CLASSICAL MECHANICS AND STATISTICAL MECHANICS
(DPHY02)
(MSC PHYSICS)

ACHARYA NAGARJUNA UNIVERSITY

CENTRE FOR DISTANCE EDUCATION

NAGARJUNA NAGAR,

GUNTUR

ANDHRA PRADESH

PAPER - II : CLASSICAL MECHANICS AND STATISTICAL MECHANICS

SYLLABUS

CLASSICAL MECHANICS

UNIT - I

Mechanics of a particle : Mechanics of a system of particles constraints. Principle of virtual work; D' Alembert's principle Lagrange's equations. Velocity dependent, potential Dissipation function. Hamilto's Principle, Derivation of Lagrange's equations from Hamilton's principle. Extension to non conservative and non - holonomic systems. Hamilton's equations of motion and their derivation from Lagrange's equation. Cyclic coordinates. Conservation theorems. Principal of least action.

Co-ordinates of a rigid body. Eulerian angles. Transformation matrix, Infinitesimal rotations. Representation as vectors Rate of change of a vector in a moving frame of reference. Centripetal acceleration, Coriolis force.

Angular momentum and kinetic energy of a rotating rigid body. The inertia tensor and moment of inertia. Euler equation of motion. Torque free motion of rigid body. Heavy symmetrical top with one point fixed.

UNIT II

Special theory of relativity. Lorentz transformation - Covariant four dimensional formulations : Force and energy equations in relativistic mechanics. Lagrangian formulation of relativistic mechanics. Lagrangian formulations. Canonical transformations. Generating functions. Simple examples of Lagrange and Poisson brackets. Their canonical invariance.

Hamilton - Jocobi equations. Hamilton's principle and characteristic functions. Separation of variables, Action - Angle variables, Kepler problem.

Formulation of the small oscillation problem. Eigen value equation and principle axis transformation. Frequencies of free vibration and normal co-ordinates. Examples of a linear triatomic molecule. Forced vibrations Effect of dessipative force.

TEXT BOOK :

1. Classical Mechanics by Herbert Goldstein (Narosa)
2. Classical Mechanics by Gupta, Kumar & Sharma (Pragathi Prakashan)

STATISTICAL MECHANICS

UNIT - III

Classical Statistical Mechanics : The Postulate of classical mechanics, Micro canonical ensemble, Derivation of thermodynamics, Equi-partition theorem.

Classical ideal gas (on the basis of micro canonical ensemble) Derivation of entropy for classical ideal gas and Gibb's paradox.

Canonical ensemble and grand canonical ensemble : Canonical ensemble, energy fluctuations in the canonical ensemble, Grand canonical ensemble, Density. fluctuations in the grand canonical ensemble, Equivalence between the canonical ensemble and grand canonical ensemble.

UNIT - IV

Quantum Statistical Mechanics : The Postulates of Quantum Statistical Mechanics, third law of thermodynamics. The ideal gases : Determination of thermodynamic parameters (micro canonical ensemble and grand canonical ensemble), foundations of statistical mechanics.

The partition function : Darwin - Fowler Method, classical limit of the partition function and the variational principle.

Ideal Fermi Gas : Equation of state of an ideal Fermi gas. Theory of white dwarf stars.

Ideal Bose gas : Photons, Phonons, Bose - Einstein condensation.

TEXT BOOK :

1. Statistical Mechanics by Kerson Huang, Wiley Eastern (pvt) ltd, New Delhi.

CONTENTS

CLASSICAL MECHANICS

STATISTICAL MECHANICS

CLASSICAL MECHANICS

Classical Mechanics Part
Unit -I
Lesson 1

MECHANCS OF SYSTEM OF PARTICLES

Objective: In this lesson we learn about
  1) Principles of mechanics of particles
  2) Conservation laws
  3) Constraints

Structure:
1.1 Mechanics of a particle
1.2 Conservation of energy
1.3 Mechanics of system of particles
1.4 Conservation of linear momentum for system of particles
1.5 To prove conservation of total Energy
1.6 Constraints
1.7 Types of constraints.

Introduction: Mechanics is the study of motion of physical bodies. Classical mechanics deals with the situations involving all bodies we come across in our daily life. However when the particles are too small or when they are travelling with velocities comparable to that of light, it was found that normal classical laws do not hold good and we have to apply other types of mechanics like quantum mechanics and relativistic mechanics. Even systems for which quantum mechanics holds, they are modeled according to classical mechanics and transformed to quantum mechanical form for further analysis. The progress of various sciences like astronautics, stellar dynamics, robotics, aerodynamics, and various branches of engineering still depend heavily on the foundation of classical mechanics. Hence it is very important for every student of physics to have considerable knowledge in classical mechanics and how the classical ideas were extended into the realm of quantum mechanics.

1.1 : Mechanics of a particle:

The position of any particle can be represented with three coordinates like x,y,z called Cartesian coordinates When the particle is in motion its position changes with time. The general displacement can be resolved into components along x,y and z axes. As we already know, the rate of change of displacement with time is called velocity. If the particle possesses a mass m it is said to possess momentum. Newton formulated the laws o motion for a particle in motion. These laws are known as Newton's laws of motion are stated as

First law: A body continues in its state of rest or uniform motion, unless not disturbed by
  some external influence. It is also called Law of Inertia.
Second law: The time-rate of change of momentum is proportional to the impressed force. It
  is also called Law of Force.

Third law: To every action there is always equal and opposite reaction. This is known as
  Law of action and reaction.

: From Newton's second law of motion·

$$F = \frac{d}{dt}(m\mathbf{v}) = \frac{d\mathbf{p}}{dt} \qquad ...(1)$$

When the total force F is zero, then $\frac{d\mathbf{p}}{dt} = 0$, so the linear momentum is conserved ie. When a particle is moving in zero force field its linear momentum remains as constant. This is called law of <u>Conservation of linear momentum</u>

A particle motion need not be confined to a direction. The particles may be revolving around a centre of attraction (Central force field). In this case also the particle possess momentum, but, as the direction may change continuously, it is called angular momentum.

Consider a particle of mass 'm' and linear momentum $\mathbf{p}$ at position vector $\mathbf{r}$ relative to the origin 'O', of an inertial frame of reference. The moment of momentum is defined as the angular momentum. $\therefore$ $\mathbf{L} = \mathbf{r} \times \mathbf{p}$

$\therefore$ Torque about the origin 'O' is $\mathbf{N} = \mathbf{r} \times \mathbf{F}$

$$\therefore \quad \mathbf{N} = \mathbf{r} \times \frac{d\mathbf{p}}{dt} = \frac{d}{dt}(\mathbf{r} \times \mathbf{p}) - \frac{d\mathbf{r}}{dt} \times \mathbf{p}$$

$$= \frac{d}{dt}(\mathbf{r} \times \mathbf{p}) - (\mathbf{v} \times m\mathbf{v})$$

$$\therefore \quad \mathbf{N} = \frac{d}{dt}(\mathbf{r} \times \mathbf{p}) = \frac{d\mathbf{L}}{dt}$$

So the rate of change of angular momentum is equal to the Torque acting on the particle.

If the total torque $\mathbf{N} = 0$, then $\frac{d\mathbf{L}}{dt} = 0 \Rightarrow \mathbf{L} =$ constant. Hence the angular momentum is conserved in the absence of an external torque. This is called the law of <u>Conservation of angular momentum</u>

1.2 Conservation of energy: A force, which is derivable from scalar potential energy function in the manner $F = -\Delta V$ then it is called conservative. If the forces, which are conservative act upon the particle, then the total energy of the particle $(KE + PE)$ is conserved.

Under the action of such force let the particle moves from position 1 to position 2. Then workdone by the particle be,

$$W_{12} = \int_1^2 \mathbf{F}.d\mathbf{r} = \int_1^2 \frac{d\mathbf{p}}{dt}.d\mathbf{r}$$

But $\mathbf{p} = m\dot{\mathbf{r}}$ and $d\mathbf{r} = \dot{\mathbf{r}}\,dt$

$$\therefore \quad W_{12} = \int_1^2 \frac{d}{dt}(m\dot{\mathbf{r}}).\dot{\mathbf{r}}\,dt$$

$$= \int_1^2 \frac{d}{dt}\left(\frac{1}{2}m\dot{\mathbf{r}}^2\right)dt = T_2 - T_1 \qquad ...(3)$$

Where $T_1 = KE$ of the particle at the position 1. $T_2 = KE$ of the particle at the position 2.

We have $F = -\nabla v$

$$\therefore \quad W_{12} = \int_1^2 -\nabla V.d\mathbf{r}$$

$$= \int_1^2 -\frac{dV}{dr}dr = -\int_1^2 dV = V_1 - V_2 \quad ...(4)$$

from (3) & (4) : $T_2 - T_1 = V_1 - V_2 \Rightarrow T_1 + V_1 = T_2 + V_2$

$\Rightarrow T + V$ = constant. So the total energy of the particle is conserved even when it is moving in a conservative force field.

Till now we dealt with the motion of a single particle. But, in real situations we have to deal with not a single particle but a system of particles either independent or interacting with one another and moving under some system constraints. We now try to understand the relevant principles.

1.3. Mechanics of system of particles:

Consider a system of 'n' particles. The equation of motion interms of Newton's second law can be easily written as,

$$m_i a_i = \dot{p}_i = F_i^j + \sum_{j \neq i} F_i^j \qquad (\because i = 1,2,3....N) \qquad ...(1)$$

$F_i^{(e)}$ = external force acting on i^{th} particle.

F_i^j = internal force on the ith particle due to j^{th} particle.

All the particles of the system exert forces on one another. So the internal force on i^{th} particle must be the sum of forces due to all other particles is $\sum_{j=1}^{n} F_i^j$ excluding the term $j = i$, (Because from the definition

$F_i^i = 0$)

Form Newton's third law the force F_i^j must be equal and opposite in direction to the force F_j^i.

$$\therefore \qquad F_i^j = F_j^i \qquad ...(2)$$

Hence the internal forces occur in pairs and act along the line joining the two particles.
By considering all the particles of the system, the equation of motion of the total system is,

$$\sum_i \dot{p}_i = \frac{d^2}{dt^2} \sum_i m_i \bar{r}_i$$

$$= \sum_i F_i^{(e)} + \sum_{i,j}^1 F_i^j = \sum_1 F_i^{(e)} \qquad ...(3)$$

we have $\sum_{i,j}^1 F_i^j = \sum_{i,j}^1 F_j^i = \frac{1}{2} \sum_{i,j}^1 [F_i^j + F_j^i] = 0$

$$\left(\sum_{i,j}^1 \text{ means } i = j \text{ term is excluded} \right)$$

Let **R** be the position vector of center of mass.

$$\therefore \qquad R = \frac{\sum_i m_i \bar{r}_i}{\sum_i m_i} = \frac{\sum_i m_i \bar{r}_i}{M}$$

Where $M = \sum_i m_i$ is the total mass of the system.

$$\therefore \qquad (3) \Rightarrow M \frac{d^2 R}{dt^2} = \sum_i F_i^{(e)} = F^{(e)} \qquad ...(4)$$

Total linear momentum of the system

$$P = \sum_i m_i \bar{r}_i$$

$$= \frac{d}{dt} \sum_i m_i \bar{r}_i = M \dot{R} \qquad \ldots(5)$$

1.4: Conservation of angular momentum for system of particles:

The total angular momentum of the system of particles $L = \sum_i (r_i \times p_i)$

$$\therefore \quad \frac{dL}{dt} = \frac{d}{dt} \sum_i (r_i \times p_i) = \sum_i [(r_i \times \dot{p}_i) + (\dot{r}_i \times p_i)]$$

$$= \sum_i (r_i \times \dot{p}_i) \qquad (\because \dot{r}_i \times p_i = 0)$$

$$\Rightarrow \quad \dot{L} = \sum_i r_i \times F_i^{(e)} + \sum_{i,j} r_i \times F_i^j \qquad \ldots(7)$$

The second term is due to the sum of internal torques. If the interacting forces are Newtonian in nature, the second term vanishes.

$$\therefore \quad \frac{dL}{dt} = N^{(e)} = \sum_i r_i \times F_i^{(e)} \qquad \ldots(8)$$

When there is no external torque $N^{(e)} = 0$.

$$\therefore \quad \frac{dL}{dt} = 0 \Rightarrow L = \text{Constant.} \qquad \ldots(9)$$

This is the <u>law of conservation of angular momentum for a system of particles</u>.

1.5. The law of conservation of total energy: For a system of particles $V = V^{(e)} + V^{int}$ ...(10)

and also, $F_i^{(e)} = -\nabla_i V^{(e)}$ and $F_i^{int} = -\nabla_i V^{int}$

$$\ldots(11)$$

Here V^{int} is the internal PE function.

$$V^{int} = \sum_j V_i^j = \sum_j V_j^i (|r_i - r_1|) = \sum_j V_i^j (|r_j - r_i|)$$

$$= \sum_i V_j^i (|r_i - r_j|) = V_j^{int}$$

Total PE $V^{int} = \sum_i V_i^{int} = \frac{1}{2} \sum_{i,j} V_j^i (|r_i - r_i|)$

The factor $\frac{1}{2}$ is taken because, while summing the mutual potential energies, a pair of particles i, j appear twice.

