic energy $\dfrac{(\Delta p)^2}{2m} = \dfrac{\hbar^2}{2ma^2}$ which is in qualitative agreement with the exact value $E_1 = \dfrac{\pi^2 \hbar^2}{2ma^2}$.

The minimum momentum uncertainty given by $\Delta p \sim \dfrac{\hbar}{a}$, is inversely proportional to the width of the well. Smaller the width, more the particle becomes localized, and Δp increases. This causes the particle to move faster thereby increasing the zero point energy. If on the other hand, width of the well increases, the zero point energy decreases but never becomes zero. Thus localization of a particle forces a minimum motion and hence a minimum energy to the particle.

Symmetric Square-Well Potential of Infinite Depth

A symmetric infinite square well potential is defined as

$$V(x) = +\infty \quad \text{for} \quad x < -a$$
$$= 0 \quad \text{for} \quad -a \le x \le a$$
$$= +\infty \quad \text{for} \quad x > a$$

and is represented in the Figure (5.17)

Consider the motion of a particle of mass m in the one-dimensional potential described above.
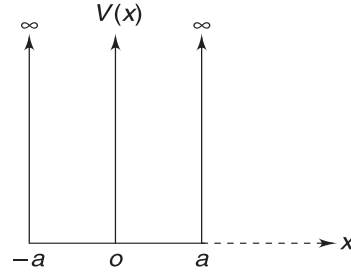
Fig. 5.17 *Symmetric Infinite Square*

If $\psi(x)$ is the wavefunction describing the state of the particle in the region $-a \leq x \leq a$ then it satisfies the time-independent Schrödinger equation,

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

or
$$\frac{d^2\psi(x)}{dx^2} + k^2\psi(x) = 0 \qquad\qquad ...(5.146)$$

Where
$$k = \sqrt{\frac{2m}{\hbar^2} E} \qquad\qquad ...(5.147)$$

The most general solution of Equation (5.146) is given by,

$$\psi(x) = A\,\sin(kx) + B\,\cos(kx) \qquad\qquad ...(5.148)$$

Where A and B are constants.

Since $V(x) = \infty$ for $x < -a$ and $x > a$, the wavefunctions in these two regions vanish giving,

$$\psi(-a) = 0 \quad \text{and} \quad \psi(+a) = 0 \qquad\qquad ...(5.149)$$

Using the conditions given by Equation (5.149) in Equation (5.148) we get

$$A\,\sin ka + B\,\cos ka = 0 \qquad\qquad ...(5.150)$$

and

$$-A\,\sin ka + B\,\cos ka = 0 \qquad\qquad ...(5.151)$$

For the above two equations to hold simultaneously we must have

$$A\,\sin ka = 0 \qquad\qquad ...(5.152)$$

and $\quad B\,\cos ka = 0 \qquad\qquad ...(5.153)$

In view of Eqations (5.151) and (5.153) we may have $A = 0$ and $B = 0$ but these are physically unacceptable because $\psi(x)$ given by Equation (5.148) would then vanish.

Since $B \neq 0$, we have from Equation (5.153),

$$\cos ka = 0 = \cos \frac{n\pi}{2}, \quad n = 1, 3, 5, \ldots$$

or $\qquad ka = \frac{n\pi}{2} \quad$ or $\quad k = \frac{n\pi}{2a}$ ...(5.154)

Using Equation (5.147) in the above we obtain the energy eigenvalues

$$E_n = \frac{\hbar^2}{2m}\frac{n^2\pi^2}{a^2} = \frac{\pi^2\hbar^2}{8ma^2}n^2, \quad n = 1, 3, 5, \ldots \qquad \ldots(5.155)$$

The energy eigenfunctions corresponding to the above energy eigenvalues are

$$\psi_n(x) = B \cos kx = B\cos\left(\frac{n\pi}{2a}x\right), \quad n = 1, 3, 5 \qquad \ldots(5.156)$$

The condition given by Equation (5.152) gives

$$\sin ka = 0 = \sin n\pi \qquad (\text{since } A \neq 0)$$

or $\qquad ka = n\pi \quad$ or $\quad k = \frac{n\pi}{a}, \quad n = 2, 4, 6$ ...(5.157)

Using the above value of k in Equation (5.147) we get the energy eigenvalues

$$E_n = \frac{\hbar^2}{2m}k^2 = \frac{\pi^2\hbar^2 n^2}{8ma^2}, \quad n = 2, 4, 6, \ldots \qquad \ldots(5.158)$$

The corresponding energy eigenfunctions are

$$\psi_n(x) = A\sin\left(\frac{n\pi x}{2a}\right), \quad n = 2, 4, 6, \ldots \qquad \ldots(5.159)$$

The normalization conditions of the wavefunctions,

$$\int_{-a}^{+a} \psi_n^*(x)\,\psi_n(x)\,dx = 1$$

lead to

$$A = \frac{1}{\sqrt{a}}, \quad B = \frac{1}{\sqrt{a}} \qquad \ldots(5.160)$$

We can thus write the set of energy eigenfunctions for the particle in the symmetric infinite square well potential as,

$$\psi_n(x) = \frac{1}{\sqrt{a}}\sin\left(\frac{n\pi}{2a}x\right), \quad n = 2, 4, \ldots \qquad \ldots(5.161)$$

$$\psi_n(x) = \frac{1}{\sqrt{a}}\cos\left(\frac{n\pi}{2a}x\right), \quad n = 1, 3, 5, \ldots \qquad \ldots(5.162)$$

and the discrete set of energy eigenvalues as

$$E_n = \frac{\pi^2\hbar^2}{8ma^2}n^2, \quad n = 1, 2, 3, \ldots \qquad \ldots(5.163)$$

Discussion

The wavefunctions corresponding to $n = 1, 3, 5, ...$, i.e., corresponding to odd quantum numbers are symmetric, $\psi(-x) = \psi(x)$.

- The wavefunctions corresponding to $n = 2, 4, 6, ...$ i.e., corresponding to even quantum numbers are antisymmetric, $\psi(-x) = -\psi(x)$

- In other words, for symmetric potentials $V(-x) = V(x)$, the wavefunctions of bound states are either even (symmetric) or odd (antisymmetric).

- The energy spectrum for the particle is discrete and non-degenerate.

- The ground state energy or the zero point energy is,

$$E_1 = \frac{\pi^2 \hbar^2}{8ma^2}$$

Corresponding to the eigenfunction,

$$\psi_1(x) = \frac{1}{\sqrt{a}} \cos \frac{\pi x}{2a}$$

Symmetric Square-Well Potential of Finite Depth

A symmetric square well potential of finite depth is described by potential function $V(x)$ of the form

$$V(z) = V_0 \quad \text{for} \quad x < -a \qquad \text{(Region I)}$$
$$= 0 \text{for} \quad -a \leq x \leq a \qquad \text{(Region II)}$$
$$= V_0 \quad \text{for} \quad x > a \qquad \text{(Region III)}$$

The potential function is shown in the Figure (5.18)
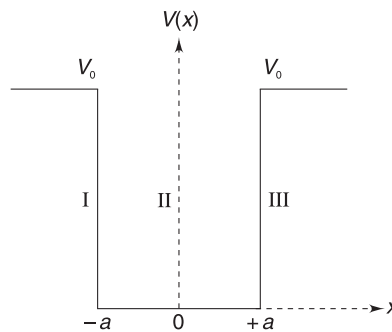
Fig. 5.18 *Potential Function*

Consider the motion of a particle of mass m in the potential well described above.

The Schrödinger equation in Regions I and III is,

$$\frac{-\hbar^2}{2m} \frac{d^2 \psi(x)}{dx^2} + V_0 \psi(x) = E \psi(x)$$

or $$\frac{d^2\psi(x)}{dx^2} + \frac{2m}{\hbar^2}(E - V_0)\,\psi(x) = 0 \qquad \text{...(5.164)}$$

In Region II the Schrödinger equation is

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

which can be put in the form

$$\frac{d^2\psi(x)}{dx^2} + k^2\psi(x) = 0 \qquad \text{...(5.165)}$$

where $$k = \sqrt{\frac{2m}{\hbar^2}E} \qquad \text{...(5.166)}$$

Let us consider the cases where: $E < V_0$, and $E > V_0$

Case $E < V_0$: We may write Equation (5.164) in the form

$$\frac{d^2\psi(x)}{dx^2} - \alpha^2\psi(x) = 0 \qquad \text{...(5.167)}$$

where $\alpha = \sqrt{\dfrac{2m}{\hbar^2}(V_0 - E)}$ is real positive $\qquad \text{...(5.168)}$

The most general solution of Equation (5.167) is

$$\psi(x) = Ae^{+\alpha x} + Be^{-\alpha x}, \quad A \text{ and } B \text{ are constants} \qquad \text{...(5.169)}$$

Specific solution in Region I $\psi_1(x) = A\,e^{\alpha x} \qquad \text{...(5.170)}$

Specific solution in Region III $\psi_3(x) = B\,e^{-\alpha x} \qquad \text{...(5.171)}$

Solution of Equation (5.165) gives the wavefunction in region II

$$\psi_2(x) = C\sin(kx) + D\cos(kx) \qquad \text{...(5.172)}$$

$\psi_2(x)$ is either symmetric or antisymmetric about $x = 0$. The first term in Equation (5.172) is antisymmetric because $\sin(kx) = -\sin(-kx)$. The second term is symmetric because $\cos(kx) = \cos(-kx)$.