Proof:

Consider a system of particles in which i^{th} particle is displaced through $d\overline{\mathbf{r}}_i$ by the force $\mathbf{F}_i$. Then the amount of work done on the particle is,

$$m_i\ddot{\mathbf{r}}_i.d\mathbf{r}_i = \mathbf{F}_i.d\mathbf{r}_i = \left(\mathbf{F}_i^{(e)} + \sum_j \mathbf{F}_i^j\right).d\mathbf{r}_i \qquad ...(12)$$

But $d\mathbf{r}_i = \dot{\mathbf{r}}_i.dt$.

$$\therefore \quad (12) \Rightarrow m_i\ddot{\mathbf{r}}_i\dot{\mathbf{r}}_idt = \mathbf{F}_i^e.d\mathbf{r}_i + \sum_j \mathbf{F}_i^j.d\mathbf{r}_i$$

Hence for a system of particles,

$$\frac{d}{dt}\sum_i\left(\frac{1}{2}m_i\dot{\mathbf{r}}_i^{\,2}\right)dt = \sum_i\mathbf{F}_i^{(e)}.d\mathbf{r}_i + \sum_{i,j}\mathbf{F}_i^j.d\mathbf{r}_i \quad ...(13)$$

Consider $\sum_{i,j}^{1}\mathbf{F}_i^j.d\mathbf{r}_i = \frac{1}{2}\sum_{i,j}^{1}\left(\mathbf{F}_i^j.d\mathbf{r}_i + \mathbf{F}_j^i.d\mathbf{r}_j\right)...(14)$

For the forces obeying Newton's third law, $\mathbf{F}_i^j = -\mathbf{F}_j^i$

And also, $V_i^j = V_i^j\left(|\mathbf{r}_i - \mathbf{r}_j|\right)$ and $V_i^j = V_j^i$

$\therefore \mathbf{F}_i^j = -\nabla_i V_i^j, \mathbf{F}_j^i = -\nabla_j V_i^j$ and $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$

$$\therefore F_i^j = -\nabla_i V_i^j = -\frac{\partial}{\partial\mathbf{r}_i}\left(V_i^j\right) = -\frac{\partial}{\partial\mathbf{r}_{i,j}}V_i^j\left(\frac{\partial\mathbf{r}_{ij}}{\partial\mathbf{r}_i}\right)$$

$$= -\nabla_{ij}V_i^j.\left(\frac{\partial}{\partial\mathbf{r}_i}\left(\mathbf{r}_i - \mathbf{r}_j\right)\right) = -\nabla_{ij}V_i^j \quad ...(15)$$

we have $\mathbf{F}_j^i = -\nabla_j V_i^j = -\frac{\partial}{\partial\mathbf{r}_j}\left(V_i^j\right)_i = -\frac{\partial V_i^j}{\partial\mathbf{r}_{ij}}\left(\frac{\partial\mathbf{r}_{ij}}{\partial\mathbf{r}_j}\right)$

$$= -\nabla_{ij}V_i^j\left[\frac{\partial}{\partial\mathbf{r}_j}\left(\mathbf{r}_i - \mathbf{r}_j\right)\right] = \nabla_{ij}V_i^j$$

$$...(16)$$

Substituting (15), (16) in (14)

$$\sum_{i,j}^{1}\mathbf{F}_i^j.d\mathbf{r}_i = \frac{1}{2}\sum_{i,j}^{1}\left(-\nabla_{ij}V_i^j.d\mathbf{r}_i + \nabla_{ij}V_i^j.d\mathbf{r}_j\right)$$

$$= -\frac{1}{2}\sum_{i,j}^{1}\nabla_{ij}V_i^j.d\left(\mathbf{r}_i - \mathbf{r}_j\right)$$

$$= -\frac{1}{2}\sum_{i,j}^{1}\nabla_{ij}V_i^j.d\mathbf{r}_{i,j} \qquad ...(17)$$

If the external force is also conserved,
$$F_i^{(e)} = -\nabla_i V_i^{(e)} \qquad ...(18)$$
from the equation (17) & (18) we can write (13) as,

$$\frac{d}{dt}\sum_i\left(\frac{1}{2}m_i\,\dot{r}_i^2\right)dt=-\sum_i\nabla_iV_i^e\,.d\mathbf{r}_i-\frac{1}{2}\sum_{i,j}^{!}\nabla_{i,j}V_i^j\,.d\mathbf{r}_{ij}$$

On integration $\displaystyle\sum_i\frac{1}{2}m_i\,\dot{r}_i^2=-\sum_iV_i^e-\frac{1}{2}\sum^{!}V_i^j+$ constant

$\Rightarrow\quad\displaystyle\sum_i\frac{1}{2}m_i\,\dot{r}_i^2+\sum_iV_i^e+\frac{1}{2}\sum_{i,j}V_i^j=$ constant

$\Rightarrow\quad T+V=$ constant $\qquad\qquad$ ...(19)

where $\displaystyle V=\sum_iV_i^{(e)}+\frac{1}{2}\sum_{i,j}^{!}V_i^j$

V is the sum of external and internal potential energies of the system. Eq. (19) is the conservation of energy.

1.6 Constraints :

A motion, which cannot proceed arbitrarily in any manner, is called constrained motion. Motion along a specified path is an example of constrained motion. Depending on the nature of the motion of particle, the coordinates are considered to describe the motion. If the particle is free to move in space, then three coordinates are needed to describe its motion. Thus imposing constraints on a mechanical system is to simplify the configuration of the system.

Example 1: Consider a particle moving in space. It requires three coordinates to determine its position at any instant. If its movement is restricted on the surface of a sphere, there exists a relation between its coordinates as,

$$x^2+y^2+z^2=a^2\;\Rightarrow r=a\,.\qquad\qquad...(1)$$

Eqn. (1) is the eqn. of a sphere with its center at origin. By using the equations of constraints a coordinate can be eliminated from the set of three coordinates. Instead of Cartesian coordinates if we express the problem interms of spherical polar coordinates (r, θ and ϕ), then the two coordinates θ and ϕ will be sufficient to describe the position of the particle completely.

Example 2: Consider the simple pendulum whose motion is confined in the vertical plane. To locate the position of the bob in motion only two coordinates are required. Let the motion of the bob taken placed under a constraint that the distance 'l' of the bob is to remain same at all time. This condition imposed by the constraint can be expressed in the form of an equation either between x and y or r and θ.

$\therefore$ Constraints $x^2+y^2=l^2$, or $r=l$ $\qquad$ ... (2)

The eqn (2) is used to reduce the number of coordinates, which other wise would have been two.

Example 3. Consider a rigid body. It is defined as a system of particles in which the relative distance of the constituent particles are fixed and cannot vary with time. In this case constraints are expressed by equations of the form $r_{ij}=C_{ij}$. In this r_{ij} denote the distance between i^{th} and j^{th} particles and C_{ij} are constants. If $\mathbf{r}_i(x_i,y_i,z_i)$ and $\mathbf{r}_j(x_j,y_j,z_j)$ are the coordinates w.r.t. the origin, then the conditions can be expressed as,

$$(x_i-x_j)^2+(y_i-y_j)^2+(z_i+z_j)^2=(C_{ij})^2\quad...(3)$$

In general for a system of N particles,

$$f(x_1y_1z_1:x_2y_2z_2:x_ny_nz_n:t)=0\qquad...(4)$$

1.7: Types of constraints:

1) Holonomic constraints 2) Non-Holonomic constraints

1. Holonomic Constraints: Constraints that can be expressed in the form of
$$f(X_1, Y_1, Z_1, X_2, Y_2, Z_2 \ldots\ldots X_n Y_n Z_n \cdot t) = 0$$
where time 't' may occur in case of constraints which may vary with time.

Ex : Let us consider the motion of a simple pendulum confined to move in the vertical plane. We would need only two co-ordinates (contain (X, Y, Z) or polar co-ordinates (r, θ) with respect to the point of suspension with 'O' as)

To locate the position of the bob in motion, however motion of the bob is not free but takes place under a constraint that the distance 'l' or the bob B to remain the same from 'O' at all times. This condition imposed by the constraint can be expressed in the form of an equation either by X and Y (or) r and θ

$$x^2 + y^2 = l^2$$
$$r = l \quad \ldots(1)$$

In plane polar co-ordinates the equation looks simpler again one co-ordinate either X or Y in Cartesian co-ordinates, r (or) θ in polar co-ordinates would be sufficient to describe the motion. Now that we have utilized eq. (1) to reduce the number of co-ordinates which otherwise would have been two.

2. Non-Holonomic Constraints :

Constraints, which are not expressible in the form, $f(X_1, Y_1, Z_1, X_2, Y_2, Z_2 \ldots\ldots X_n Y_n Z_n \cdot t) = 0$ are termed as non-holonomic constraints.

The motion of the particle placed on the surface of sphere under the action of gravitational force is bound by non-holonomic constraints, for it can be expressed as an inequality.

$$r^2 - a^2 \geq 0$$

Equality sign holds until the particle rolls on the sphere and when it leaves the sphere we must have
$$r^2 - a^2 > 0$$

Ex : The constraints involved in the motion of the molecules in a gas container are non-holonomic.

Scleronomic & Rheonomic Constraints:

If the constraints are independent of time they are termed as scleronomic but if they contain time explicitly they are called Rheonomic.
A bead sliding on a moving wire is an example of Rheonomic constraint.

Force of constraints:

Constraints not only interfere with the solution or the problems in that the co-ordinates are no longer independent but they are always associated with the forces by virtue of which they restrict the motion of the system. Such forces are termed as forces of constraints. We generally formulate the laws of mechanics in a way that the work done by the forces of constraints is zero, when the system is in motion.
Forces of constrains in the case of a bead sliding on a wire is the reaction by the wire exerted on the bead at each point. The surface of sphere exerts a reaction force on the particle normally at each point.

Summary: For a particle moving in a force free space the linear momentum and energy are conserved. When a particle is moving in conservative force field linear momentum, angular

momentum, and total energy of the system are conserved. These conservation laws hold good even for a system of particles moving in a conservative force field. In general in a multi-particle system a particle may have specific relation ship with its neighbors and so cannot move arbitrarily. We say that its motion is constrained. The relations, which restrict the motion of particles, are called constraints. The constraints are divided in to holonomous and non-holonomous types. Non-holonomous constraints are further divided into scleronomous and rheonomous constraints. Constraints are associated with forces called constraint forces. However the laws of mechanics are so formulated so that the work done by the forces of constraints are zero.

Key Words:

Angular momentum, linear momentum, total energy (T+V), and constraints, Holonomic, Non-holonomic constraints, forces of constraints.

Self Assessment Questions:
1. What are constraints? Give specific examples to explain the forces of constraints.
2. Prove the laws of conservation of linear momentum, angular momentum and energy for a system of particles.
3. State and prove work-energy theorem.

Reference books:
1. Classical Mechanics: H.Goldstein
2. Mechanics: Simon
3. Mechanics: Gupta, Kumar and Sharma.

Classical Mechanics Part
Unit I
Lesson 2

Lagrangian Formulation

Objective: To learn about
1) The D'Alembert's Principle and derivation of Lagrangian equation of motion from it
2) The concept of virtual displacement and virtual work.
3) Hamilton's variational principle and deriving Hamilton's canonical equations of motion.
4) Conservation theorems.
5) Principle of Least Action.

Structure:
2.1 D'Alembert's Principle
2.2 Lagrange's equations
2.3 Virtual displacement and work
2.4 Principle of virtual work
2.5 Hamilton's approach
2.6 Hamiltons variational principle
2.7 Canonical equations of motion
2.8 Conservation theorems
2.9 Principle of least action.

2.1: D'Alembert's Principle

This method is based on the principle of virtual work. The system is subjected to an infinitesimal displacement consistent with the forces and constraints imposed on the system at the given instant 't'. This change in the configuration of the system is not associated with a change in time (i.e.) there is no actual displacement during which forces and constraints may change and hence the displacement is termed as virtual displacement.

Now suppose the system is in equilibrium (i.e.) the total force F_i on every particle is zero. Then work done by this force in a small virtual displacement δr_i will also vanish i.e.,

$$\sum_i F_i \cdot \delta r_i = 0$$

Let this total force be expressed as sum of applied force F_i^a and forces of constraints f_i then above equation takes the form

$$\sum_i F_i^a . \delta r_i + \sum_i F_i . \delta r_i = 0$$

We now consider the systems for which the virtual work of the forces of constraints is zero. An example of such a system can be had in mind is that a particle i.e. constrained to move on a smooth surface. So that the forces of constraints being $\perp$ to the surface while virtual displacement is tangential to it. Then virtual work done by forces of constraints is zero. Thus

$$\sum_i F_i^a \, \delta r_i = 0 \qquad \qquad ...(1)$$

This equation. is termed as principle of virtual work. To interpret the equilibrium of the system D'Alembert adopted an idea of a reversed force. D'Alembert conceived that a system would remain in equilibrium under the action of a force equal to the actual force F_i.