For the *symmetric function* in Region II, the coefficient $C = 0$ so that we may write the symmetric wavefunction in Region II as

$$(\psi_2(x))_{\text{symmetric}} = D\cos(kx) \qquad \text{...(5.173)}$$

At $x = \pm a$, we have, using the single valuedness of wavefunction

$$Ae^{-\alpha a} = D\cos ka \qquad \text{...(5.174)}$$

$$Be^{-\alpha a} = D\cos ka \qquad \text{...(5.175)}$$

Similarly, using the continuity of wavefunction at $x = \pm a$ gives

$$\alpha A\,e^{-\alpha a} = +D\,k\sin(ka) \qquad \text{...(5.176)}$$

$$+ \alpha B \, e^{-\alpha a} = D \, k \sin (ka) \qquad \qquad \text{...(5.177)}$$

The above equations give

$$A = B \qquad \qquad \text{...(5.178)}$$

and $\qquad k \tan ka = \alpha \qquad \qquad \text{...(5.179)}$

Let us now consider the antisymmetric wavefunctions in Region II. In the case $D = 0$ so that we may write the antisymmetric wavefunction in Region II as

$$(\psi_2(x))_{\text{antisymmetric}} = C \sin (kx) \qquad \qquad \text{...(5.180)}$$

Using the single valuedness and continuity of wavefunction at the boundaries at $x = \pm a$ we get

$$A \, e^{-\alpha a} = - C \sin (ka) \qquad \qquad \text{...(5.181)}$$

$$B \, e^{-\alpha a} = C \sin (ka) \qquad \qquad \text{...(5.182)}$$

$$\alpha A \, e^{-\alpha a} = - C \, k \cos (ka) \qquad \qquad \text{...(5.183)}$$

$$- \alpha B e^{-\alpha a} = C \, k \cos (ka) \qquad \qquad \text{...(5.184)}$$

From the above four equations, we obtain

$$A = - B \qquad \qquad \text{...(5.185)}$$

and

$$k \cot ka = - \alpha \qquad \qquad \text{...(5.186)}$$

The energy eigenvalues for the particle can be obtained by solving Equations (5.177) and (5.186) graphically as explained in the following:

Let us put $\qquad ka = x \qquad \qquad \text{...(5.187)}$

$$\alpha a = y \qquad \qquad \text{...(5.188)}$$

From the above we get

$$x^2 + y^2 = (k^2 + \alpha^2) \, a^2$$

Substituting for k and α from Equations (5.166) and (5.168), the above becomes

$$x^2 + y^2 = \left[\frac{2m}{\hbar^2} E + \frac{2m}{\hbar^2} (V_0 - E) \right] a^2$$

or $\qquad \qquad x^2 + y^2 = \dfrac{2ma^2}{\hbar^2} V_0 \qquad \qquad \text{...(5.189)}$

Substituting Equations (5.187) and (5.188) in Eqations (5.179) and (5.186), respectively, we obtain

$$x \tan x = y \qquad \qquad \text{...(5.190)}$$

$$-x \cot x = y \qquad \qquad \text{...(5.191)}$$

We plot $x \tan x$ against x, $x \cot x$ against x and $x^2 + y^2$ for different values of $V_0 a^2$ (which are circles of different radii). Since both x and y can take only positive values, the sections of the circles have been shown in the first quadrant only in Figure (5.19).
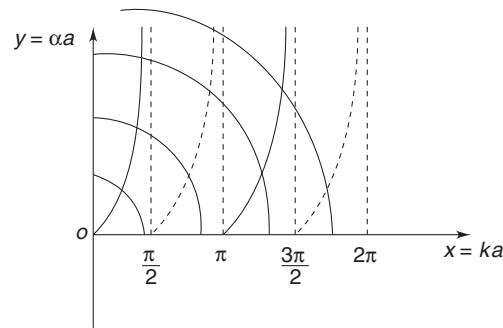
Fig. 5.19 *Graph*

In the Figure 5.19,

Full line curves $\rightarrow x \tan x$ against x plots

Dashed curves $\rightarrow -x \cot x$ against x plots

Circular sections $\rightarrow$ Different values of $V_0 a^2$

The energy levels and the energy eigenvalues for the symmetric wavefunction are given by the intersections of the $x \tan x$ against x curves and the circular sections. Similarly, the energy eigenfunctions and the energy eigenvalues when the wavefunction in the well is antisymmetric are given by the intersections of $-x \cot x$ against x curves and the circular sections.

If the intersections of $x \tan x$ against x curves and circles occur at values of x equal to $x_1, x_2, ..., x_n, ...$ then we get

$$x_n^2 = k^2 a^2 = \frac{2m E_n}{\hbar^2} a^2$$

or $\qquad E_n = \frac{\hbar^2}{2ma^2} x_n^2, \quad n = 1, 3, 5, ...$

Similarly, if the intersections of $-x \cot x$ against x curves and the circles occur at values of x equal to $x_1, x_2, ... x_{n'}, ...,$ then we get

$$x_{n'}^2 = k^2 a^2 = \frac{2mE_{n'}}{\hbar^2} a^2$$

or $\qquad E_{n'} = \frac{\hbar^2}{2ma^2} x_{n'}^2; \quad n' = 2, 4, 6, ...$

The number of bound states are seen to depend upon the height V_0 and the width a of the well through the factor $V_0 a^2$. From the Figure (5.19) we find the following

(i) Only one energy level of symmetric type, if $0 < V_0 a^2 < \dfrac{\pi^2 \hbar^2}{8m}$

(ii) Two energy levels of which one is of symmetric type and the other of antisymmetric type, if $\dfrac{\pi^2 \hbar^2}{8m} < V_0 a^2 < \dfrac{4\pi^2 \hbar^2}{8m}$

(iii) Three energy levels of which two are of symmetric type and one of antisymmetric type, if $\dfrac{4\pi^2\hbar^2}{8m} < V_0 a^2 < \dfrac{9\pi^2\hbar^2}{8m}$ and so on.

Some of the energy eigenfunctions corresponding to bound states are shown in the Figure (5.20).
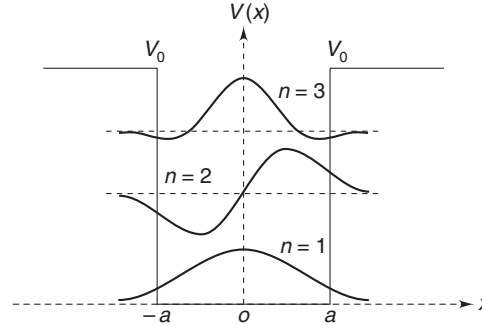
Fig. 5.20 Eigenfunctions for Bound State

Unlike in the case of infinite potential well, both the symmetric as well as the antisymmetric eigenfunctions extend beyond the well, i.e., in the regions $x < -a$ and $x > a$ which define the classical turning points. Clearly, there exists finite probability of finding the particle outside the well. This is a quantum mechanical effect.

Case $E > V_0$:

The Schordinger equation in Regions I and III is given by,

$$\frac{d^2\psi(x)}{dx^2} + \frac{2m}{\hbar^2}(E - V_0)\,\psi(x) = 0$$

Since E is greater than V_0, $\dfrac{2m}{\hbar^2}(E - V_0)$ is a real positive quantity. As such the solution of the above equation is sinusoidal in nature. The probability density for the particle is distributed over all space in regions I and III. It is also distributed in Region II, i.e., within the well. Thus we do not get bound state for the particle.

5.10 LANDAU THEORY OF PHASE TRANSITION

Landau theory in physics is a theory that Lev Landau introduced in an attempt to formulate a general theory of continuous (i.e., second-order) phase transitions. It can also be adapted to systems under externally-applied fields, and used as a quantitative model for discontinuous (i.e., first-order) transitions.

Mean-Field Formulation (no long-range correlation)

Landau was motivated to suggest that the free energy of any system should obey two conditions:

- It is analytic.
- It obeys the symmetry of the Hamiltonian.

Given these two conditions, one can write down (in the vicinity of the critical temperature, T_c) a phenomenological expression for the free energy as a Taylor expansion in the order parameter.

Second-Order Transitions

Consider a system that breaks some symmetry below a phase transition, which is characterized by an order parameter η. This order parameter is a measure of the order before and after a phase transition; the order parameter is often zero above some critical temperature and non-zero below the critical temperature. In a simple ferromagnetic system like the Ising model, the order parameter is characterized by the net magnetization m, which becomes spontaneously non-zero below a critical temperature T_c. In Landau theory, one considers a free energy functional that is an analytic function of the order parameter. In many systems with certain symmetries, the free energy will only be a function of even powers of the order parameter, for which it can be expressed as the series expansion

$$F(T, \eta) - F_0 = a(T)\eta^2 + \frac{b(T)}{2}\eta^4 + \cdots$$

In general, there are higher order terms present in the free energy, but it is a reasonable approximation to consider the series to fourth order in the order parameter, as long as the order parameter is small. For the system to be thermodynamically stable (that is, the system does not seek an infinite order parameter to minimize the energy), the coefficient of the highest even power of the order parameter must be positive, so $b(T) > b_0$, a constant, near the critical temperature. Furthermore, since $a(T)$ changes sign above and below the critical temperature, one can likewise expand $a(T) \approx a_0(T - T_c)$, where it is assumed that $a > 0$ for the high-temperature phase while $a < 0$ or the low-temperature phase, for a transition to occur. With these assumptions, minimizing the free energy with respect to the order parameter requires

$$\frac{\partial F}{\partial \eta} = 2a(T)\eta + 2b(T)\eta^3 = 0$$

The solution to the order parameter that satisfies this condition is either $\eta = 0$, or

$$\eta_0^2 = -\frac{a}{b} = -\frac{a_0}{b_0}(T - T_c)$$

It is clear that this solution only exists for $T < Tc$, otherwise $\eta = 0$ is the only solution. Indeed, $\eta = 0$ is the minimum solution for $T > Tc$, but the solution h_0 minimizes the free energy for $T < Tc$, and thus is a stable phase. Furthermore, the order parameter follows the relation

$$\eta(T) \propto |T - T_c|^{1/2}$$

below the critical temperature, indicating a critical exponent $\beta = 1/2$ for this Landau mean-theory model.

The free-energy will vary as a function of temperature given by

$$F - F_0 = \begin{cases} -\dfrac{a_0^2}{2b_0}(T - T_c)^2, & T < T_c \\ 0, & T > T_c \end{cases}$$

From the free energy, one can compute the specific heat,

$$c_p = -T\frac{\partial^2 F}{\partial T^2} = \begin{cases} \dfrac{a_0^2}{b_0}T, & T < T_c \\ 0, & T > T_c \end{cases}$$

which has a finite jump at the critical temperature of size $\Delta c = a_0^2 T_c / b_0$.

This finite jump is therefore not associated with a discontinuity that would occur if the system absorbed latent heat, since T_c DS = 0. It is also noteworthy that the discontinuity in the specific heat is related to the discontinuity in the second derivative of the free energy, which is characteristic of a second-order phase transition. Furthermore, the fact that the specific heat has no divergence or cusp at the critical point indicates its critical exponent for $c \sim |T - T_c|^{-\alpha}$ is $\alpha = 0$.