F_i plus reversed affective force

$\dot{P}_i$ thus $F_i + (-\dot{P}_i) = 0$ (or) $F_i - \dot{P}_i = 0$

Thus the principle of virtual work takes the form

$$\sum_i (F_i - \dot{P}_i) \delta r_i = 0$$

Again writing $F_i = F_i^a + f_i$

$$\sum_i (F_i^a - \dot{P}_i) \delta r_i + \sum_i f_i \cdot \delta r_i = 0$$

Dealing with the system for which the virtual work of the forces of constraints is zero. We write

$$\sum_i (F_i^a - \dot{P}_i) \delta r_i = 0$$

(Force of constraints is no more in picture)

It is better to drop subscript 'a'. Thus

$$\sum_i (F_i - \dot{P}_i) \delta r_i = 0 \qquad ...(2)$$

Which is called D'Alembert's principle. To satisfy eq. (2) we can't equate the coefficients of δr_i to zero. Since δr_i are not independent of each other and hence it is necessary to transform δr_i changes into the changes of G. Co-ordinates δq_j which are independent of each other. The coefficient of every δq_i will then equated to zero.

2.2: Lagrange's Equation from D'Alembert's Principle:

The co-ordinate transformation eq's are

$$r_i = r_i (q_1, q_2, ... q_{n_2}, t) \qquad ...(1)$$

So that $\dfrac{dri}{dt} = \dfrac{\partial ri}{\partial q_1} \dfrac{dq_1}{dt} + \dfrac{\partial r_i}{\partial q_2} \dfrac{dq_2}{dt} + ... \dfrac{\partial r_i}{\partial t} \dfrac{dt}{dt}$

$$V_i = \sum_j \frac{\partial r_i}{\partial q_j} \dot{q}_j + \frac{\partial r_i}{\partial t} \qquad ...(3)$$

Further infinitesimal displacement δr_i can be connected with δq_j as

$$\delta r_i = \sum_i \frac{\partial r_i}{\partial q_j} \delta q_j + \frac{\partial r_i}{\partial t} \delta t$$

But last term is zero since in virtual displacement only co-ordinate displacement is considered and not that of time.

$$\therefore \quad \delta r_i = \sum_j \frac{\partial r_i}{\partial q_j} \delta q_j$$

Now we write eq (2) as

$$\sum_i (F_i - \dot{P}_i) \delta r_i = 0$$

$$\sum_{ij} F_i \frac{\partial r_i}{\partial q_j} \delta q_j - \sum_{ij} \dot{P}_i \frac{\partial r_i}{\partial q_j} \delta q_j = 0$$

we define $\sum_i F_i \frac{\partial r_i}{\partial q_j} = Q_j$

(The components of generalized force)

As discussed under generalized force q's need not have dimensions of length and similarly it is not necessary for the Q's to have the dimensions of force. But it is necessary that the product $Q_j \ \delta q_j$ must have the dimensions of work. Thus above eq. takes the form

$$\sum_j Q_j \delta q_j - \sum_{ij} \dot{P}_i \left(\frac{\partial r_i}{\partial q_j} \right) \delta q_j = 0 \quad ...(4)$$

Let us evaluate the second term of eq. (4)

$$\sum_{ij} \dot{P}_i \left(\frac{\partial r_i}{\partial q_j} \right) \delta q_j = \sum_{ij} m_i \ddot{r}_i \left(\frac{\partial r_i}{\partial q_j} \right) \delta q_j$$

$$= \sum_{ij} \left\{ \frac{d}{dt} \left(m_i r_i \frac{\partial r_i}{\partial q_j} \right) - m_i r_i \frac{d}{dt} \left(\frac{\partial r_i}{\partial q_j} \right) \right\} \delta q_j$$

$$= \sum_{ij} \left\{ \frac{d}{dt} \left(m_i V_i \frac{\partial r_i}{\partial q_j} \right) - m_i V_i \frac{d}{dt} \left(\frac{\partial r_i}{\partial q_j} \right) \right\} \delta q_j \quad ...(5)$$

Further

$$\frac{d}{dt} \left(\frac{\partial r_i}{\partial q_j} \right) = \sum_k \frac{\partial}{\partial q_k} \left(\frac{\partial r_i}{\partial q_j} \right) \frac{\partial q_k}{\partial t} + \frac{\partial}{\partial t} \left(\frac{\partial r_i}{\partial q_j} \right) \frac{dt}{dt}$$

$$= \sum_k \frac{\partial^2 r_i}{\partial q_k \partial q_j} \dot{q}_k + \frac{\partial^2 r_i}{\partial t . \partial q_j}$$

$$= \sum_k \frac{\partial}{\partial q_j} \left(\frac{\partial r_i}{\partial q_k} \dot{q}_k + \frac{\partial r_i}{\partial t} \right)$$

$$= \frac{\partial}{\partial q_j} \left(\frac{dr_i}{dt} \right) = \frac{\partial V_i}{\partial q_j} \quad ...(6)$$

Also differentiating eq. (3) write $\dot{q}_j$ we get

$$\frac{\partial V_i}{\partial \dot{q}_j} = \frac{\partial r_i}{\partial q_j} \quad ...(7)$$

Putting eq. (6) and (7) in eq. (5), we get

$$\sum_{ij} \dot{P}_i \left(\frac{\partial r_i}{\partial q_j} \right) \delta q_j = \sum_{ij} \left\{ \frac{d}{dt} m_i V_i \left(\frac{\partial V_i}{\partial \dot{q}_j} \right) - m_i V_i \left(\frac{\partial V_i}{\partial q_j} \right) \right\} \delta q_j = \sum_j \left\{ \frac{d}{dt} \left[\frac{\partial}{\partial \dot{q}_j} \left(\sum_i \frac{1}{2} m_i V_i^2 \right) \right] - \frac{\partial}{\partial q_j} \left(\sum_i \frac{1}{2} m_i V_i^2 \right) \right\} \delta q_j$$

With this substitution eq. (4) becomes

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

Where for $\sum_i \frac{1}{2} m_i V_i^2$ 'T' is written since it represents the total $K.E.$ of the system further

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

Since the constraints are holomonic.

q_j are independent of each other and hence to satisfy above eq. the coefficient of each δq_j should separately vanish (i.e.)

$$\left[\sum_j \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} \right] = Q_j \quad ...(8)$$

As 'j' from 1 to n there will be 'N' such second order eq's.

2.3:Virtual Displacement:

Any imaginary displacement, which is considered with the constrained relation at any given instant (The real time doesn't change) is, called virtual displacement.

Virtual work:

Work done by the forces due to virtual displacement is called virtual work.

Virtual work $\delta W = \sum_i F_i \,\delta r_i$

Here $F_i = F_1^{(a)} + f_i$

Where F_i^a is applied force and f_i is forces of constraints

'δr_i' is displacement of position co-ordinate only

It does not involve variation of time

$$\delta r_i = \delta r_i \left(q_1, q_2 \ldots\ldots q_n \right)$$

2.4 Principle of virtual work :

Suppose the system is in equilibrium the resultant force on any particle vanishes
(i.e.) $F_i = 0$ for all values of 'i'.

Now virtual work

$$\delta w_i = \sum_i F_i \,\delta r_i = 0$$

$$\sum_i F_i \,\delta r_i = 0$$

$$\sum_i F_i^{(a)} \,\delta r_i + \sum_i f_i \,\delta r_i = 0$$

If the virtual work done by the force of constraints is zero;

(i.e.) $\sum_i f_i \,\delta r_i = 0$ then virtual work done by the applied force on a system in equilibrium state vanishes.

(i.e.) $\delta w_i = \sum_i F_i^{(a)} \delta r_i = 0$

The necessary condition for state equilibrium is that the virtual work done by all the constrained forces vanishes. This is called the principle of virtual work.

2. 5 Hamiltonian Approach

In the Lagrangian formulation, the equations of motion for a system are obtained in terms of generalized coordinates and generalized velocities. These equations of motion form a set of second order differential equations. Hamilton proposed an alternative formulation by using generalized coordinates and generalized momenta. This formulation results in two sets of first order differential equations. Both the formulations are identical, but the Hamiltonian formulation is more fundamental to the foundations of statistical and quantum mechanics. This formulation is particularly useful when some of the generalized momenta are the constants of motion.

Hamiltonian observed ambiguity of Lagrangian function for which it consists of derivative co-ordinates $(\dot{q}_j)$ and underivative co-ordinates are not given equal status (or) equal flooding. Some co-ordinates are not derivative co-ordinates and some other co-ordinates are underivative co-ordinates.

Hamiltonian simply place momenta co-ordinate in the case of $\dot{q}_j$ co-ordinate and the entire co-ordinates are given equal status. The new function would be written as $H(q_j, P_j, t)$

$H(q_j, p_j, t)$ Hamiltonian quantity is nothing but energy quantity and it is sum of $K.E$ and $P.E$

$\therefore \quad H = T + V$

According to Hamiltonian approach if any particle moves in free space requires '6' co-ordinates (3 are Generalized co-ordinates and rest of '3' are momenta co-ordinates). If 'N' particle body moves in the free space it requires 6N co-ordinates. Out of '6N' co-ordinates 3N are G. co-ordinates and remaining 3N are momenta co-ordinates. The 'N' particle body reduced to single particle known as *system point*. This system point can travel through '6N' dimensional space. Then the space is known as *phase space*.

2.6 Hamilton variational principle

The principle states that the integral $\int_{t_1}^{t_2} (T - V) dt$ shall have a stationary value (extremum). Where T is the $K.E$ of the mechanical system. It is a function of position   co-ordinates and their derivatives. V is the P.E. of the mechanical system, it is a function of co-ordinates only. Such a system for which V is purely a function of co-ordinates is called conservative system.

Statement : Hamilton's principle for conservative system is stated as follows.

The motion of the system from time 't_1' to time t_2 is such that the line integral

$$I = \int_{t_1}^{t_2} L dt \quad \text{where}$$

$L = T - V$ is an extremum for the path of motion.

Deduction: Let us consider a conservative system of particles employing the generalized co-ordinates. The integral can be written as

$$\int_{t_1}^{t_2} [T(q_j, \dot{q}_j) - V(q_j)] dt$$

$\therefore$ According to Hamilton's principle we have

$$\delta \int_{t_1}^{t_2} [T(q_j, \dot{q}_j) - V(q_j)] dt = 0$$

$$\Rightarrow \int_{t_1}^{t_2} \sum_j \left[\left(\frac{\partial T}{\partial q_j} \delta q_j + \frac{\partial T}{\partial \dot{q}_j} \delta \dot{q}_j \right) - \frac{\partial V}{\partial q_j} \delta q_j \right] dt = 0$$

$$\Rightarrow \int_{t_1}^{t_2} \sum_j \left(\frac{\partial T}{\partial q_j} - \frac{\partial V}{\partial q_j} \right) \delta q_j \, dt + \int_{t_1}^{t_2} \sum_j \frac{\partial T}{\partial \dot{q}_j} \delta \dot{q}_j \, dt = 0$$

$$\Rightarrow \int_{t_1}^{t_2} \sum_j \left(\frac{\partial T}{\partial q_j} - \frac{\partial V}{\partial q_j} \right) \delta q_j \, dt + \int_{t_1}^{t_2} \sum_j \frac{\partial T}{\partial \dot{q}_j} \frac{d}{dt}(\delta q_j) dt = 0$$

Integrating by parts the second term we get

$$\Rightarrow \int_{t_1}^{t_2} \sum_j \left(\frac{\partial T}{\partial q_j} - \frac{\partial V}{\partial q_j} \right) \delta q_j \, dt + \left[\sum_j \frac{\partial T}{\partial q_j} \delta q_j \right]_{t_1}^{t_2} - \int_{t_1}^{t_2} \sum_j \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) \delta q_j dt = 0$$

Since in such a variation there is no co-ordinate variation at end points,

$$\delta q_j \Big|_{t_1}^{t_2} = 0$$

Hence eq. Reduces to

$$\int_{t_1}^{t_2} \sum_j \left(\frac{\partial T}{\partial q_j} - \frac{\partial V}{\partial q_j} \right) \delta q_j \, dt - \int_{t_1}^{t_2} \sum_j \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) \delta q_j \, dt = 0$$

$$\int_{t_1}^{t_2} \sum_j \left[\frac{\partial T}{\partial q_j} - \frac{\partial V}{\partial q_j} - \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) \right] \delta q_j \, dt = 0$$

Since each δq_j are independent of each other, the coefficient of every δq_j should be equated to zero to satisfy above equation.