Applied Fields

In many systems, one can consider a perturbing field *h* that couples linearly to the order parameter. For example, in the case of a classical dipole moment µ, the energy of the dipole-field system is -µB. In the general case, one can assume an energy shift of -m*h* due to the coupling of the order parameter to the applied field *h*, and the Landau free energy will change as a result:

$$F(T,\eta) - F_0 = a_0(T - T_c)\eta^2 + \frac{b_0}{2}\eta^4 - \eta h$$

In this case, the minimization condition is

$$\frac{\partial F}{\partial \eta} = 2a(T)\eta + 2b_0\eta^3 - h = 0$$

One immediate consequence of this equation and its solution is that, if the applied field is non-zero, then the magnetization is non-zero at any temperature. This implies there is no longer a spontaneous symmetry breaking that occurs at any temperature. Furthermore, some interesting thermodynamic and universal quantities can be obtained from this above condition. For example, at the critical temperature where $a(T_c) = 0$, one can find the dependence of the order parameter on the external field:

$$\eta(T_c) = \left(\frac{h}{2b_0}\right)^{1/3} \propto h^{1/\delta}$$

indicating a critical exponent $\delta = 3$.

Furthermore, from the above condition, it is possible to find the zero-field susceptibility $\chi \equiv \partial\eta/\partial h|_{h=0}$.

which must satisfy

$$0 = 2a\frac{\partial \eta}{\partial h} + 6b\eta^2 \frac{\partial \eta}{\partial h} - 1$$

$$[2a + 6b\eta^2]\frac{\partial \eta}{\partial h} = 1$$

In this case, recalling in the zero-field case that $\eta^2 = - a / b$ at low temperatures, while $\eta^2 = 0$ for temperatures above the critical temperature, the zero-field susceptibility therefore has the following temperature dependence:

$$\chi(T, h \to 0) = \begin{cases} \frac{1}{2a_0(T-T_c)}, & T > T_c \\ \frac{1}{-4a_0(T-T_c)}, & T < T_c \end{cases} \propto |T - T_c|^{-\gamma}$$

which is reminiscent of the Curie-Weiss law for the temperature dependence of magnetic susceptibility in magnetic materials, and yields the mean-field critical exponent $\lambda = 1$.

First-Order Transitions

Landau theory can also be used to study first-order transitions. There are two different formulations, depending on whether or not the system is symmetric under a change in sign of the order parameter.

Symmetric Case

Here we consider the case where the system has a symmetry and the energy is invariant when the order parameter changes sign. A first-order transition will arise if the quartic term in F is negative. To ensure that the free energy remains positive at large η, one must carry the free-energy expansion to sixth-order,

$$F(T, \eta) = A(T)\eta^2 - B_0\eta^4 + C_0\eta^6.$$

Nonsymmetric Case

Next we consider the case where the system does not have a symmetry. In this case there is no reason to keep only even powers of η in the expansion of F. And a cubic term must be allowed (The linear term can always be eliminated by a shift $\eta \to + \chi$ constant.) We thus consider a free energy functional

$$A(T) = A_0(T - T_0), \text{ and } A_0, B_0, C_0$$

Once again $F(T, \eta) = A(T)\eta^2 - C_0\eta^3 + B_0\eta^4 + \cdots.$ are all positive. The sign of the cubic term can always be chosen to be negative as we have done by reversing the sign of η if necessary.

Applications of Landau Theory of Phase Transition

It was known experimentally that the liquid–gas coexistence curve and the ferromagnet magnetization curve both exhibited a scaling relation of the form $|T - Tc|^\beta$, where β was mysteriously the same for both systems. This is

the phenomenon of universality. It was also known that simple liquid–gas models are exactly mappable to simple magnetic models, which implied that the two systems possess the same symmetries. It then followed from Landau theory why these two apparently disparate systems should have the same critical exponents, despite having different microscopic parameters. It is now known that the phenomenon of universality arises for other reasons (see Renormalization group). In fact, Landau theory predicts the incorrect critical exponents for the Ising and liquid–gas systems.

The great virtue of Landau theory is that it makes specific predictions for what kind of non-analytic behavior one should see when the underlying free energy is analytic. Then, all the non-analyticity at the critical point, the critical exponents, are because the equilibrium value of the order parameter changes non-analytically, as a square root, whenever the free energy loses its unique minimum.

The extension of Landau theory to include fluctuations in the order parameter shows that Landau theory is only strictly valid near the critical points of ordinary systems with spatial dimensions higher than 4. This is the upper critical dimension, and it can be much higher than four in more finely tuned phase transition. In Mukhamel's analysis of the isotropic Lifschitz point, the critical dimension is 8. This is because Landau theory is a mean field theory, and does not include long-range correlations.

This theory does not explain non-analyticity at the critical point, but when applied to superfluid and superconductor phase transition, Landau's theory provided inspiration for another theory, the Ginzburg–Landau theory of superconductivity.

Including Long-Range Correlations

Consider the Ising model free energy above. Assume that the order parameter Ψ and external magnetic field, H, may have spatial variations. Now, the free energy of the system can be assumed to take the following modified form:

$$F := \int d^D x \ \left(a(T) + r(T)\psi^2(x) + s(T)\psi^4(x) \ + f(T)(\nabla\psi(x))^2 \ + h(x)\psi(x) \ + \mathcal{O}(\psi^6; (\nabla\psi)^4) \right)$$

where D is the total spatial dimensionality. So,

$$\langle \psi(x) \rangle := \frac{\mathrm{Tr} \ \psi(x)\mathrm{e}^{-\beta H}}{Z}$$

Assume that, for a localized external magnetic perturbation $h(x) \to 0 + h_0 \delta(x)$, the order parameter takes the form $\psi(x) \to \psi_0 + \phi(x)$. Then,

$$\frac{\delta\langle\psi(x)\rangle}{\delta h(0)} = \frac{\phi(x)}{h_0} = \beta\left(\langle\psi(x)\psi(0)\rangle - \langle\psi(x)\rangle\langle\psi(0)\rangle\right)$$

That is, the fluctuation $\phi(x)$ in the order parameter corresponds to the order-order correlation. Hence, neglecting this fluctuation (like in the earlier mean-field approach) corresponds to neglecting the order-order correlation, which diverges near the critical point.

One can also solve for ϕ (*x*) from which the scaling exponent, *v*, for correlation length $\xi \sim (T - T_c)^{-\nu}$ can deduced. From these, the Ginzburg criterion for the upper critical dimension for the validity of the Ising mean-field Landau theory (the one without long-range correlation) can be calculated as:

$$D \geq 2 + 2\frac{\beta}{\nu}$$

In our current Ising model, mean-field Landau theory gives β = 1/2 = *v* and so, it (the Ising mean-field Landau theory) is valid only for spatial dimensionality greater than or equal to 4 (at the marginal values of *D*=4, there are small corrections to the exponents). This modified version of mean-field Landau theory is sometimes also referred to as the Landau–Ginzburg theory of Ising phase transitions. As a clarification, there is also a Landau-Ginzburg theory specific to superconductivity phase transition, which also includes fluctuations.

5.11 CRITICAL INDICES

Critical exponents describe the behavior of physical quantities near continuous phase transitions. It is believed, though not proven, that they are universal, i.e., they do not depend on the details of the physical system, but only on some of its general features. For instance, for ferromagnetic systems, the critical exponents depend only on:

- The dimension of the system
- The range of the interaction
- The spin dimension

These properties of critical exponents are supported by experimental data. Analytical results can be theoretically achieved in mean field theory in high dimensions or when exact solutions are known such as the two-dimensional Ising model. The theoretical treatment in generic dimensions requires the renormalization group approach or the conformal bootstrap techniques. Phase transitions and critical exponents appear in many physical systems such as water at the liquid-vapor transition, in magnetic systems, in superconductivity, in percolation and in turbulent fluids. The critical dimension above which mean field exponents are valid varies with the systems and can even be infinite. It is 4 for the liquid-vapor transition, 6 for percolation and probably infinite for turbulence. Mean field critical exponents are also valid for random graphs, such as Erdős–Rényi graphs, which can be regarded as infinite dimensional systems.

Definition of Critical Indices: The control parameter that drives phase transitions is often temperature but can also be other macroscopic variables like pressure or an external magnetic field. For simplicity, the following discussion works in terms of temperature; the translation to another control parameter is straightforward. The temperature at which the transition occurs is called the critical temperature Tc. We want to describe the behavior of a

physical quantity f in terms of a power law around the critical temperature; we introduce the reduced temperature

$$\tau := \frac{T - T_c}{T_c}$$

which is zero at the phase transition, and define the critical exponent k:

$$k \overset{\text{def}}{=} \lim_{\tau \to 0} \frac{\log |f(\tau)|}{\log |\tau|}$$

This results in the power law we were looking for:

$$f(\tau) \propto \tau^k, \quad \tau \to 0$$

It is important to remember that this represents the asymptotic behavior of the function f(τ) as τ → 0.

More generally one might expect

The most important critical exponents

Let us assume that the system has two different phases characterized by an order parameter Ψ, which vanishes at and above Tc.

Consider the disordered phase (τ > 0), ordered phase (τ < 0) and critical temperature (τ = 0) phases separately. Following the standard convention, the critical exponents related to the ordered phase are primed. It is also another standard convention to use superscript/subscript + (−) for the disordered (ordered) state. In general spontaneous symmetry breaking occurs in the ordered phase.

Definitions

Ψ	order parameter (e.g. $\frac{\rho - \rho_c}{\rho_c}$ for the liquid–gas critical point, magnetization for the Curie point, etc.)
τ	$\frac{T - T_c}{T_c}$
f	specific free energy
C	specific heat; $-T\frac{\partial^2 f}{\partial T^2}$
J	source field (e.g. $\frac{P - P_c}{P_c}$ where P is the pressure and P_c the critical pressure for the liquid-gas critical point, reduced chemical potential, the magnetic field H for the Curie point)
χ	the susceptibility, compressibility, etc.; $\frac{\partial \psi}{\partial J}$
ξ	correlation length
d	the number of spatial dimensions
$\langle \psi(x)\, \psi(y) \rangle$	the correlation function
r	spatial distance

The following entries are evaluated at J = 0 (except for the δ entry)

Critical exponents for $\tau > 0$ (disordered phase)	
Greek letter	relation
α	$C \propto \tau^{-\alpha}$
γ	$\chi \propto \tau^{-\gamma}$
ν	$\xi \propto \tau^{-\nu}$

Critical exponents for $\tau < 0$ (ordered phase)	
Greek letter	relation
α'	$C \propto (-\tau)^{-\alpha'}$
β	$\Psi \propto (-\tau)^{\beta}$
γ'	$\chi \propto (-\tau)^{-\gamma'}$
ν'	$\xi \propto (-\tau)^{-\nu'}$

Critical exponents for $\tau = 0$	
Greek letter	relation
δ	$J \propto \psi^{\delta}$
η	$\langle \psi(0)\, \psi(r) \rangle \propto r^{-d+2-\eta}$

The critical exponents can be derived from the specific free energy f(J,T) as a function of the source and temperature. The correlation length can be derived from the functional F[J;T].

These relations are accurate close to the critical point in two- and three-dimensional systems. In four dimensions, however, the power laws are modified by logarithmic factors. These do not appear in dimensions arbitrarily close to but not exactly four, which can be used as a way around this problem.

Mean Field Critical Exponents of Ising-Like Systems

The classical Landau theory (also known as mean field theory) values of the critical exponents for a scalar field (of which the Ising model is the prototypical example) are given by

If we add derivative terms turning it into a mean field Ginzburg–Landau theory, we get

One of the major discoveries in the study of critical phenomena is that mean field theory of critical points is only correct when the space dimension of the system is higher than a certain dimension called the upper critical dimension which excludes the physical dimensions 1, 2 or 3 in most cases. The problem with mean field theory is that the critical exponents do not depend on the space dimension. This leads to a quantitative discrepancy below the critical dimensions, where the true critical exponents differ from the mean field values. It can even lead to a qualitative discrepancy at low space dimension, where a critical point in fact can no longer exist, even though mean field theory still predicts there is one. This is the case for the Ising model in dimension 1 where there is no phase transition. The space dimension where mean field theory becomes qualitatively incorrect is called the lower critical dimension.

Experimental Values

The most accurately measured value of α is −0.0127 for the phase transition of superfluid helium (the so-called lambda transition). The value was measured on a space shuttle to minimize pressure differences in the sample. This value is in a significant disagreement with the most precise theoretical determinations coming from high temperature expansion techniques, Monte Carlo methods and the conformal bootstrap.

Theoretical Predictions

Critical exponents can be evaluated via Monte Carlo simulations of lattice models. The accuracy of this first principle method depends on the available computational resources, which determine the ability to go to the infinite volume limit and to reduce statistical errors. Other techniques rely on theoretical understanding of critical fluctuations. The most widely applicable technique is the renormalization group. The conformal bootstrap is a more recently developed technique, which has achieved unsurpassed accuracy for the Ising critical exponents.

Scaling Functions

In light of the critical scalings, we can reexpress all thermodynamic quantities in terms of dimensionless quantities. Close enough to the critical point, everything can be reexpressed in terms of certain ratios of the powers of the reduced quantities. These are the scaling functions.

The origin of scaling functions can be seen from the renormalization group. The critical point is an infrared fixed point. In a sufficiently small neighborhood of the critical point, we may linearize the action of the renormalization group. This basically means that rescaling the system by a factor of a will be equivalent to rescaling operators and source fields by a factor of aΔ for some Δ. So, we may reparameterize all quantities in terms of rescaled scale independent quantities.

Scaling Relations

It was believed for a long time that the critical exponents were the same above and below the critical temperature, e.g., $\alpha \equiv \alpha'$ or $\gamma \equiv \gamma'$. It has now been shown that this is not necessarily true: When a continuous symmetry is explicitly broken down to a discrete symmetry by irrelevant (in the renormalization group sense) anisotropies, then the exponent's γ and γ' are not identical.

Critical exponents are denoted by Greek letters. They fall into universality classes and obey the scaling and hyperscaling relations

$$\nu d = 2 - \alpha = 2\beta + \gamma = \beta(\delta + 1) = \gamma \frac{\delta + 1}{\delta - 1}$$

$$2 - \eta = \frac{\gamma}{\nu} = d \frac{\delta - 1}{\delta + 1}$$

These equations imply that there are only two independent exponents, e.g., ν and η. All this follows from the theory of the renormalization group.

Anisotropy

There are some anisotropic systems where the correlation length is direction dependent. Directed percolation can be also regarded as anisotropic percolation. In this case the critical exponents are different and the upper critical dimension is 5.

Multicritical Points

More complex behavior may occur at multicritical points, at the border or on intersections of critical manifolds. They can be reached by tuning the value of two or more parameters, such as temperature and pressure.

Static Versus Dynamic Properties

The above examples exclusively refer to the static properties of a critical system. However dynamic properties of the system may become critical, too. Especially, the characteristic time, τ_{char}, of a system diverges as $\tau_{char} \propto \xi$ z, with a dynamical exponent z. Moreover, the large static universality classes of equivalent models with identical static critical exponents decompose into smaller dynamical universality classes, if one demands that also the dynamical exponents are identical. The critical exponents can be computed from conformal field theory.

Transport Properties

Critical exponents also exist for transport quantities like viscosity and heat conductivity. A recent study suggests that critical exponents of percolation play an important role in urban traffic.

Self-Organized Criticality

Critical exponents also exist for self organized criticality for dissipative systems.

Percolation Theory

Phase transitions and critical exponents appear also in percolation processes where the concentration of occupied sites or links play the role of temperature. The simplest example is perhaps percolation in a two dimensional square lattice. Sites are randomly occupied with probability p. For small values of p the occupied sites form only small clusters. At a certain threshold pc a giant cluster is formed, and we have a second-order phase transition. Refer percolation critical exponents. For percolation the critical exponents are different from Ising. For example, in the mean field $\delta = 2$for percolation compared to $\delta = 3$ for Ising. In network theory, the strength of interactions between communities has been found to behave analogous to an external field in magnets near the phase transition or as ghost field in percolation.

5.12 SCALE TRANSFORMATION

In physics, mathematics and statistics, scale invariance is a feature of objects or laws that do not change if scales of length, energy, or other variables, are multiplied by a common factor, and thus represent a universality. The technical term for this transformation is a dilatation (also known as dilation), and the dilatations can also form part of a larger conformal symmetry.

- In mathematics, scale invariance usually refers to an invariance of individual functions or curves. A closely related concept is self-similarity, where a function or curve is transformation under a discrete subset of the dilations. It is also possible for the probability distributions of random processes to display this kind of scale invariance or self-similarity.

- In classical field theory, scale invariance most commonly applies to the invariance of a whole theory under dilatations. Such theories typically describe classical physical processes with no characteristic length scale.

- In quantum field theory, scale invariance has an interpretation in terms of particle physics. In a scale-transformation theory, the strength of particle interactions does not depend on the energy of the particles involved.

- In statistical mechanics, scale invariance is a feature of phase transitions. The key observation is that near a phase transition or critical point, fluctuations occur at all length scales, and thus one should look for an explicitly scale-transformation theory to describe the phenomena. Such theories are scale-transformation statistical field theories, and are formally very similar to scale-transformation quantum field theories.

- Universality is the observation that widely different microscopic systems can display the same behaviour at a phase transition. Thus phase transitions in many different systems may be described by the same underlying scale-transformation theory.

- In general, dimensionless quantities are scale transformation. The analogous concept in statistics are standardized moments, which are scale transformation statistics of a variable, while the unstandardized moments are not.

Scale-Transformation Curves and Self-Similarity

In mathematics, one can consider the scaling properties of a function or curve $f(x)$ under rescalings of the variable x. That is, one is interested in the shape of $f(\lambda x)$ for some scale factor λ, which can be taken to be a length or size rescaling. The requirement for $f(x)$ to be transformation under all rescalings is usually taken to be

$$f(\lambda x) = \lambda^\Delta f(x)$$

for some choice of exponent Δ, and for all dilations λ. This is equivalent to f being a homogeneous function of degree Δ.

Examples of scale-transformation functions are the monomials for which $\Delta = n$, in that clearly

$$f(\lambda x) = (\lambda x)^n = \lambda^n f(x) \,.$$

An example of a scale-transformation curve is the logarithmic spiral, a kind of curve that often appears in nature. In polar coordinates (r, θ), the spiral can be written as

$$\theta = \frac{1}{b} \ln(r/a) \,.$$

Allowing for rotations of the curve, it is transformation under all rescalings λ; that is, $\theta(\lambda r)$ is identical to a rotated version of $\theta(r)$.

Projective Geometry

The idea of scale invariance of a monomial generalizes in higher dimensions to the idea of a homogeneous polynomial, and more generally to a homogeneous function. Homogeneous functions are the natural denizens of projective space, and homogeneous polynomials are studied as projective varieties in projective geometry. Projective geometry is a particularly rich field of mathematics; in its most abstract forms, the geometry of schemes, it has connections to various topics in string theory.

Fractals

It is sometimes said that fractals are scale-transformation, although more precisely, one should say that they are self-similar. A fractal is equal to itself typically for only a discrete set of values λ, and even then a translation and rotation may have to be applied to match the fractal up to itself. Thus, for example, the Koch curve scales with $\Delta = 1$, but the scaling holds only for values of $\lambda = 1/3n$ for integer n. In addition, the Koch curve scales not only at the origin, but, in a certain sense, 'Everywhere': miniature copies of itself

can be found all along the curve. Some fractals may have multiple scaling factors at play at once; such scaling is studied with multi-fractal analysis. Periodic external and internal rays are transformation curves.

Scale Invariance in Stochastic Processes

If P(f) is the average, expected power at frequency f , then noise scales as

$$P(f) = \lambda^{-\Delta} P(\lambda f)$$

with $\Delta = 0$ for white noise, $\Delta = -1$ for pink noise, and $\Delta = -2$ for Brownian noise (and more generally, Brownian motion). More precisely, scaling in stochastic systems concerns itself with the likelihood of choosing a particular configuration out of the set of all possible random configurations. This likelihood is given by the probability distribution.

Examples of scale-transformation distributions are the Pareto distribution and the Zipfian distribution.