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

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

Part now for conservative systems 'V' is not a function of velocity '$\dot{q}_j$' but only of co-ordinates

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

Here we introduce the concept of scalar function called Lagrangian 'L' for a conservative system and is equal to $(T - V)$ thus the above equation then takes the form

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

Thus set of eq's. is called Lagrangian eqs.of motion.

2.7 Hamilton's canonical equations of motion

In Hamiltonian postulation we provide generalized momenta, an independent status placing it on equal footing with the G.co-ordinates. Hamiltonian is then to be regarded in general as a function of the position co-ordinates q_j, the momenta P_j and the time 't' (i.e.)

$$H = (q_j, P_j, t)$$

The differential of 'H' gives

$$dH = \sum_j \frac{\partial H}{\partial q_j} dq_j + \sum_j \frac{\partial H}{\partial P_j} dP_j + \frac{\partial H}{\partial t} dt \qquad ...(1)$$

Further 'H' is defined as

$$H = \sum_j P_j \dot{q}_j - L$$

So that

$$dH = \sum_j \dot{q}_j \, dP_j + \sum P_j d\dot{q}_j - dL \quad ...(2)$$

But Lagrangian is

$$L = L(q_j, \dot{q}_j, t)$$

So that

$$dL = \sum_j \frac{\partial L}{\partial q_j} dq_j + \sum_j \frac{\partial L}{\partial \dot{q}_j} d\dot{q}_j + \frac{\partial L}{\partial t} dt$$
$$...(3)$$

Putting the value of dL from eq. (3) in eq. (2), we get

$$dH = \sum_j \dot{q}_j \, dP_j + \sum_j P_j \, d\dot{q}_j - \sum_j \frac{\partial L}{\partial q_j} dq_j - \sum_j \frac{\partial L}{\partial \dot{q}_j} d\dot{q}_j - \frac{\partial L}{\partial t} dt$$
$$...(4)$$

Recognising $\dfrac{\partial L}{\partial \dot{q}_j} = P_j$ and $\dfrac{\partial L}{\partial q_j} = \dot{P}_j$ and putting these in eq. (4), we get

$$dH = \sum_j \dot{q}_j \, dP_j + \sum_j P_j d\dot{q}_j - \sum_j \dot{P}_j \, dq_j - \sum_j P_j \, d\dot{q}_j - \frac{\partial L}{\partial t} dt$$

$$dH = \sum_j \dot{q}_j \, dP_j - \sum_j \dot{P}_j \, dq_j - \frac{\partial L}{\partial t} dt \quad ...(5)$$

Comparing coefficients in eq. (5) and eq. (1), we get

$$\left. \begin{aligned} \dot{q}_j &= \frac{\partial H}{\partial P_j} \\[2mm] \dot{P}_j &= -\frac{\partial H}{\partial q_j} \end{aligned} \right\} \qquad ...(6)$$

$$-\frac{\partial L}{\partial t} = \frac{\partial H}{\partial t} \qquad ...(7)$$

eq. (6) is known as Hamilton's canonical eqs. of motion.

2.8 Conservation Theorems

In many problems a number of first integrals of the equations of motion can be obtained immediately; by this we mean relations of the type

$f(q_1, q_2, \ldots \dot{q}_1, \dot{q}_2, \ldots, t) = $ Constant, which are first-order differential equations.

Consider as an example a system of mass points under the influence of forces derived from potentials dependent on position only. Then

$$\frac{\partial L}{\partial \dot{x}_i} \equiv \frac{\partial T}{\partial \dot{x}_i} - \frac{\partial V}{\partial \dot{x}_i} = \frac{\partial T}{\partial \dot{x}_i} = \frac{\partial}{\partial \dot{x}_i} \Sigma \frac{1}{2} m_i \left(\dot{x}_i^2 + \dot{y}_1^2 + \dot{z}_i^2 \right)$$

$$= m_i \dot{x}_i = p_{ix}, \text{------(1)}$$

which is the x component of the linear momentum associated with the ith particle. This result suggests an obvious extension to the concept of momentum. The generalized momentum associated with the coordinate q_j shall be defined as

$$p_j = \frac{\partial L}{\partial \dot{q}_j}$$

The terms canonical momentum or conjugate momentum are often also used for p_j. If q_j is not a Cartesian coordinate, q_j does not necessarily have the dimensions of a linear momentum. If there is a velocity-dependent potential, then even with a Cartesian coordinate q_j the associated generalized momentum will not be identical with the usual mechanical momentum. Thus is the case of a group of particles in the electromagnetic field the Lagrangian is

$$L = \sum_i \frac{1}{2} m_i \dot{r}_i^2 - \sum_i q_i \phi(x_i) + \sum_i \frac{q_i}{c} A(x_i) \cdot \dot{r}_i \text{ -------(2)}$$

(q_i here denotes charge) and the generalized momentum conjugate to x_i is

$$p_{ix} = \frac{\partial L}{\partial \dot{x}_i} = m_i \dot{x}_i + \frac{q_i A_x}{c}, \text{ -------(3)}$$

that is, mechanical momentum plus an additional terms.

If the Lagrangian of a system does not contain a given coordinate q_j (although it may contain the corresponding velocity $\dot{q}_j$), then the coordinate is said to be cyclic or ignorable. The Lagrange equation of motion,

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_j} = 0,$$

reduces, for a cyclic coordinate, to

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} = 0 \quad \Rightarrow \quad \frac{dp_j}{dt} = 0,$$

which means that $p_j = $ constant. ------(4)

Hence we can state as a general conservation theorem that the generalized momentum conjugate to a cyclic coordinate is conserved.

Consider first a generalized coordinate q_j, for which a change dq_j represents a translation of the system as a whole in some given direction. An example would be one of the Cartesian coordinates of the center of mass of the system. Then clearly q_j cannot appear in T, for velocities are not affected by a shift in the origin, and therefore the partial derivative of T with respect to q_j must be zero. Further we will assume

conservative systems for which V is not a function of the velocities, so as to eliminate such anomalies as electromagnetic forces. The Lagrange equation of motion for a coordinate so defined then reduces to

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} = \dot{p}_j = -\frac{\partial V}{\partial q_j} = Q_j \qquad \text{-------(5)}$$

We will now show that the above eqn. is the equation of motion for the total linear momentum, i.e., that Q_j represents the component of the total force along the direction of translation of q_j and p_j is the component of the total linear momentum along this direction. In general, the generalized force Q_j is given by

$$Q_j = \sum_j \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j}$$

Since dq_j corresponds to a translation of the system along some axis, the vectors $\mathbf{r}_i\left(q_j\right)$ and $\mathbf{r}_i\left(q_j + dq_j\right)$ are related as show in fig. (1). By definition of derivative we have

$$\frac{\partial \mathbf{r}_i}{\partial q_j} = \underset{dq_j \to 0}{\mathbf{L}} \frac{\mathbf{r}_i\left(q_j + dq_j\right) - \mathbf{r}_i\left(q_j\right)}{dq_j}$$

$$= \frac{dq_j}{dq_j}\mathbf{n} = \mathbf{n} \ .$$

where $\mathbf{n}$ is the unit vector along the direction of translation. Hence

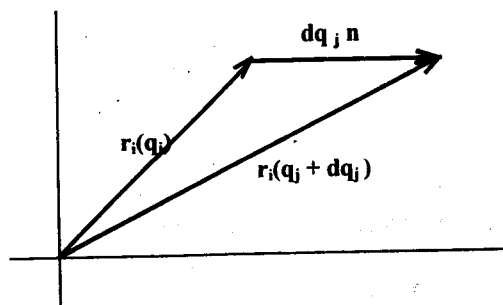$$Q_j = \sum \mathbf{F}_i \cdot \mathbf{n} = \mathbf{n} \cdot \mathbf{F} \ ,$$

which, as was stated, is the component of the total force in the direction of $\mathbf{n}$. To prove the other half of the statement notes that with the kinetic energy in the form

$$T = \frac{1}{2}\sum m_i \dot{\mathbf{r}}_i^2$$

the conjugate momentum is

$$p_j = \frac{\partial T}{\partial \dot{q}_j} = \sum_i m_i \dot{\mathbf{r}}_i \cdot \frac{\partial \dot{\mathbf{r}}_i}{\partial \dot{q}_j},$$

$$= \sum_i m_i \mathbf{v}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j},$$

$$p_j = \mathbf{n} \cdot \sum_i m_i \mathbf{v}_i$$

which again, as predicted, is the component of the total system linear momentum along $\mathbf{n}$

Fig 2.1: Change in a position vector under translation
of the system

Si)ose now that the translation coordinate q_j that we have been discussing is cyclic. Then q_j cannot
an ar in V and therefore

$$- \frac{\partial V}{\partial q_j} \equiv Q_j = 0.$$

But this is just the familiar conservation theorem for linear momentum — that if a given component of the total applied force vanishes, the corresponding component of the linear momentum is conserved.

In a similar fashion it can be shown that if a cyclic coordinate q_j is such that dq_j corresponds to a rotation of the system of particles around some axis, then the conservation of its conjugate momentum corresponds to conservation of an angular momentum. By the same argument as used above, T cannot contain q_j, for a rotation of the coordinate system cannot affect the magnitude of the velocities. Hence the partial derivative of T with respect to q_j must again be zero, and since V is independent of $\dot{q}_j$ we once again get equation (5).

The generalized force Q_j is again given by

$$Q_j = \sum_i \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j}$$

only the derivative now has a different meaning.
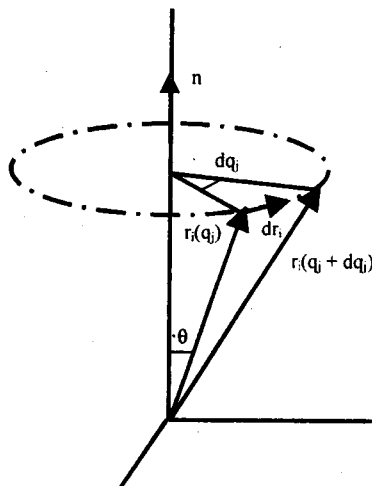
Fig 2.2: Change of a position vector under rotation of the system

Here the change in q_j must correspond to an infinitesimal rotation of the vector $\mathbf{r}_i$, keeping the magnitude of the vector constant. Form fig. (2) the magnitude of the derivative can easily be obtained:

$$\left| d\mathbf{r}_i \right| = r_i \sin\theta \, dq_j$$

and

$$\left| \frac{\partial \mathbf{r}_i}{\partial q_j} \right| = r_i \sin\theta$$

and its direction is perpendicular to both $\mathbf{r}_i$ and $\mathbf{n}$. Clearly the derivative can be written in vector form as

$$\frac{\partial \mathbf{r}_i}{\partial q_j} = \mathbf{n} \times \mathbf{r}_i$$

With this result the generalized force becomes

$$Q_j = \sum_i \mathbf{F}_i \cdot \mathbf{n} \times \mathbf{r}_i$$

$$= \sum_i \mathbf{n} \cdot \mathbf{r}_i \times \mathbf{F}_i$$

reducing to $Q_j = \mathbf{n} \cdot \sum_i \mathbf{N}_i = \mathbf{n} \cdot \mathbf{N}$

which proves the first part. A similar manipulation of p_j provides proof of the second part of the statement.

$$p_j = \frac{\partial T}{\partial \dot{q}_j} = \sum_i m_i \mathbf{v}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = \sum_i \mathbf{n} \cdot \mathbf{r}_i \times m_i \mathbf{v}_i$$

$$\mathbf{n} \cdot \sum_i \mathbf{L}_i = \mathbf{n} \cdot \mathbf{L}$$

Summarizing these results we see that if the rotation coordinate q_j is cyclic, then Q_j, which is the component of the applied torque along $\mathbf{n}$, vanishes, and the component of $\mathbf{L}$ along $\mathbf{n}$ is constant. We have recovered the angular momentum conservation theorem out of the general conservation theorem relating to cyclic coordinates.

If a coordinate corresponding to a displacement is cyclic, it means that a translation of the system, as if rigid, has no effect on the problem. In other words, if the system is invariant under translation along a given direction, the corresponding linear momentum is conserved. Similarly, the fact that a rotation coordinate is cyclic (and therefore the conjugate angular momentum conserved) indicates that the system is invariant under rotation about the given axis. Thus the momentum conservation theorems are closely connected with the symmetry properties of the system.

Another conservation theorem we should expect to obtain in the lagrangian formulation is the conservation of total energy for systems where the forces are derivable from potentials dependent only on position. It is possible to demonstrate a conservation theorem for which conservation of total energy represents only a special case. Consider a general Lagrangian, which will be a function of the coordinates q_j and the velocities $\dot{q}_j$ and may also depend explicitly on the time. Then the total time derivative of L is

$$\frac{dL}{dt} = \sum_j \frac{\partial L}{\partial q_j} \frac{dq_j}{dt} + \sum_j \frac{\partial L}{\partial \dot{q}_j} \frac{d\dot{q}_j}{dt} + \frac{\partial L}{\partial t}$$

Form Lagrange's equations

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

and it can be rewritten as

$$\frac{dL}{dt} = \sum_j \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right)\dot{q}_j + \sum_j \frac{\partial L}{\partial \dot{q}_j} \frac{d\dot{q}_j}{dt} + \frac{\partial L}{\partial t}$$

$$\Rightarrow \frac{dL}{dt} = \sum_j \frac{d}{dt}\left(\dot{q}_j \frac{\partial L}{\partial \dot{q}_j}\right) + \frac{\partial L}{\partial t}$$

It therefore follows that

$$\frac{d}{dt}\left(\sum_j \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} - L\right) + \frac{\partial L}{\partial t} = 0$$

The quantity in the brackets is often times called the energy function and will be denoted by h :

$$h(q, \dot{q}, t) = \sum_j \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} - L$$

and eq. can be looked on as giving the total time derivative of h :

$$\frac{dh}{dt} = -\frac{\partial L}{\partial t}$$

If the Lagrangian is not an explicit function of time, that is, if t does not appear in L explicitly but only implicitly through the time variation of q and $\dot{q}$, then the above Eq. says that h is conserved. The function h is the total energy of the system.

2.9 Principle of Least Action

·The Important variational principle associated with Hamiltonian formulation is the principle of least action. This is a more general type of variation of the path of a system which allows time as well as position co-ordinates to vary. At the end points of the path, position co-ordinates are held fixed but changes in the time are permitted. Such a variation is shown in fig.(3) and is denoted by a letter Δ instead of δ.

In mechanics the quantity $A = \int_{1}^{t_2} \sum_{j} p_j \dot{q}_j \, dt$ is defined as action.

The principle of least action for conservative system is then expressed as

$$\Delta \int_{1}^{t_2} \sum_{j} p_j \dot{q}_j \, dt = 0 \qquad \text{where } \Delta \text{ is the variation.}$$

Proof of Principle of Least Action
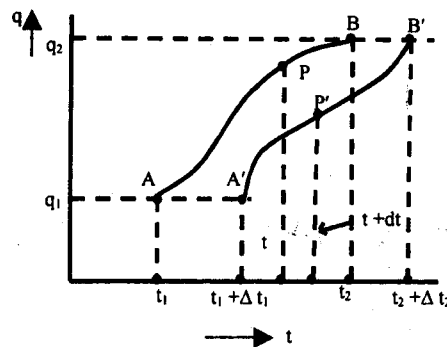
Fig 2.3: Δ - variation

The action A, can be written as,

$$A = \int_{t_1}^{t_2} \sum_{j} p_j \dot{q}_j \, dt = \int_{t_1}^{t_2} (L+H) dt$$

$$= \int_{t_1}^{t_2} L \, dt + H(t_2 - t_1) \qquad ...(2)$$

Since H is conserved, Δ- variation of action is

$$\Delta A = \Delta \int_{t_1}^{t_2} L \, dt + H \Delta(t_2 - t_1)$$

$$= \Delta \int_{t_1}^{t_2} L \, dt + H \Delta t \Big|_{t_1}^{t_2} \qquad ...(3)$$

Now we proceed to solve the remaining integral $\Delta \int_{t_1}^{t_2} L \, dt$,

Since t_1 and t_2 limits are also subject to change in this variation, Δ cannot be taken inside the integral. Let

$$\int_{t_1}^{t_2} L \, dt = I, \quad \text{so that } \dot{I} = L$$

Applying eq. (1), we get

$$\Delta I = \delta I + \dot{I} \, \Delta t . \quad \Rightarrow \quad \Delta \int_{t_1}^{t_2} L \, dt = \delta \int_{t_1}^{t_2} L \, dt + L \, \Delta t \Big|_{t_1}^{t_2} \quad \dots (4)$$

Putting eq. (4) in eq. (3), we get

$$\Delta A = \delta \int_{t_1}^{t_2} L \, dt + L \, \Delta t \Big|_{t_1}^{t_2} + H \, \Delta t \Big|_{t_1}^{t_2} \quad \dots (5)$$

The term i.e. $\delta \int_{t_1}^{t_2} L \, dt$ cannot be zero in consequence of Hamilton's principle. Hamilton's principle requires that $\delta q_j = 0$ at end points of the paths but in this variation $\Delta q_j = 0$ at the end points and not δq_j. Therefore the integral will not vanish. Using the nature of δ – variation, the integral can be expressed as

$$\delta \int_{t_1}^{t_2} L \, dt = \int_{t_1}^{t_2} \sum_j \left(\frac{\partial L}{\partial q_j} \delta q_j + \frac{\partial L}{\partial \dot{q}_j} \delta \dot{q}_j \right) dt$$

$$= \int_{t_1}^{t_2} \sum_j \left[\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j} \right) \delta q_j + \frac{\partial L}{\partial \dot{q}_j} \frac{d}{dt} \left(\delta q_j \right) \right] dt,$$

after putting $\dfrac{\partial L}{\partial q_j} = \dfrac{d}{dt} \left(\dfrac{\partial L}{\partial \dot{q}_j} \right)$ from Lagrange's equation of motion.

Thus $\delta \int_{t_1}^{t_2} L \, dt = \int_{t_1}^{t_2} \sum_j \left[\dfrac{d}{dt} \left(\dfrac{\partial L}{\partial \dot{q}_j} \delta q_j \right) \right] dt.$

Putting $\delta q_j = \Delta q_j - \dot{q}_j \, \Delta t$, we get

$$\delta \int_{t_1}^{t_2} L \, dt = \int_{t_1}^{t_2} \sum_j \left[\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j} \Delta q_j - \frac{\partial L}{\partial \dot{q}_j} \dot{q}_j \, \Delta t \right) \right] dt = \sum_j \left(\frac{\partial L}{\partial \dot{q}_j} \Delta q_j - \frac{\partial L}{\partial \dot{q}_j} \dot{q}_j \, \Delta t \right) \Big|_{t_1}^{t_2}$$

At end points $\Delta q_j = 0$. Therefore $\delta \int_{t_1}^{t_2} L \, dt = -\sum_j \frac{\partial L}{\partial \dot{q}_j} \dot{q}_j \, \Delta t \Big|_{t_1}^{t_2} = -\sum_j p_j \, \dot{q}_j \, \Delta t \Big|_{t_1}^{t_2}$

With this substitution, eq. (5) becomes

$$\Delta A = -\sum_j p_j \, \dot{q}_j \, \Delta t \Big|_{t_1}^{t_2} + L \, \Delta t \Big|_{t_1}^{t_2} + H \, \Delta t \Big|_{t_1}^{t_2}$$

$$= \left(H + L - \sum_j p_j \, \dot{q}_j \right) \Delta t \Big|_{t_1}^{t_2} = 0$$

since $H = \sum_j p_j \, \dot{q}_j - L$. Thus $\Delta A = \Delta \int_{t_1}^{t_2} \sum_j p_j \, \dot{q}_j \, dt = 0$

Which proves the principle of least action.

Summary :

1. If a system is subjected to a infinitesimal displacement consistent with the forces and constrains imposed on the system at the given instant 't' and if this change in the configuration of the system is not associated with a change in time (i.e.) there is no actual displacement during which forces and constraints may change and hence the displacement, δr_i is termed as virtual displacement.

2. The equation $\sum_i F_i^a \, \delta r_i = 0$ is termed as principle of virtual work.

3. The equation $\sum_i (F_i - \dot{P}_i).\delta r_i = 0$ is called D'Alembert's principle

4. The scalar function called Lagrangian 'L' for a conservative system is equal to $(T - V)$.

In terms of L, the equations of motion can be obtained as $\left[\dfrac{d}{dt}\left(\dfrac{\partial L}{\partial \dot{q}_j} \right) - \dfrac{\partial L}{\partial q_j} \right] = 0$

This set of equations are called Lagrangian equations of motion.

5. The generalized momentum associated with the coordinate q_j shall be defined as

$$p_j = \frac{\partial L}{\partial \dot{q}_j}$$

The terms canonical momentum or conjugate momentum are often also used for p_j .

If q_j is not a Cartesian coordinate, q_j does not necessarily have the dimensions of a linear momentum.

6. The generalized force Q_j is again given by

$$Q_j = \sum_i \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j}$$

7. The equations of motion defined in terms of generalized coordinates and momenta

$$\dot{q}_j = \frac{\partial H}{\partial P_j} \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\}$$
$$\dot{P}_j = -\frac{\partial H}{\partial q_j}$$

$$-\frac{\partial L}{\partial t} = \frac{\partial H}{\partial t}$$

are called Hamilton's equations of motion.

8. In mechanics the quantity $A \int_{t_1}^{t_2} \sum_j p_j \dot{q}_j \, dt$ is defined as action.

The principle of least action for conservative system is then expressed as

$$\Delta \int_{t_1}^{t_2} \sum_j p_j \dot{q}_j \, dt = 0$$

Key words: D'Alembert's Principle , Lagrangian equations . virtual displacement , virtual work. Hamilton's variational principle , Hamilton's canonical equations of motion , Principle of Least Action.

Self Assessment Questions:
1. State and prove the principle of least action.
2. Derive Lagrange equations of motion from Hamiltons principle.
3. Derive Lagrange equations of motion from D'Alembert's principle.

4 Derive Lagrange equations of motion

5.Derive Hamilton's equations of motion.

Reference books:

1.Classical Mechanics: H.Goldstein

2.Mechanics: Simon

3. Mechanics: Gupta, Kumar and Sharma

Classical Mechanics Part
Unit I
Lesson 3

Kinematics of a Rigid body

Objective: To learn the dynamics of a rigid body interms of Eular's angles, writing equations of motion, concept of infinitesimal rotations, Coriolis force and its influence in daily life

Structure:

3.1 The independent coordinates of a rigid body
3.2 Transformation matrix
3.3 The Euler angles
3.4 Infinitesimal rotations
3.5 Coriolis force

3.1 The Independent Coordinates of a Rigid Body

A rigid body with N particles can at most have $3N$ degrees of freedom, but these are greatly reduced by the constraints, which can be expressed as equations of the form

$$r_{ij} = c_{ij} \qquad \qquad ...(1)$$

Here r_{ij} is the distance between the ith particles and the $c's$ are constants. The actual number of degrees of freedom cannot be obtained simply by subtracting the number of constraint equations from $3N$, for there are $\frac{1}{2}N(N-1)$ possible equations of the form of eq. (1), which exceeds $3N$ for large N. In truth, the eq. (1) are not all independent. To fix a point in the rigid body it is not necessary to specify its distances to all other points in the body. One need only state the distances to any three other non-collinear points. Thus, once the positions of three of the particles of the rigid body are determined, the constraints fix the positions of all remaining particles. The number of degrees of freedom therefore cannot be more than nine.

There are in fact three equations of rigid constraints imposed on them,

$$r_{12} = c_{12}, \qquad r_{23} = c_{23} \qquad r_{13} = c_{13}$$

that reduce the number of degrees of freedom to six.
A rigid body in space thus needs six independent generalized coordinates to specify its configuration.

The configuration of a rigid body is completely specified by locating a Cartesian set of coordinates fixed in the rigid body. Three of the coordinates are needed to specify the coordinates of the origin of this "body" set of axes. The relative to a coordinate system parallel to the external axes, but with the same origin as the primed axes.
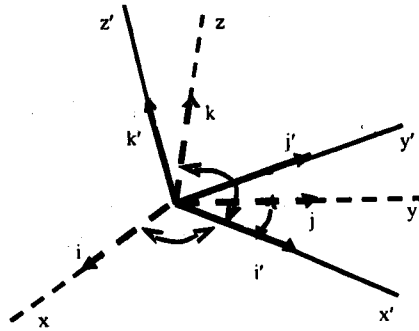
Fig 3.1: Direction Cosines of the body set of axes relative to an external set of axes.

There are many ways of specifying the orientation of a Cartesian set of axes relative to another set with common origin. A most fruitful procedure is to state the direction cosines of the primed axes relative to the unprimed. Thus the x' axis could be specified by its three direction cosines x_1, x_2, x_3 with respect to the x, y, z axes. If, as customary i, j, k are three unit vectors along x, y, z and i', j', k' perform the same function in the primed system, then these direction cosines are defined as

$$\alpha_1 = \cos(i', i) = i' \cdot i$$
$$\alpha_2 = \cos(i', j) = i' \cdot j \qquad ...(2)$$
$$\alpha_3 = \cos(i', k) = i' \cdot k$$

The vector i' can be expressed in terms of i, j, k by the relation

$$i' = (i' \cdot i)i + (i' \cdot j)j + (i' \cdot k)k$$
$$\Rightarrow i' = \alpha_1 i + \alpha_2 j + \alpha_3 k \qquad ...(3)$$

Similarly the direction cosines of the y' axis with x, y, z may be designated by β_1, β_2 and β_3, and these will be the components of j' in the unprimed reference frame.

$$j' = \beta_1 i + \beta_2 j + \beta_3 k \qquad ...(4)$$

An equation analogous to (4) can be written for k', with the direction cosines of the z' axis designated by γ's. Thus we can write

$$i = (i \cdot i')i' + (i \cdot j')j' + (i \cdot k')k'$$
$$\Rightarrow i = \alpha_1 i' + \beta_1 j' + \gamma_1 k' \qquad ...(5)$$

with analogous equations for j and k.