Scale Transformation Tweedie Distributions

Tweedie distributions are a special case of exponential dispersion models, a class of statistical models used to describe error distributions for the generalized linear model and characterized by closure under additive and reproductive convolution as well as under scale transformation. These include a number of common distributions: the normal distribution, Poisson distribution and gamma distribution, as well as more unusual distributions like the compound Poisson-gamma distribution, positive stable distributions, and extreme stable distributions. Consequent to their inherent scale invariance Tweedie random variables Y demonstrate a variance var(Y) to mean E(Y) power law:

$$\text{var}(Y) = a[\text{E}(Y)]^p,$$

where a and p are positive constants. This variance to mean power law is known in the physics literature as fluctuation scaling, and in the ecology literature as Taylor's law.

Random sequences, governed by the Tweedie distributions and evaluated by the method of expanding bins exhibit a biconditional relationship between the variance to mean power law and power law autocorrelations. The Wiener–Khinchin theorem further implies that for any sequence that exhibits a variance to mean power law under these conditions will also manifest 1/f noise.

The Tweedie convergence theorem provides a hypothetical explanation for the wide manifestation of fluctuation scaling and 1/f noise. It requires, in essence, that any exponential dispersion model that asymptotically manifests a variance to mean power law will be required express a variance function that comes within the domain of attraction of a Tweedie model. Almost all distribution functions with finite cumulant generating functions qualify as exponential dispersion models and most exponential dispersion models manifest variance functions of this form. Hence many probability distributions

have variance functions that express this asymptotic behavior, and the Tweedie distributions become foci of convergence for a wide range of data types.

Much as the central limit theorem requires certain kinds of random variables to have as a focus of convergence the Gaussian distribution and express white noise, the Tweedie convergence theorem requires certain non-Gaussian random variables to express 1/f noise and fluctuation scaling.

Cosmology

In physical cosmology, the power spectrum of the spatial distribution of the cosmic microwave background is near to being a scale-transformation function. Although in mathematics this means that the spectrum is a power-law, in cosmology the term 'scale-transformation' indicates that the amplitude, P(k), of primordial fluctuations as a function of wave number, k, is approximately constant, i.e. a flat spectrum. This pattern is consistent with the proposal of cosmic inflation.

Scale Invariance in Classical Field Theory

Classical field theory is generically described by a field, or set of fields, φ, that depend on coordinates, x valid field configurations are then determined by solving differential equations for φ, and these equations are known as field equations.

For a theory to be scale-transformation, its field equations should be transformation under a rescaling of the coordinates, combined with some specified rescaling of the fields,

$$x \rightarrow \lambda x ,$$
$$\varphi \rightarrow \lambda^{-\Delta} \varphi .$$

The parameter Δ is known as the scaling dimension of the field, and its value depends on the theory under consideration. Scale invariance will typically hold provided that no fixed length scale appears in the theory. Conversely, the presence of a fixed length scale indicates that a theory is not scale-transformation. A consequence of scale invariance is that given a solution of a scale-transformation field equation, we can automatically find other solutions by rescaling both the coordinates and the fields appropriately. In technical terms, given a solution, $\varphi(x)$, one always has other solutions of the form

Scale Invariance of Field Configurations

For a particular field configuration, $\varphi(x)$, to be scale-transformation, we require that

$$\varphi(x) = \lambda^{-\Delta} \varphi(\lambda x)$$

where Δ is, again, the scaling dimension of the field.

We note that this condition is rather restrictive. In general, solutions even of scale-transformation field equations will not be scale-transformation, and in such cases the symmetry is said to be spontaneously broken.

Classical Electromagnetism

An example of a scale-transformation classical field theory is electromagnetism with no charges or currents. The fields are the electric and magnetic fields, E(x,t) and B(x,t), while their field equations are Maxwell's equations.

With no charges or currents, these field equations take the form of wave equations

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

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

where c is the speed of light. These field equations are transformation under the transformation

$$x \to \lambda x,$$

$$t \to \lambda t.$$

Moreover, given solutions of Maxwell's equations, E(x, t) and B(x, t), it holds that E(λx, λt) and B(λx, λt) are also solutions.

Massless Scalar Field Theory

Another example of a scale-transformation classical field theory is the massless scalar field (note that the name scalar is unrelated to scale invariance). The scalar field, φ(x, t) is a function of a set of spatial variables, x, and a time variable, t.

Consider first the linear theory. Like the electromagnetic field equations above, the equation of motion for this theory is also a wave equation,

$$\frac{1}{c^2} \frac{\partial^2 \varphi}{\partial t^2} - \nabla^2 \varphi = 0,$$

and is transformation under the transformation

$$x \to \lambda x,$$

$$t \to \lambda t.$$

The name massless refers to the absence of a term $\propto m^2 \varphi$ in the field equation. Such a term is often referred to as a 'Mass' term, and would break the invariance under the above transformation. In relativistic field theories, a mass-scale, m is physically equivalent to a fixed length scale through

$$L = \frac{\hbar}{mc},$$

and so it should not be surprising that massive scalar field theory is not scale-transformation.

φ4 Theory: The field equations in the examples above are all linear in the fields, which has meant that the scaling dimension, Δ, has not been so important. However, one usually requires that the scalar field action is dimensionless, and this fixes the scaling dimension of φ. In particular,

$$\Delta = \frac{D-2}{2},$$

where D is the combined number of spatial and time dimensions.

Given this scaling dimension for φ, there are certain nonlinear modifications of massless scalar field theory which are also scale-transformation. One example is massless φ4 theory for D = 4. The field equation is

$$\frac{1}{c^2}\frac{\partial^2 \varphi}{\partial t^2} - \nabla^2 \varphi + g\varphi^3 = 0.$$

(Note that the name φ4 derives from the form of the Lagrangian, which contains the fourth power of φ.)

When D = 4 (e.g., three spatial dimensions and one time dimension), the scalar field scaling dimension is Δ=1. The field equation is then transformation under the transformation

$$x \rightarrow \lambda x,$$
$$t \rightarrow \lambda t,$$
$$\varphi(x) \rightarrow \lambda^{-1}\varphi(x).$$

The key point is that the parameter g must be dimensionless, otherwise one introduces a fixed length scale into the theory: For φ4 theory, this is only the case in D = 4. Note that under these transformations the argument of the function φ is unchanged.

Scale Invariance in Quantum Field Theory

The scale-dependence of a Quantum Field Theory (QFT) is characterised by the way its coupling parameters depend on the energy-scale of a given physical process. This energy dependence is described by the renormalization group, and is encoded in the beta-functions of the theory. For a QFT to be scale-transformation, its coupling parameters must be independent of the energy-scale, and this is indicated by the vanishing of the beta-functions of the theory. Such theories are also known as fixed points of the corresponding renormalization group flow.

Quantum Electrodynamics

A simple example of a scale-transformation QFT is the quantized electromagnetic field without charged particles. This theory actually has no coupling parameters (since photons are massless and non-interacting) and is therefore scale-transformation, much like the classical theory.

However, in nature the electromagnetic field is coupled to charged particles, such as electrons. The QFT describing the interactions of photons and charged particles is Quantum Electrodynamics (QED), and this theory is not scale-transformation. We can see this from the QED beta-function. This tells us that the electric charge (which is the coupling parameter in the theory) increases with increasing energy. Therefore, while the quantized electromagnetic field without charged particles is scale-transformation, QED is not scale-transformation.

Massless Scalar Field Theory

Free, massless quantized scalar field theory has no coupling parameters. Therefore, like the classical version, it is scale-transformation. In the language of the renormalization group, this theory is known as the Gaussian fixed point. However, even though the classical massless $\varphi 4$ theory is scale-transformation in D = 4, the quantized version is not scale-transformation. We can see this from the beta-function for the coupling parameter, g.

Even though the quantized massless $\varphi 4$ is not scale-transformation, there do exist scale-transformation quantized scalar field theories other than the Gaussian fixed point. One example is the Wilson-Fisher fixed point, below.

Conformal Field Theory

Scale-transformation QFTs are almost always transformation under the full conformal symmetry, and the study of such QFTs is Conformal Field Theory (CFT). Operators in a CFT have a well-defined scaling dimension, analogous to the scaling dimension, Δ, of a classical field discussed above. However, the scaling dimensions of operators in a CFT typically differ from those of the fields in the corresponding classical theory. The additional contributions appearing in the CFT are known as anomalous scaling dimensions.

Scale and Conformal Anomalies

The $\varphi 4$ theory example above demonstrates that the coupling parameters of a quantum field theory can be scale-dependent even if the corresponding classical field theory is scale-transformation (or conformally transformation). If this is the case, the classical scale (or conformal) invariance is said to be anomalous. A classically scale transformation field theory, where scale invariance is broken by quantum effects, provides an explication of the nearly exponential expansion of the early universe called cosmic inflation, as long as the theory can be studied through perturbation theory.

Phase Transitions

In statistical mechanics, as a system undergoes a phase transition, its fluctuations are described by a scale-transformation statistical field theory. For a system in equilibrium (i.e., time-independent) in D spatial dimensions, the corresponding statistical field theory is formally similar to a D-dimensional CFT. The scaling dimensions in such problems are usually referred to as critical exponents, and one can in principle compute these exponents in the appropriate CFT.

The Ising Model

An example that links together many of the ideas in this article is the phase transition of the Ising model, a simple model of ferromagnetic substances. This is a statistical mechanics model, which also has a description in terms of conformal field theory. The system consists of an array of lattice sites, which form a D-dimensional periodic lattice. Associated with each lattice site is a magnetic moment, or spin, and this spin can take either the value +1 or −1. (These states are also called up and down, respectively.)

The key point is that the Ising model has a spin-spin interaction, making it energetically favourable for two adjacent spins to be aligned. On the other hand, thermal fluctuations typically introduce a randomness into the alignment of spins. At some critical temperature, Tc, spontaneous magnetization is said to occur. This means that below Tc the spin-spin interaction will begin to dominate, and there is some net alignment of spins in one of the two directions.

An example of the kind of physical quantities one would like to calculate at this critical temperature is the correlation between spins separated by a distance r. This has the generic behaviour:

$$G(r) \propto \frac{1}{r^{D-2+\eta}},$$

for some particular value of η, which is an example of a critical exponent.

CFT Description

The fluctuations at temperature T_c are scale-transformation, and so the Ising model at this phase transition is expected to be described by a scale-transformation statistical field theory. In fact, this theory is the Wilson-Fisher fixed point, a particular scale-transformation scalar field theory.