The coordinates of a point in a given reference frame are the components of the position vector, r, along the axes of the system. The x' coordinate is then given in terms of x, y, z by

$$x' = (r \cdot i') = a_1 x + a_2 y + a_3 z,$$

while for the other coordinates we obtain

$$y' = \beta_1 x + \beta_2 y + \beta_3 z$$
$$z' = \gamma_1 x + \gamma_2 y + \gamma_3 z \qquad ...(6)$$

What has been done here for the components of the r vector can obviously be done for any arbitrary vector. If G is some vector, then the component of G along the x' axis will be related to its x, y, z components by

$$G_{x'} = (G \cdot i') = a_1 G_x + a_2 G_y + a_3 G_z,$$

and so on. The set of nine direction cosines thus completely spells out the transformation between the two coordinate systems.

The connections between the direction cosines arise from the fact that the basis vectors in both coordinate systems are orthogonal to each other and have unit magnitude, in symbols,

$$\mathbf{i} \cdot \mathbf{j} = \mathbf{j} \cdot \mathbf{k} = \mathbf{k} \cdot \mathbf{i} = 0,$$

and $\mathbf{i} \cdot \mathbf{i} = \mathbf{j} \cdot \mathbf{j} = \mathbf{k} \cdot \mathbf{k} = 1$ ...(7)

with similar relations for $\mathbf{i'}, \mathbf{j'}$ and $\mathbf{k'}$. We can obtain the conditions satisfied by the nine coefficients by forming all possible dot products among the three equations for $\mathbf{i}, \mathbf{j}$ and $\mathbf{k}$ in terms of $\mathbf{i'}, \mathbf{j'}$ and $\mathbf{k'}$ (as in eq. 5), making use of the eq.7)

$$\alpha_l \alpha_m + \beta_l \beta_m + \gamma_l \gamma_m = 0, \quad l, m = 1,2,3; l \neq m,$$
$$\alpha_l^2 + \beta_l^2 + \gamma_l^2 = 1 \qquad l = 1,2,3 \qquad ...(8)$$

These two sets of three equations each are exactly sufficient to reduce the number of independent quantities from nine to three. Formally, the six equations can be combined into one by using the Kronecker δ-symbol δ_{lm}, defined by

$$\delta_{lm} = 1 \quad l = m$$
$$= 0 \quad l \neq m$$

Equation (8) can then be written as

$$\alpha_l \alpha_m + \beta_l \beta_m + \gamma_l \gamma_m = \delta_{lm} \qquad ...(9)$$

The use of direction cosines to describe the connections between two Cartesian coordinate systems nevertheless has a number of important advantages. With their aid many of the theorems about the motion of rigid bodies can be expressed with great elegance and generality, and in a form naturally leading to the procedures necessarily used in special relativity and quantum mechanics.

3.2 Transformation Matrix

Consider two successive transformations are made corresponding to two successive displacements of the rigid body. Let the first transformation from r to r' be denoted by **B** :

$$x'_k = b_{kj} x_j, \qquad\qquad (10)$$

and the succeeding transformation from r' to a third coordinate set r'' by **A** :

$$x''_i = a_{ik} x'_k \qquad\qquad (11)$$

The relation between x''_i and x_j can then be obtained by combining the two eqs (10) and (11)

$$x''_i = a_{ik} b_{kj} x_j$$

This may also be written as

$$x''_i = c_{ij} x_j \qquad\qquad (12)$$

where $c_{ij} = a_{ik} b_{kj}$ (13)

The successive application of two orthogonal transformation **A, B** is thus equivalent to a third linear transformation **C**. The resultant operator **C** can be considered as the product of the two operators **A** and **B** :

$$\mathbf{C} = \mathbf{AB},$$

and the matrix elements c_{ij} are by definition the elements of the square matrix obtained by multiplying the two square matrices **A** and **B**.

The elements of the transformation **D = BA** are

$$d_{ij} = b_{ik} a_{kj}, \qquad\qquad (14)$$

which generally do not agree with the matrix elements of **C**, matrix multiplication is associative, in a product of three or more matrices the order of the multiplications is unimportant :

$$(AB)C = A(BC) \qquad (15)$$

The sum $A+B$ is a matrix C whose elements are the sum of the corresponding elements of A and B:

$$c_{ij} = a_{ij} + b_{ij},$$

Of greater importance is the transformation inverse to A, the operation that changes r' back to r. This transformation will be called A^{-1} and its matrix elements designated by a'_{ij}. We then have the set of equations

$$x_i = a'_{ij} x'_j, \qquad (\because r' = Ar) \qquad (16)$$

which must be consistent with

$$x'_k = a_{ki} x_i \qquad (17)$$

Substituting x_i from (16), eq. (17) becomes

$$x'_k = a_{ki} a'_{ij} x'_j \qquad (18)$$

Since the components of r' are independent, eq. (18) is correct only if the summation reduces identically to x'_k. The coefficient of x'_j must be 1 for $j=k$ and zero for $j \neq k$; in symbols,

$$a_{ki} a'_{ij} = \delta_{kj} \qquad (19)$$

The left-hand side of eq. (19) is easily recognized as the matrix element for the product AA^{-1}, while the right-hand side is the element of the matrix known as the unit matrix 1:

$$1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad ...(20)$$

Equation (30) can therefore be written as

$$AA^{-1} = 1 \qquad (21)$$

which indicates the reason for the designation of the inverse matrix by A^{-1}. The transformation corresponding to 1 is known as the identity transformation, producing no change in the coordinate system:

$$x = 1x .$$

Similarly multiplying any matrix A by 1, in any order, leaves A unaffected:

$$1A = A1 = A$$

Instead of substituting x_i in eq. (18) in terms of x', one could equally as well demand consistency by eliminating x' from the two equations, leading in analogous fashion to

$$a'_{ij} a_{jk} = \delta_{ik}$$

In matrix notation this reads

$$A^{-1}A = 1 \qquad ---- (22)$$

Consider now the double sum

$$a_{kl} a_{ki} a'_{ij}$$

which can be written either as

$$c_{li} a'_{ij} \quad \text{with} \quad c_{li} = a_{kl} a_{ki}$$

or as $\quad a_{kl} d_{kj} \quad \text{with} \quad d_{kj} = a_{ki} a'_{ij}$

Applying the orthogonality conditions, $i.e. a_{ij} a_{ik} = \delta_{jk} \; j,k = 1,2,3$ the sum in the first form reduces to

$$\delta_{li} a'_{ij} = a'_{lj}$$

On the other hand, the same sum from the second point of view, and with the help of eq. (19) can be writte

$$a_{kl} \delta_{kj} = a_{jl}$$

Thus the elements of the direction matrix A and the reciprocal A^{-1} are related by

$$a'_{ij} = a_{ji} \qquad (23)$$

In general, the matrix obtained from A by interchanging rows and columns is known as the transposed matrix, indicated by $\tilde{A}$. Symbolically

$$A^{-1} = \tilde{A} \qquad (24)$$

If this result is substituted in eq. (22), we obtain

$$\tilde{A}A = 1, \qquad (25)$$

which is identical with the set of orthogonality conditions,

A rectangular matrix is said to be of dimension $m \times n$ if it has m rows and n columns, i.e. if the matrix element is a_{ij}, then i runs from 1 to m, and j from 1 to n. Clearly, the transpose of such a matrix has the dimension $n \times m$. If a vector column matrix is considered as a rectangular matrix of dimension $m \times 1$, the transpose of a vector is of dimension $1 \times m$, that is, a one-row matrix. The product AB of two rectangular matrices exists only if the number of columns of A is the same as the number of rows of B. This is an obvious consequence of the definition of the multiplication operation leading to a matrix element:

$$c_{ij} = a_{ik}b_{kj}$$

From this viewpoint, the product of a vector column matrix with a square matrix does not exist.

A useful commutation property of the product of a vector and a square matrix that

$$Ax = x\tilde{A}$$

A square matrix that is the same as its transpose,

$$A_{ij} = A_{ji}$$

is said to be symmetric. When the transpose is the negative of the matrix.

$$A_{ij} = -A_{ji}$$

the matrix is antisymmetric or skew symmetric. It is clear that in an antisymmetric matrix, the diagonal elements are always zero for any square matrix A, the matrix A_s defined as

$$A_s = \frac{1}{2}\left(A + \tilde{A}\right)$$

is symmetric, and a corresponding antisymmetric matrix can be defined as

$$A_a = \frac{1}{2}\left(A - \tilde{A}\right)$$

It obviously follows that

$$A = A_s + A_a$$

and $\qquad \tilde{A} = A_s + A_a$

Analogous to the definition (25) for an orthogonal matrix, a unitary matrix A satisfies the condition

$$A^+A = 1$$

Let A be considered an operator acting upon a vector F (or a single-column matrix F) to produce a vector G:

$$G = AF$$

If the coordinate system is transformed by a matrix B the components of the vector G in the new system will be given by

$$BG = BAF$$

which can also be written

$$BG = BAB^{-1}BF \qquad (26)$$

Eq. (26) can be interpreted as stating that the operator BAB^{-1} acting upon the vector F, expressed in the new system, produces the vector G, likewise expressed in the new coordinates. Therefore consider BAB^{-1} to be the form taken by the operator A when transformed to a new set of axes:

$$A' = BAB^{-1} \qquad (27)$$

Any transformation of a matrix having the form of eq. (27) is known a similarity transformation. The definition of matrix multiplication is identical with that for the multiplication of determinants. Hence

$$|AB| = |A| \cdot |B|$$

Since the determinant of the unit matrix is 1, the determinantal form of the orthogonality conditions, can be written

$$|\tilde{A}| \cdot |A| = 1$$

As the value of a determinant is unaffected by interchanging rows and columns, we can write·

$$|A|^2 = 1,$$

which implies that the determinant of an orthogonal matrix can only be +1 or -1 .

3.3 The Euler Angles

The six relations that express the orthogonality conditions, reduce the number of independent elements to three. But in order to characterize the motion of a rigid body there is an additional requirement the matrix elements must satisfy, beyond those implied by orthogonality.

Consider a simple matrix with the determinant -1 :

$$S = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = -1$$
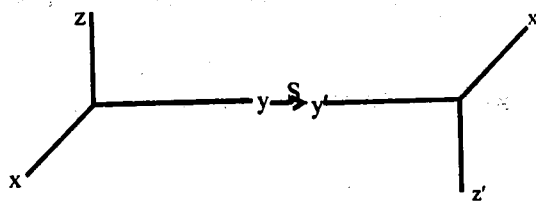

Fig 3.2: Inversion of the coordinate axes.

The transformation S has the effect of changing the sign of each of the components or coordinate axes (fig.1). Such an operation transforms a right-handed coordinate system into a left-handed one and is known as an inversion or reflection of the coordinate axes.

From the nature of this operation it is clear that an inversion of a right handed system into a left-handed one cannot be accomplished by any rigid change in the orientation of the coordinate axes. An inversion therefore never corresponds to a physical displacement of a rigid body.

In order to describe the motion of rigid bodies in the Lagrangian formulation of mechanics, it will therefore be necessary to seek three independent parameters specifying the orientation of a rigid body in such a manner that the corresponding orthogonal matrix of transformation has the determinant +1 . Only when such generalized coordinates have been found can one write a Lagrangian for the system and obtain the Lagrangian equations of motion. A number of such sets of parameters have been described in the literature. But the most common and useful are the Euler angles.

The Euler angles are then defined as the three successive angles of rotation. Within limits, the choice of rotation angles is arbitrary. The main convention that will be followed here is used widely in celestial

mechanics, applied mechanics, and frequently in molecular and solid state physics. Other conventions will be described below.
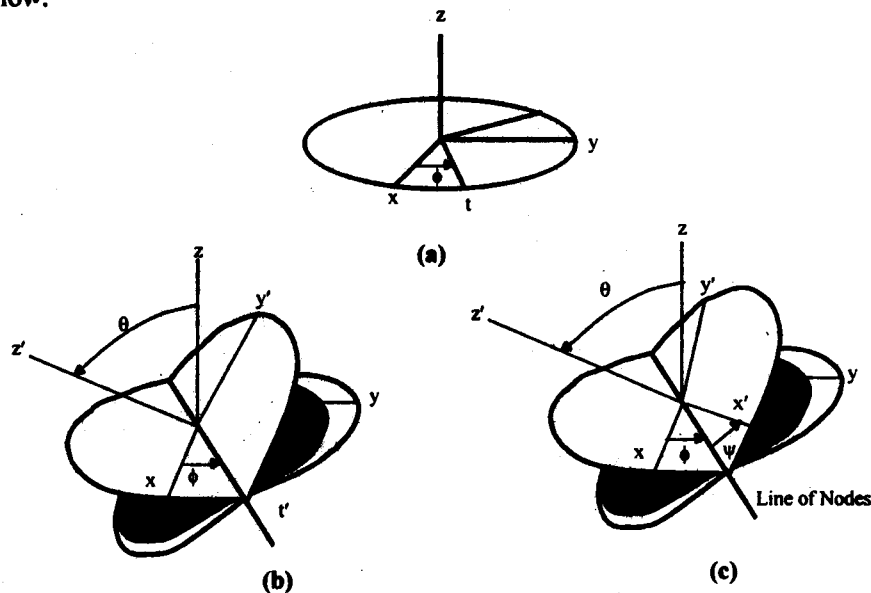
Fig 3.3: The rotations defining the Eulerian angles.

The sequence employed here is started by rotating the initial system of axes, xyz, by an angle ϕ counterclockwise about the z axis, and the resultant coordinate system is labeled the $\xi\eta\zeta$ axes. In the second stage the intermediate axes, $\xi\eta\zeta$ are rotated about the ξ axis counterclockwise by an angle θ to produce another intermediate set, the $\xi'\eta'\zeta'$ axes. The ξ' axis is at the intersection of the xy and $\xi'\eta'$ planes and is known as the line of nodes. Finally the $\xi'\eta'\zeta'$ axes are rotated counterclockwise by the angle ψ thus the ζ' axis to produce the desired $x'y'z'$ system of axes. Figure 2 illustrates the various stages of the sequence. The Euler angles θ, ϕ and ψ thus completely specify the orientation of the $x'y'z'$ system relative to the xyz and can therefore act as the three needed generalized coordinates.

The elements of the complete transformation **A** can be obtained by writing the matrix as the triple product of the separate rotations, each of which has a relatively simple matrix form. Thus, the initial rotation about z can be described by a matrix **D**

$$\xi = Dx$$

where ξ and x stand for column matrices. Similarly the transformation from $\xi\eta\zeta$ to $\xi'\eta'\zeta'$ can be described by a matrix **C**.

$$\xi' = C\xi$$

and the last rotation to $x'y'z'$ by a matrix **B**

$$x' = B\xi'$$

Hence the matrix of the complete transformation

$$x' = Ax$$

is the product of the successive matrices,

$$A = BCD$$

Now the **D** transformation is a rotation about z, and hence has a matrix of the form is

$$D = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The C transformation corresponds to a rotation about ξ, with the matrix

$$C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}$$

and finally B is a rotation about ζ' and therefore has the same form as D :

$$B = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The product matrix A = BCD then follows as

$$A = \begin{pmatrix} \cos\psi\cos\phi - \cos\theta\sin\phi\sin\psi & \cos\psi\sin\phi + \cos\theta\cos\phi\sin\psi & \sin\psi\sin\theta \\ -\sin\psi\cos\phi - \cos\theta\sin\phi\cos\psi & -\sin\psi\sin\phi + \cos\theta\cos\phi\cos\psi & \cos\psi\sin\theta \\ \sin\theta\sin\phi & -\sin\theta\cos\phi & \cos\theta \end{pmatrix}$$

The inverse transformation from body coordinates to space axes

$$x = A^{-1}x'$$

is then given immediately by the transposed matrix $\tilde{A}$:

$$A^{-1} = \tilde{A} = \begin{pmatrix} \cos\psi\cos\phi - \cos\theta\sin\phi\sin\psi & -\sin\psi\cos\phi + \cos\theta\sin\phi\cos\psi & \sin\psi\sin\theta \\ \cos\psi\sin\phi + \cos\theta\cos\phi\sin\psi & -\sin\psi\sin\phi + \cos\theta\cos\phi\cos\psi & -\sin\psi\cos\phi \\ \sin\theta\sin\psi & \sin\theta\cos\psi & \cos\theta \end{pmatrix}$$

Verification of the multiplication, and demonstration that A represents a proper, orthogonal matrix will be left to the exercises.

The initial rotation could be taken about any of the three Cartesian axes. In the subsequent two rotations, the only limitation is that no two successive rotations can be about the same axis. A total of twelve conventions is therefore possible in defining the Euler angles (in a right-handed coordinate system). The two most frequently used conventions differ only in the choice of axis for the second rotation. A third convention is commonly used in engineering applications relating to the orientation of moving vehicles such as aircraft and satellites.

3.4 Infinitesimal Rotations

Let A and B are two 'vectors' associated with transformations A and B . Then to qualify as vectors they must be commutative in addition:

$$A + B = B + A$$

But the addition of two rotations, i.e., one rotation performed after another, it corresponds to the product AB of the two matrices. This conclusion, that the sum of finite rotations depends on the order of the rotations, is strikingly demonstrated by a simple experiment. Illustrates the sequence of events in rotating a block first through 90 about the z axis fixed in the block, and then 90 about the y' axis, while presents the same rotations in reverse order.
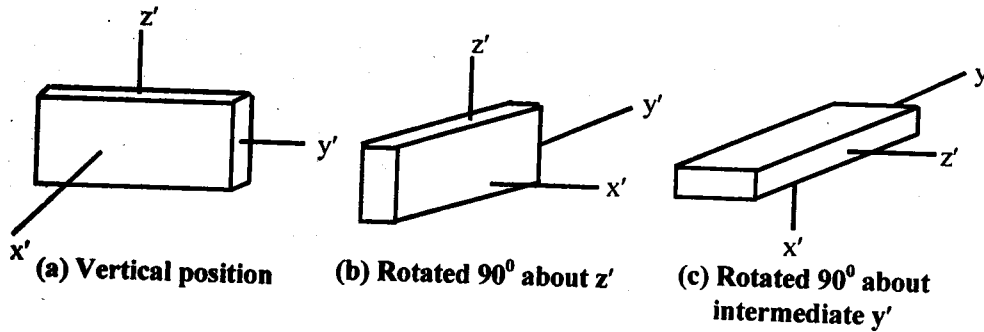
(a) Vertical position   (b) Rotated 90^0 about z'   (c) Rotated 90^0 about intermediate y'

Fig 3.4: Illustrating the effect of two rotations performed in a given order

(a) Vertical position   (b) Rotated 90^0 about y'   (c) Rotated 90^0 about
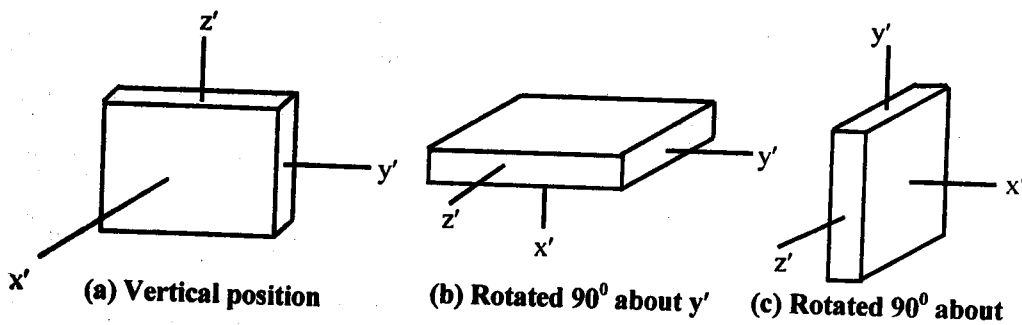
Fig 3.5: The two rotations shows in fig 3.4, but performed in reverse order.

An infinitesimal rotation is an orthogonal transformation of coordinate axes in which the components of a vector are almost the same in both sets of axes- the change is infinitesimal. Thus the x_1' component of some vector r would be practically the same as x_1, the difference being extremely small:

$$x_1' = x_1 + \epsilon_{11} x_1 + \epsilon_{12} x_2 + \epsilon_{13} x_3$$

The matrix elements $\epsilon_{11}, \epsilon_{12}$ and so on are to be considered as infinitesimals, so that in subsequent calculations only the first nonvanishing order in ϵ_{ij} need be retained. For any general component x_i' the equations of infinitesimal transformation can be written as

$$x_i' = x_i + \epsilon_{ij} x_j.$$

$$\Rightarrow x_i' = \left(\delta_{ij} + \epsilon_{ij} \right) x_j \qquad (1)$$

The quantity δ_{ij} will be recognized as the element of the unit matrix, and eq. (101) appears in matrix notation as

$$x' = (1 + \epsilon) x \qquad (2)$$

States that the typical form for the matrix of an infinitesimal transformation is $1 + \epsilon$, that is, it is almost the identity transformation, differing at most by an infinitesimal operator.

Infinitesimal transformation; words are commute. If $1 + \epsilon_1$ and $1 + \epsilon_2$ are two infinitesimal transformation, then one of the possible products is

$$(1 + \epsilon_1)(1 + \epsilon_2) = 1^2 + \epsilon_1 1 + 1 \epsilon_2 + \epsilon_1 \epsilon_2$$

$$= 1 + \epsilon_1 + \epsilon_2 \qquad (3)$$

neglecting higher order infinitesimals. The product in reverse order merely interchanges ϵ_1 and ϵ_2, this has no effect on the result, as matrix addition is always commutative.

If $A = 1 + \epsilon$ is the matrix of the transformation, then the inverse is

$$A^{-1} = 1 - \epsilon \qquad (4)$$

$$AA^{-1} = (1 + \epsilon)(1 - \epsilon) = 1$$

in agreement with the definition for the inverse matrix, the orthogonality of A implies that $\tilde{A} = (1 + \tilde{\epsilon})$ must be equal to A^{-1}. Hence the infinitesimal matrix is antisymmetric:

$$\tilde{\epsilon} = -\epsilon$$

Since the diagonal elements of an antisymmetric matrix are necessarily zero there can be only three distinct elements in any 3×3 antisymmetric matrix. Hence there is no loss of generality in writing ϵ in the form

$$\epsilon = \begin{pmatrix} 0 & d\Omega_3 & -d\Omega_2 \\ -d\Omega_3 & 0 & d\Omega_1 \\ d\Omega_2 & -d\Omega_1 & 0 \end{pmatrix} \qquad (5)$$

The three quantities $d\Omega_1, d\Omega_2, d\Omega_3$ are clearly to be identified with the three independent parameters specifying the rotation. The change in the components of a vector under the infinitesimal transformation of the coordinates system can be expressed by the matrix equation

$$x' - x \equiv dx = \epsilon x \qquad (6)$$

with in expanded form, with ϵ given by (5) becomes

$$dx_1 = x_2 d\Omega_3 - x_3 d\Omega_2$$

$$dx_2 = x_3 d\Omega_1 - x_1 d\Omega_3 \qquad (7)$$

$$dx_3 = x_1 d\Omega_2 - x_2 d\Omega_1$$

The right-hand side of each in the form of a component of the cross product of two vectors, namely the cross product of r with a vector $d\Omega$ having components $d\Omega_1, d\Omega_2, d\Omega_3$. Therefore write equivalently as

$$dr = r \times d\Omega \qquad (8)$$

The vector r transforms under an orthogonal matrix B according to the relations .

$$x_i' = b_{ij} x_j \qquad (9)$$

Let Φ go to the limit of an infinitesimal angle $d\Phi$ the corresponding formula for an infinitesimal rotation can be obtained. In this limit $\cos \Phi$ approaches unity, and $\sin \Phi$ goes to Φ; the resultant expression for the infinitesimal charge in r is then

$$r' = r \cos \Phi + n(n \cdot r)(1 - \cos \Phi) + (r \times n)\sin \Phi$$

$$r' - r \equiv dr = r \times n \, d\Phi \qquad (10)$$

Comparison with indicates that $d\Omega$ is indeed a vector and is determined by

$$d\Omega = n \, d\Phi \qquad (11)$$

The concept of an infinitesimal rotation provides a powerful tool for describing the motion of a rigid body in time. Consider some arbitrary vector G involved in the mechanical problem, such as the position vector of a point in the body, or the total angular momentum. Usually such a vector will vary in time as the body moves, but the change will often depend on the coordinate system to which the observations are referred. For example, if the vector happens to be the radius vector from the origin of the body set of axes to a point in the rigid body then, clearly, such a vector appears constant when measured by the body set of axes.