In this context, G(r) is understood as a correlation function of scalar fields,

$$\langle \phi(0)\phi(r) \rangle \propto \frac{1}{r^{D-2+\eta}}.$$

Now we can fit together a number of the ideas seen already.

From the above, one sees that the critical exponent, η, for this phase transition, is also an anomalous dimension. This is because the classical dimension of the scalar field,

$$\Delta = \frac{D-2}{2}$$

is modified to become

$$\Delta = \frac{D-2+\eta}{2},$$

where D is the number of dimensions of the Ising model lattice. So this anomalous dimension in the conformal field theory is the same as a particular critical exponent of the Ising model phase transition. Note that for dimension $D \equiv 4-\varepsilon$, η can be calculated approximately, using the epsilon expansion, and one finds that

$$\eta = \frac{\epsilon^2}{54} + O(\epsilon^3).$$

In the physically interesting case of three spatial dimensions, we have ε = 1, and so this expansion is not strictly reliable. However, a semi-quantitative prediction is that η is numerically small in three dimensions. On the other hand, in the two-dimensional case the Ising model is exactly soluble. In particular, it is equivalent to one of the minimal models, a family of well-understood CFTs, and it is possible to compute η (and the other critical exponents) exactly,

$$\eta_{D=2} = \frac{1}{4}.$$

Schramm–Loewner Evolution

The anomalous dimensions in certain two-dimensional CFTs can be related to the typical fractal dimensions of random walks, where the random walks are defined via Schramm–Loewner Evolution (SLE). As we have seen above, CFTs describe the physics of phase transitions, and so one can relate the critical exponents of certain phase transitions to these fractal dimensions. Examples include the 2d critical Ising model and the more general 2d critical Potts model. Relating other 2d CFTs to SLE is an active area of research.

Universality

A phenomenon known as universality is seen in a large variety of physical systems. It expresses the idea that different microscopic physics can give rise to the same scaling behaviour at a phase transition. A canonical example of universality involves the following two systems:

- The Ising model phase transition, described above.

- The liquid-vapour transition in classical fluids.

Even though the microscopic physics of these two systems is completely different, their critical exponents turn out to be the same. Moreover, one can calculate these exponents using the same statistical field theory. The key observation is that at a phase transition or critical point, fluctuations occur at all length scales, and thus one should look for a scale-transformation statistical field theory to describe the phenomena. In a sense, universality is the observation that there are relatively few such scale-transformation theories.

The set of different microscopic theories described by the same scale-transformation theory is known as a universality class. Other examples of systems which belong to a universality class are:

- Avalanches in piles of sand. The likelihood of an avalanche is in power-law proportion to the size of the avalanche, and avalanches are seen to occur at all size scales.

- The frequency of network outages on the Internet, as a function of size and duration.

- The frequency of citations of journal articles, considered in the network of all citations amongst all papers, as a function of the number of citations in a given paper.[citation needed]

- The formation and propagation of cracks and tears in materials ranging from steel to rock to paper. The variations of the direction of the tear, or the roughness of a fractured surface, are in power-law proportion to the size scale.

- The electrical breakdown of dielectrics, which resemble cracks and tears.

- The percolation of fluids through disordered media, such as petroleum through fractured rock beds, or water through filter paper, such as in chromatography. Power-law scaling connects the rate of flow to the distribution of fractures.

- The diffusion of molecules in solution, and the phenomenon of diffusion-limited aggregation.

- The distribution of rocks of different sizes in an aggregate mixture that is being shaken (with gravity acting on the rocks).

The key observation is that, for all of these different systems, the behaviour resembles a phase transition, and that the language of statistical mechanics and scale-transformation statistical field theory may be applied to describe them.

Other Examples of Scale Transformation

Newtonian Fluid Mechanics with No Applied Forces: Under certain circumstances, fluid mechanics is a scale-transformation classical field theory. The fields are the velocity of the fluid flow, u(x, t), the fluid density, ρ(x, t), and the fluid pressure, P(x, t). These fields must satisfy both the Navier–Stokes equation and the continuity equation. For a Newtonian fluid these take the respective forms

$$\rho\frac{\partial \mathbf{u}}{\partial t} + \rho\mathbf{u}\cdot\nabla\mathbf{u} = -\nabla P + \mu\left(\nabla^2\mathbf{u} + \frac{1}{3}\nabla\left(\nabla\cdot\mathbf{u}\right)\right)$$

$$\frac{\partial \rho}{\partial t} + \nabla\cdot(\rho\mathbf{u}) = 0$$

where μ is the dynamic viscosity.

In order to deduce the scale invariance of these equations we specify an equation of state, relating the fluid pressure to the fluid density. The equation

of state depends on the type of fluid and the conditions to which it is subjected. For example, we consider the isothermal ideal gas, which satisfies

where c_s is the speed of sound in the fluid. Given this equation of state, Navier–Stokes and the continuity equation are transformation under the transformations

$$x \rightarrow \lambda x,$$
$$t \rightarrow \lambda^2 t,$$
$$\rho \rightarrow \lambda^{-1} \rho,$$
$$\mathbf{u} \rightarrow \lambda^{-1} \mathbf{u}.$$

Given the solutions u(x, t) and ρ(x, t), we automatically have that λu (λx, λ²t) and ρ(λx, λ²t), are also solutions.

5.13 DIMENSIONAL ANALYSIS

Dimensional analysis is a branch of mathematics which deals with dimensions of quantities. Each physical phenomenon can be expressed by an equation that represents the relationship between the variables governing the phenomenon. In general, such quantities may be dimensional or non-dimensional (absolute numerals). Any physical variable can be described using qualitative and quantitative approaches. The qualitative description is known as dimension and quantitative description is known as unit.

The dimensions of basic quantities are selected as basic dimensions and dimensions of other variables are derived. Dimensions are classified as *Absolute system* (*MLT system*) and Gravitational System (*FLT System*),

Where   MLT is Mass, Length and Time and

FLT is Force, Length and Time.

Mass is independent of location and hence MLT system is known as absolute system. Force is based on gravitational acceleration which depends upon the location and hence FLT system is known as gravitational system.

Using Newton's second law,

In MLT system,

$$F = \frac{ML}{T^2} \qquad \qquad ...(5.192)$$

In FLT system,

$$M = \frac{FT^2}{L} \qquad \qquad ...(5.193)$$

The basic and derived dimensions of various quantities in MLT and FLT systems are shown in Table 5.2

Table 5.2 *Dimensions of quantities in MLT and FLT systems*

Sl. No	Quantity	MLT	FLT
1	Acceleration	LT^{-2}	LT^{-2}
2	Angular acceleration	T^{-2}	T^{-2}
3	Angular velocity	T^{-1}	T^{-1}
4	Area	L^2	L^2
5	Density	ML^{-3}	FT^3L^{-4}
6	Discharge	L^3T^{-1}	L^3T^{-1}
7	Energy	ML^2T^{-2}	FL
8	Force	MLT^{-2}	F
9	Impulse	MLT^{-1}	FT
10	Length	L	L
11	Mass	M	$FL^{-1}T^2$
12	Moment of a force	ML^2T^{-2}	FL
13	Moment of inertia (area)	L^4	L^4
14	Moment of inertia (mass)	ML^2	FLT^2
15	Momentum	MLT^{-1}	FT
16	Modulus of elasticity	$ML^{-1}T^{-2}$	FL^{-2}
17	Modulus of rigidity	$ML^{-1}T^{-2}$	FL^{-2}
18	Power	ML^2T^{-3}	FLT^{-1}
19	Pressure	$ML^{-1}T^{-2}$	FL^{-2}
20	Specific weight	$ML^{-2}T^{-2}$	FL^{-3}
21	Stress	$ML^{-1}T^{-2}$	FL^{-2}
22	Surface tension	MT^{-2}	FL^{-1}
23	Viscosity (absolute)	$ML^{-1}T^{-1}$	$FL^{-2}T$
24	Viscosity (kinematic)	L^2T^{-1}	L^2T^{-1}
25	Velocity	LT^{-1}	LT^{-1}
26	Volume	L^3	L^3
27	Weight	MLT^{-2}	F

Dimensional Homogeneity

An equation is said to be dimensionally homogeneous, if the dimensions of various terms on the left and right side of the equation are identical.

In other words, if $\delta = \alpha + \beta$ ...(5.194)

Dimension of δ = Dimension of α ≡ Dimension of β

If $\delta = \alpha (\kappa + r)$ ...(5.195)

where 'k' = constant,

then 'r' = dimensionless quantity

Therefore dimension of δ = Dimension of α.

Example 5.1: Verify whether the following equations are dimensionally homogeneous or not.

(a) $S = ut + \left(\dfrac{1}{2}\right) at^2$;

(b) $S = ut + \left(\dfrac{1}{2}\right)(9.81)t^2$;

(c) Curvature of a curve $y = f(x)$ is given by

$$\left(\frac{1}{\rho}\right) = \frac{\left(d^2 y / dx^2\right)}{\left[1 + (dy / dx)^2\right]^{3/2}}$$

where S is distance, u is velocity, t is time, a is acceleration and ρ is radius of curvature.

Solution:

(a) Substituting dimensions in the above eqn, we get

$$L=\left(LT^1\right)(T)+\left(LT^2\right)\left(T^2\right)$$
$$L \equiv L \equiv L$$

Hence it is dimensionally homogeneous.

(b) Substituting dimensions in the above eqn, we get

$$L=\left(LT^1\right)(T)+\left(T^2\right)$$
$$L \equiv L \not\equiv L^2$$

Hence it is dimensionally not homogeneous.

$$\frac{dy}{dx}=\frac{L}{L}=\text{constant}$$

(c) $\dfrac{d^2y}{dx^2}=\dfrac{d}{dx}\left(\dfrac{dy}{dx}\right)=\left(\dfrac{1}{L}\right)\left(\dfrac{L}{L}\right)=\left(\dfrac{1}{L}\right)$

$$\left(\frac{1}{L}\right)=\left(\frac{1}{L}\right)$$

Hence it is dimensionally homogeneous.

Dimensions of Unknown Variables

In some physical phenomena, dimensions of unknown variables can be determined using dimensional analysis. Similarly, a homogeneous equation can also be converted to non-dimensional form. For example, discharge equation is expressed as,

$$Q=(8/15)C_d \sqrt{2g}\tan\theta (H)^{5/2} \qquad \qquad ...(5.196)$$

where Q is discharge, H is height of water level and C_d is Coefficient of discharge (dimensionless constant).

Substituting dimensions in the above eqn. 5.196,

$$L^3 T^{-1}=\left[LT^{-2}\right]^{1/2}[L]^{5/2}$$
$$L^3 T^{-1} \equiv L^3 T^{-1} \qquad \qquad ...(5.197)$$

Hence it is dimensionally homogeneous.

The same equation can be written in non-dimensional form as,

$$\frac{Q}{\sqrt{g}\,H^{5/2}}=\left(\frac{8}{15}\right)\sqrt{2C_d}\,\tan\theta \qquad \qquad ...(5.198)$$

Example 5.2: A physical phenomenon is expressed as,

$$\delta=p+\left(\frac{1}{2}\right)\alpha V^2+\rho g\beta$$

where,   r = mass density,

             p = pressure,

             g = acceleration due to gravity and

             V = velocity

Find the dimensions of δ, α and β. Use (a) MLT System (b) FLT System.

Solution:

(a) MLT System (Refer Table 5.2)

$$\rho = ML^{-3} \qquad V = LT^{-1}$$
$$p = ML^{-1}T^{-2} \qquad g = LT^{-2}$$

Dimension of δ ≡ Dimension of p ≡ Dimension of αV^2 ≡ Dimension of $\rho g \beta$

$$\delta \equiv ML^{-1}T^{-2} \equiv \alpha \left(LT^{-1}\right)^2 \equiv \left(ML^{-3}\right)\left(LT^{-2}\right)\beta$$

Hence

$$\delta = ML^{-1}T^{-2}$$
$$\alpha = ML^{-3}$$
$$\beta = L$$

(b) *FLT* System (Refer Table 5.2)

$$\rho = FT^3 L^{-4} \qquad V = LT^{-1}$$
$$p = FL^{-2} \qquad g = LT^{-2}$$
$$\delta \equiv FL^{-2} \equiv \alpha \left(LT^{-1}\right)^2 \equiv \left(FT^3L^{-4}\right)\left(LT^{-2}\right)\beta$$

Hence

$$\delta = FL^{-2}$$

$$\alpha = FL^{-4}T^{-2}$$

$$\beta = L$$

5.14 DENSITY AND ENERGY FUNCTION WITH ELECTRON SPIN IN HYDROGEN LIKE ATOM

Consider a H_2 atom, which has the simplest atomic configuration with a single electron revolving round the nucleus. The total energy of such an electron is given by $-\dfrac{me^4}{8\varepsilon_0 h^2 n^2}$, where m and e are mass and charge of the electron respectively, ε_0 is the permittivity of free-space, h is the Planck's constant and n is known as principal quantum number. Thus, the total energy for such an electron solely depends on n. We know that for a given value of n, we get other three quantum numbers whose values depend on n in the following manner:

l (orbital quantum number) = 0, 1, 2, 3, ..., $(n-1)$

m1 (magnetic orbital quantum number) = -1, $(-1+1)$, $(-1+2)$, ..., -2,
$$-1, 0, 1, 2, ..., (1-1), 1$$

m_s (magnetic spin quantum number) $= -\dfrac{1}{2}, +\dfrac{1}{2}$ (for each value of m_1)

All n, l and m_1 are integers or zero (but $n \neq 0$).

A specific energy level corresponds to a specific value of n. For different integral values of n, we get different discrete energy levels. On the other hand, each discrete set of the four quantum numbers, i.e. $\{n, l, m_1, m_s\}$, designates each energy state. Let's have an example to make these concepts clear.

Consider $n = 2$ energy level.

l-value	m_1-value	m_s-value	Energy states	No. of energy states	
2	0	0	$-\dfrac{1}{2}, \dfrac{1}{2}$	$\left(2,0,0,-\dfrac{1}{2}\right)$ $\left(2,0,0,+\dfrac{1}{2}\right)$	2
2	1	-1	$-\dfrac{1}{2}, \dfrac{1}{2}$	$\left(2,1,-1,-\dfrac{1}{2}\right)$ $\left(2,1,-1,+\dfrac{1}{2}\right)$	
		0	$-\dfrac{1}{2}, \dfrac{1}{2}$	$\left(2,1,0,-\dfrac{1}{2}\right)$ $\left(2,1,0,\dfrac{1}{2}\right)$	6
		$+1$	$-\dfrac{1}{2}, \dfrac{1}{2}$	$\left(2,1,+1,-\dfrac{1}{2}\right)$ $\left(2,1,+1,\dfrac{1}{2}\right)$	

For $n = 2$, $l = 0, 1$.

For $l = 0$, $m_1 = 0$ and $m_s = -, +$

For $l = 1$, $ml = -1, 0, +1$ and $ms = -\dfrac{1}{2}, +\dfrac{1}{2}$ for each value of ml.

Thus for $n = 2$ energy level, we get eight distinct energy states, But all these eight energy states have the same value of the total energy because, the total energy, as already stated, depends only on the principal quantum number. Such type of distinct energy states having the same total energy are known as degenerate states and it is said that the energy level $n = 2$ is 8-fold degenerate. Similarly, we can show that $n = 3$ energy level is 18-fold degenerate. Note that this type of degenerate states are direct consequence of the symmetry of the respective system.

Check Your Progress

10. How will you define the non-relativistic bosons?

11. Give the experimental verification for Einstein condensation.

12. What do you mean by the liquid helium?

13. What are the basic postulates of Dirac statistics?

14. How will you define the Fermi energy for electron gas in metals?

15. Define the term Fermi temperature.

16. What is free particle?

17. Define the Heisenberg's uncertainty principle.

18. What is potential function V(x)?

19. What is reflection coefficient?

20. How will you define the Landau theory of phase transition?

21. Give short note on critical indices.

22. What is cosmology in scale transformation?

23. Define dimensional analysis.

24. Define the density and energy function with electron spin in hydrogen like atom.

5.15 ANSWERS TO 'CHECK YOUR PROGRESS'

1. The study of quantum phenomena has come to be known as quantum physics.

2. The theory of relativity extends the range of application of physical laws to the region of very high velocities and just as the universal constant of fundamental significance 'c' (speed of light in vacuum) characterizes relativity, so a universal constant of fundamental significance 'h' (Planck's constant) characterizes quantum physics.

3. Louis de-Broglie suggested that the idea of duality should be extended not only to radiation but also to all microparticles. He hypothesized that just as a quantum of radiation has a wave associated with it which governs its motion in space, so also a quantity of matter has a corresponding wave (which may be called matter wave) that governs its motion in space.

   The universe is essentially composed of only two entities namely matter Postulates and radiation. De-Broglie agreed that since one of the entities, namely radiation, has dual nature, the other entity matter must also exhibit dual character. His hypothesis is consistent with the symmetry principle of nature.

4. De-Broglie proposed to associate, with every microparticle, corpuscular characteristics namely energy E and momentum p on the one hand,

and wave characteristics namely frequency ν and wavelength λ on the other hand. According to de-Broglie, the mutual dependence between the characteristics of the two kinds was accomplished, through the Planck's constant h as

$$E = h\nu \quad \text{and} \quad p = \frac{h\nu}{c} = \frac{h}{\lambda}$$

This relation is known as de-Broglie's equation.

The wavelength λ of matter wave associated with a microparticle is called de-Broglie wavelength of the particle.

5. The wavelength of electrons impinging the crystal are given by $\lambda = \frac{h}{p}$, according to de-Broglie's equation.

6. The basic postulates of BE statistics are:
   (i) The associated particles are identical and indistinguishable.
   (ii) Each energy state can contain any number of particles.
   (iii) Total energy and total number of particles of the entire system are constant.
   (iv) The particles have zero or integral spin, i.e., 0, 1, 5, 50, etc., where is the unit of spin.
   (v) The wave function of the system is symmetric under the positional exchange of any two particles.

   Examples: Photon, phonon, all mesons (p, k, h), etc., these are known as Bosons.

7. The total number of possible ways (microstates) in BE statistics are respectively given by

$$W_{BE} = \frac{(N_i + g_i - 1)}{N_i!(g_i - 1)!}$$

Here, $N = 2$, $Ni = 2$, $gi = 3$. Thus,

$$W_{BE} = \frac{(2 + 3 - 1)!}{2!(3 - 1)!} = \frac{4!}{2!2!} = 6$$

For BE statistics, the particles are indistinguishable and any state can contain any number of particles. Let particles be a, a.

Cell 1	Cell 2	Cell 3
aa	0	0
0	aa	0
0	0	aa
a	a	0
a	0	a
0	a	a

8. In quantum mechanics we have already established Planck's law of black-body radiation which exactly accounts for the observed energy density in case of a black-body radiation. Here shall re-derive the same Planck's law by using Bose-Einstein statistics.

9. A blackbody is defined as it absorbs all the thermal radiations, whatever may be the wavelength incident on it. It neither reflects nor transmits any radiation, and appears black.

10. Our bosons are non-relativistic particles with spin *s*, whose one-particle energies $\epsilon(k)$

$$\epsilon(\mathbf{k}) = \epsilon(k) = \frac{\hbar^2 k^2}{2m}, \qquad \epsilon_0 = \epsilon(0) = 0$$

include only the kinetic energy term.

11. The Bose-Einstein condensation was predicted by Satyendra Bose and Albert Einstein in 1924-1925. It took almost 70 years to have an experimental corroboration of this phenomenon with the ultracold gas systems. Previous experiments had been done with 4He as well as with hydrogen.

12. Liquid helium is a physical state of helium at very low temperatures at standard atmospheric pressures. Liquid helium may show superfluidity.

13. The some basis postulates of FD statistics are:
    (i) Particles are identical and indistinguishable.
    (ii) Total energy and total number of particles of the entire system is constant.

14. Metals are characterised by the presence of a good number of free electrons. These electrons move about at random within the metals. While moving the free electrons collide among themselves also encounter with the fixed ion cores. Such behaviour of the free electrons is similar to that of molecules of a gas.

15. Fermi temperature (TF) is the temperature equivalent of Fermi energy (FF) and it is defined as $T_F =$, where k (= 1.38 × 10^{-23} Joule/Kelvin) is the Boltzmann constant.

16. The study of quantum phenomena has come to be known as quantum physics.

17. The theory of relativity extends the range of application of physical laws to the region of very high velocities and just as the universal constant of fundamental significance 'c' (speed of light in vacuum) characterizes relativity, so a universal constant of fundamental significance 'h' (Planck's constant) characterizes quantum physics.

18. Louis de-Broglie suggested that the idea of duality should be extended not only to radiation but also to all microparticles. He hypothesized that just as a quantum of radiation has a wave associated with it which governs its motion in space, so also a quantity of matter has a corresponding wave (which may be called matter wave) that governs its motion in space.

The universe is essentially composed of only two entities namely matter and radiation. De-Broglie agreed that since one of the entities, namely radiation, has dual nature, the other entity matter must also exhibit dual character. His hypothesis is consistent with the symmetry principle of nature.

19. De-Broglie proposed to associate, with every microparticle, corpuscular characteristics namely energy E and momentum p on the one hand, and wave characteristics namely frequency ν and wavelength λ on the other hand. According to de-Broglie, the mutual dependence between the characteristics of the two kinds was accomplished, through the Planck's constant h as

$$E = h\nu \quad \text{and} \quad p = \frac{h\nu}{c} = \frac{h}{\lambda}$$

This relation is known as de-Broglie's equation.

The wavelength λ of matter wave associated with a microparticle is called de-Broglie wavelength of the particle.

20. Landau theory in physics is a theory that Lev Landau introduced in an attempt to formulate a general theory of continuous (i.e., second-order) phase transitions. It can also be adapted to systems under externally-applied fields, and used as a quantitative model for discontinuous (i.e., first-order) transitions.

21. Critical exponents describe the behavior of physical quantities near continuous phase transitions. It is believed, though not proven, that they are universal, i.e., they do not depend on the details of the physical system, but only on some of its general features. For instance, for ferromagnetic systems, the critical exponents depend only on:

- The dimension of the system
- The range of the interaction
- The spin dimension

22. In physical cosmology, the power spectrum of the spatial distribution of the cosmic microwave background is near to being a scale-transformation function. Although in mathematics this means that the spectrum is a power-law, in cosmology the term 'scale-transformation' indicates that the amplitude, P(k), of primordial fluctuations as a function of wave number, k, is approximately constant, i.e. a flat spectrum. This pattern is consistent with the proposal of cosmic inflation.

23. Dimensional analysis is a branch of mathematics which deals with dimensions of quantities. Each physical phenomenon can be expressed by an equation that represents the relationship between the variables governing the phenomenon.

24. Consider a H_2 atom, which has the simplest atomic configuration with a single electron revolving round the nucleus. The total energy of such an electron is given by $-\dfrac{me^4}{8\varepsilon_0 h^2 n^2}$, where m and e are mass and charge of the electron respectively, ε_0 is the permittivity of free-space, h is the Planck's constant and n is known as principal quantum number. Thus, the total energy for such an electron solely depends on n. We know that for a given value of n,

5.16 SUMMARY

- The study of quantum phenomena has come to be known as quantum physics.

- Like classical physics, quantum physics also has been provided with a mathematical apparatus. The currently accepted structure developed by Schrödinger, Heisenberg, Max Born, Jordan, Dirac and many others to deal with problems in the microdomain, i.e., at atomic and subatomic levels is termed as quantum mechanics.

- The theory of relativity extends the range of application of physical laws to the region of very high velocities and just as the universal constant of fundamental significance 'c' (speed of light in vacuum) characterizes relativity, so a universal constant of fundamental significance 'h' (Planck's constant) characterizes quantum physics which includes classical physics as a special case.

- Around 1923, Louis de-Broglie suggested that the idea of duality should be extended not only to radiation but also to all microparticles. He hypothesized that just as a quantum of radiation has a wave associated with it which governs its motion in space, so also a quantity of matter has a corresponding wave (which may be called matter wave) that governs its motion in space.

- The universe is essentially composed of only two entities namely matter and radiation. De-Broglie agreed that since one of the entities, namely radiation, has dual nature, the other entity matter must also exhibit dual character. His hypothesis is consistent with the symmetry principle of nature.

- De-Broglie proposed to associate, with every microparticle, corpuscular characteristics namely energy E and momentum p on the one hand, and wave characteristics namely frequency ν and wavelength λ on the other hand. According to de-Broglie, the mutual dependence between the characteristics of the two kinds was accomplished, through the Planck's constant h as

$$E = h\nu \quad \text{and} \quad p = \frac{h\nu}{c} = \frac{h}{\lambda}$$

  This relation is known as de-Broglie's equation.

- The wavelength λ of matter wave associated with a microparticle is called de-Broglie wavelength of the particle.

- The wavelength of electrons impinging the crystal are given by $\lambda = \frac{h}{p}$, according to de-Broglie's equation.

- In quantum mechanics we have already established Planck's law of black-body radiation which exactly accounts for the observed energy density in case of a black-body radiation. Here shall re-derive the same Planck's law by using Bose-Einstein statistics.

- The Bose-Einstein condensation was predicted by Satyendra Bose and Albert Einstein in 1924-1925. It took almost 70 years to have an

experimental corroboration of this phenomenon with the ultracold gas systems. Previous experiments had been done with 4He as well as with hydrogen.

- Liquid helium is a physical state of helium at very low temperatures at standard atmospheric pressures. Liquid helium may show superfluidity.

- Metals are characterised by the presence of a good number of free electrons. These electrons move about at random within the metals. While moving the free electrons collide among themselves also encounter with the fixed ion cores. Such behaviour of the free electrons is similar to that of molecules of a gas.

- Fermi temperature (TF) is the temperature equivalent of Fermi energy (FF) and it is defined as TF = , where k (= $1.38 \times 10{-}23$ Joule/Kelvin) is the Boltzmann constant.

- Critical exponents describe the behavior of physical quantities near continuous phase transitions. It is believed, though not proven, that they are universal, i.e., they do not depend on the details of the physical system, but only on some of its general features.

- In physical cosmology, the power spectrum of the spatial distribution of the cosmic microwave background is near to being a scale-transformation function. Although in mathematics this means that the spectrum is a power-law, in cosmology the term 'scale-transformation' indicates that the amplitude, $P(k)$, of primordial fluctuations as a function of wave number, k, is approximately constant, i.e. a flat spectrum. This pattern is consistent with the proposal of cosmic inflation.

- Dimensional analysis is a branch of mathematics which deals with dimensions of quantities. Each physical phenomenon can be expressed by an equation that represents the relationship between the variables governing the phenomenon.

- In general, such quantities may be dimensional or non-dimensional (absolute numerals). Any physical variable can be described using qualitative and quantitative approaches. The qualitative description is known as dimension and quantitative description is known as unit.

- The dimensions of basic quantities are selected as basic dimensions and dimensions of other variables are derived. Dimensions are classified as Absolute system (MLT system) and Gravitational System (FLT System).

5.17 KEY TERMS

- **Quantum physics:** The study of quantum phenomena has come to be known as quantum physics.

- **Blackbody:** A blackbody is defined as it absorbs all the thermal radiations, whatever may be the wavelength incident on it. It neither reflects nor transmits any radiation, and appears black.

- **Liquid helium:** Liquid helium is a physical state of helium at very low temperatures at standard atmospheric pressures. Liquid helium may show superfluidity.

- **Fermi temperature:** Fermi temperature (TF) is the temperature equivalent of Fermi energy (FF).

- **Free particle:** Free particle means a particle which moves freely in space without the influence of any force. Hence, for a free particle the potential energy is zero.

- **Critical indices:** Critical exponents describe the behavior of physical quantities near continuous phase transitions. It is believed, though not proven, that they are universal, i.e., they do not depend on the details of the physical system, but only on some of its general features.

- **Dimensional analysis:** Dimensional analysis is a branch of mathematics which deals with dimensions of quantities. Each physical phenomenon can be expressed by an equation that represents the relationship between the variables governing the phenomenon.

5.18 SELF-ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. When was quantum mechanics developed?
2. Name the areas where the quantum mechanics is used.
3. Define Bose-Einstein distribution law.
4. Give the application of Bose-Einstein statistics.
5. What is blackbody radiation?
6. How will you define the pressure of an ideal Bose gas?
7. Define the term Einstein condensation.
8. What is the theory of liquid helium?
9. State the Fermi Dirac statistics.
10. What do you mean by free electron theory of solids?
11. Define Landau theory of phase transition.
12. How will you define the critical indices?
13. Give a short note on scale transformation.
14. Define the term dimensional homogeneity.
15. What do you mean by the density an energy function with electron spin in hydrogen like atom?

Long-Answer Questions

1. Discuss the origin, history and significance of quantum mechanics.
2. Describe the wave nature of microparticles as stated by de-Broglie's hypothesis.

3. Explain Davisson and Germer's experiment with the help of a diagrams.

4. Prove that the Thomson's experiment is analogous to Debye-Scherrer X-ray diffraction method.

5. What do you understand by the Bose-Einstein distribution law? Discuss the application of Bose-Einstein statistics with the help of giving examples.

6. Discuss the graphical representation of Wein, Rayleigh-Jeans and Planck's distribution law. Give appropriate examples.

7. Explain briefly about the pressure of an ideal Bose gas with the help of relevant examples.

8. Analysis the theory of liquid helium with the help of giving examples.

9. What are the basic postulates of Fermi Dirac statistics? Explain with appropriate examples.

10. Describe the free electron theory of solids with the help of relevant examples.

11. What do you understand by the Landau theory of phase transition? Discuss the first and second-order transitions with the help of giving examples.

12. Discuss the definition and types of critical indices. Give appropriate examples.

13. Explain the dimensional analysis with the help of relevant examples.

5.19 FURTHER READING

Rao, K. Sankara. 2009. *Classical Mechanics*. New Delhi: PHI Learning Private Limited.

Upadhyaya, J.C. 2010. *Classical Mechanics*, 2nd Edition. New Delhi: Himalaya Publishing House.

Goldstein, Herbert. 2011. *Classical Mechanics*, 3rd Edition. New Delhi: Pearson Education India.

Gupta, S.L. 1970. *Classical Mechanics*. New Delhi: Meenakshi Prakashan.

Takwala, R.G. and P.S. Puranik. 1980. New Delhi: Tata McGraw Hill Publishing.