The change in a time dt of the components of a general vector G as seen by an observer in the body system of axes will differ from the corresponding change as seen by an observer in the space system. A relation between the two differential changes in G can be derived on the basis of physical arguments. We can write that the only difference between the two is the effect of rotation of the body axes:

$$(dG)_{space} = (dG)_{body} + (dG)_{rot}$$

Consider now a vector fixed in the rigid body. As the body rotates there is of course no change in the components of this vector as seen by the body observer, i.e., relative to body axes. The only contribution to $(d\mathbf{G})_{space}$ is then the effect of the rotation of the body. But since the vector is fixed in the body system, it rotates with it counterclockwise, and the change in the vector as observed in space is that given by and hence $(d\mathbf{G})_{rot}$ is given by

$$(d\mathbf{G})_{rot} = d\mathbf{\Omega} \times \mathbf{G}$$

For an arbitrary vector the change relative to the space axes is the sum of the two effects:

$$(d\mathbf{G})_{space} = (d\mathbf{G})_{body} + d\mathbf{\Omega} \times \mathbf{G} \qquad (12)$$

The time rate of change of the vector $\mathbf{G}$ as seen by the two observers is then obtained by dividing the terms by the differential time element dt under consideration:

$$\left(\frac{d\mathbf{G}}{dt}\right)_{space} = \left(\frac{d\mathbf{G}}{dt}\right)_{body} + \boldsymbol{\omega} \times \mathbf{G} \qquad (13)$$

Here ω is the instantaneous angular velocity of the body defined by the relation

$$\omega\, dt = d\Omega \qquad (14)$$

The vector ω lies along the axis of the infinitesimal rotation occurring between t and $t+dt$, a direction known as the instantaneous axis of rotation. In magnitude ω measures the instantaneous rate of rotation of the body.

A more formal derivation of the basic can be given in terms of the orthogonal matrix of transformation between the space and body coordinates. The component of $\mathbf{G}$ along the ith space axis is related to the components along the body axes:

$$G_i = \tilde{a}_{ij} G'_j = a_{ji} G'_j$$

As the body moves in time the components G'_j will change as will also the elements a_{ij} of the transformation matrix. Hence the change in G_i in a differential time element dt is

$$dG_i = a_{ji}\, dG'_j + da_{ji} G'_j \qquad (15)$$

Components in the two systems will then be the same instantaneously, but differentials will not be the same, since the two systems are moving relative to each other. Thus $G'_j = G_j$ but $a_{ji} dG'_j = dG'_i$, the prime emphasizing the differential is measured in the body axis system. The change in the matrix $\mathbf{A}$ in the time dt is thus a change from the unit matrix and therefore corresponds to the matrix ϵ of the infinitesimal rotation. Hence

$$da_{ji} = (\tilde{\epsilon})_{ij} - \epsilon_{ij}$$

using the antisymmetry property of ϵ. In terms of the permutation symbol ϵ_{ijk} the elements of ϵ are such that

$$-\epsilon_{ij} = -\epsilon_{ijk}\, d\Omega_k = -\epsilon_{ikj}\, d\Omega_k .$$

Equation can now be written

$$dG_i = dG'_i + \epsilon_{ikj}\, d\Omega_k G_j$$

The last term on the right will be recognized as the expression for the ith component of a cross product, so that the final expression for the relation between differentials in the two systems is

$$dG_i = dG'_i + (d\mathbf{\Omega} \times \mathbf{G}) \qquad (16)$$

which is the same as the ith component

The arbitrary nature of the vector $\mathbf{G}$ made use of in the derivation can be emphasized by writing as an operator equation acting n some given vector:

$$\left(\frac{d}{dt}\right)_s = \left(\frac{d}{dt}\right)_r + \boldsymbol{\omega} \times \qquad (17)$$

Here the subscripts s and r indicate the time derivatives are to be those observed in the space and body (rotating) system of axes, respectively.

3.5 The Coriolis Force

The Eq. $\left(\dfrac{d}{dt}\right)_s = \left(\dfrac{d}{dt}\right)_r + \omega \times$ is the basic kinematical law upon which the dynamical equations of motion for a rigid body are founded. But its validity is not restricted solely to rigid body motion. It may be used whenever we wish to discuss the motion of a particle, or system of particles, relative to a rotating coordinate system. The above equation provides the needed modifications of the equations of motion relative to the non inertial system fixed in the rotating earth.

The initial step is apply to the radius vector, $\mathbf{r}$, from the origin of the terrestrial system to the given particle:

$$\mathbf{v} = \mathbf{v}_r + \omega \times r \qquad (18)$$

where v_s and v_r are the velocities of the particle relative to the space and rotating set of axes, respectively, and ω is the (constant) angular velocity of the earth relative to the inertial system. In the second step Eq.(17) is used to obtain the time rate of change of v_s:

$$\left(\frac{d\mathbf{v}_s}{dt}\right)_s = \mathbf{a}_s = \left(\frac{d\mathbf{v}_s}{dt}\right)_r + \omega \times \mathbf{v}_s$$

$$= \mathbf{a}_r + 2(\omega \times \mathbf{v}_r) + \omega \times (\omega \times \mathbf{r}) \quad (19)$$ where v_s has been substituted from equation 18 and where a_s and a_r are the accelerations of the particle in the two systems. Finally, the equation of motion, which in the inertial system is simply

$$\mathbf{F} = m\mathbf{a}_s$$

expands, when expressed in the rotating coordinates, into the equation

$$\mathbf{F} - 2m(\omega \times \mathbf{v}_r) - m\omega \times (\omega \times \mathbf{r}) = m\mathbf{a}_r \quad (20)$$

To an observer in the rotating system it therefore appears as if the particles is moving under the influence of an effective force F_{eff}:

$$\mathbf{F}_{eff} = \mathbf{F} - 2m(\omega \times \mathbf{v}_r) - m\omega \times (\omega \times \mathbf{r}). \quad (21)$$

Let us examine the nature of the terms occurring in the last term is a vector normal to ω and pointing outward. Further. Its magnitude is $m\omega^2 r \sin\theta$. It will therefore be recognized that this term is simply the familiar centrifugal force. When the particle is stationary in the moving system the centrifugal force is the only added term in the effective force. When the particle is moving, the middle term known as the *Coriolis* force comes into play. The order of magnitude of both of these forces may easily be calculated for a particle on the earth's surface. The earth rotates counterclockwise about the North Pole with an angular velocity relative to the fixed stars:

$$\omega = \left(\frac{2\pi}{24 \times 3600}\right)\left(\frac{366.5}{365.5}\right) = 7.292 \times 10^{-5} \text{ sec}^{-1}$$
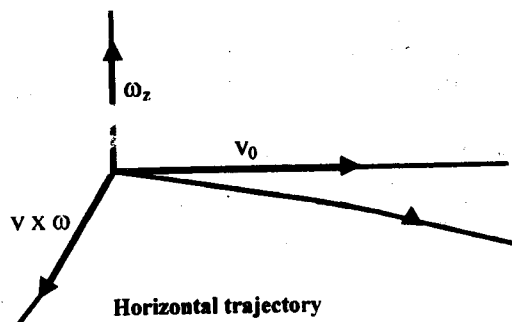

Fig 3.6: Direction of Coriolis deflection in the Northern Hemisphere.

The Coriolis force on a moving particle is perpendicular to both ω and v. In the Northern Hemisphere , where ω points out of the ground, the Coriolis force., $2m(v \times \omega)$ tends to deflect a projectile shot along the earth's surface, to the right of its direction of travel;. The Coriolis deflection reverses direction in the Southern Hemisphere and is zero at the Equator, where ω is horizontal. The magnitude of the Coriolis acceleration is always less than

$$2\omega v \simeq 1.5 \times 10^{-4} v.$$

which for a velocity of 10^5 cm/sec (roughly 2000 mph) is 15 cm/sec^2, or about 0.015g. The Coriolis force also plays a significant role in many oceanographic and meteorological phenomena involving displacements of masses of matter over long distances, such as the circulation pattern of the trade winds and the course of the Gulf stream.

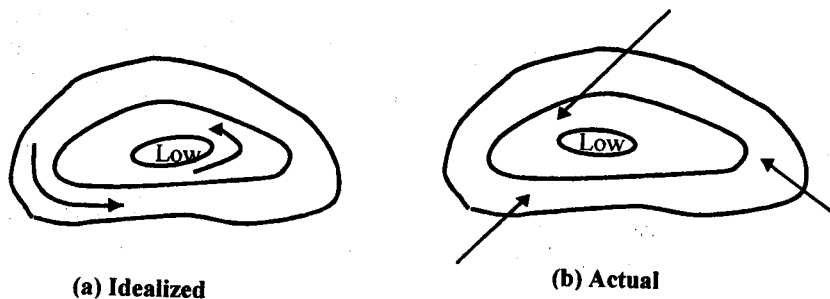(a) Idealized   (b) Actual

Fig 3.7: Cyclone Formation.

Masses of air tend to move, other things be equal, from regions of high pressure to regions of low pressure – the so-called pressure gradient flow. In the vertical direction the pressure gravitational forces roughly balance gradient so that it is only in the horizontal plane that there are persistent long-range motions of air masses – which we perceive as winds. The pressure gradient forces are quite modest, and comparable in magnitude to the Coriolis forces acting on air masses moving at usual speeds. In the absence of Coriolis forces the wind directions ideally, would be perpendicular to the isobars. However the Coriolis forces deflect the wind to the right of this direction in the sense indicated in the figure 6. The deflection to the right continues until the wind vector is parallel to the isobars and the Coriolis force is in the opposite direction to and ideally just balances, the pressure-gradient force. The wind then continues parallel to the isobars, circulating in the Northern Hemisphere in a counterclockwise direction about a center of low pressure. In the Southern Hemisphere the Coriolis force acts in the opposite direction, and the cyclonic direction.

Another classical instance where Coriolis force produces a measurable effect is in the deflection from the vertical of a freely falling particle. Since the particle velocity is almost vertical and ω lies in the north-south vertical plane, the deflecting force $2m(v \times \omega)$ is in the east-west direction. Thus, in the Northern Hemisphere, a body falling freely will be deflected to the east. Calculation of the deflection is greatly simplified by choosing the z axis of the terrestrial coordinate system to be along the direction of the upward vertical as previously defined. If the y axis is taken as pointing north, then the equation of motion in the x(east) direction is

$$m\frac{d^2 x}{dt^2} = -2m(\omega \times v)_x$$

$$= -2m\omega v_z \sin\theta, \qquad (22)$$

where θ is the co-latitude. The effect of the Coriolis force on v_z would constitute a small correction to the deflection, which itself is very small. Hence the vertical velocity appearing in (22) may be computed as if Coriolis forces were absent :

$$v_z = -gt \quad \text{and} \quad t = \sqrt{\frac{2z}{g}}.$$

With these values, Eq(22) may be easily integrated to give the deflection as

$$x = \frac{\omega g}{3} t^3 \sin\theta \quad \text{or} \quad x = \frac{\omega}{3}\sqrt{\frac{(2z)^3}{g}} \sin\theta.$$

An order or magnitude of the deflection can be obtained by assuming $\theta = \pi/2$ (corresponding to the Equator) and z = 100m.The deflection is then, roughly, x = 2.2 cm.

Effects due to the Coriolis terms also appear in atomic physics. Thus, two types of motion may occur simultaneously in polyatomic molecules: The molecule rotates as a rigid whole, and the atoms *vibrate* about their equilibrium positions. As a result of the vibrations the atoms are in motion relative to the rotating coordinate system of the molecule. The Coriolis term will then be different from zero and will cause the atoms to move in a direction perpendicular to the original oscillations. Perturbations in molecular spectra due to Coriolis forces thus appear as interactions between the rotational and vibrational levels of the molecule.

Summary :

A rigid body with N particles can at most have $3N$ degrees of freedom, but these are greatly reduced by the constraints, which can be expressed as equations of the form

$$r_{ij} = c_{ij}$$

Here r_{ij} is the distance between the ith particles and the c's are constants

A rigid body in space thus needs six independent generalized coordinates to specify its configuration The Euler angles are defined as the three successive angles of rotation. Within limits, the choice of rotation angles is arbitrary.
Euler angles θ, ϕ and ψ completely specify the orientation of the $x'y'z'$ system relative to the xyz and can therefore act as the three needed generalized coordinates.

An infinitesimal rotation is an orthogonal transformation of coordinate axes in which the components of a vector are almost the same in both sets of axes- the change is infinitesimal.
The concept of an infinitesimal rotation provides a powerful tool for describing the motion of a rigid body in time.
A general vector from the origin of the body set of axes appears constant when measured by the body set of axes. The change in a time dt of the components of a general vector G as seen by an observer in the body system of axes will differ from the corresponding change as seen by an observer in the space system. A relation between the two differential changes in G can be derived on the basis of physical arguments. We can show that the only difference between the two is the effect of rotation of the body axes:

The time rate of change of the vector G as seen by the two observers is

$$\left(\frac{dG}{dt}\right)_{space} = \left(\frac{dG}{dt}\right)_{body} + \omega \times G$$

Here ω is the instantaneous angular velocity of the body defined by the relation

$$\omega \, dt = d\Omega$$

The vector ω lies along the axis of the infinitesimal rotation occurring between t and $t+dt$, a direction known as the instantaneous axis of rotation. In magnitude ω measures the instantaneous rate of rotation of the body.

The Eq. $\left(\dfrac{d}{dt}\right)_s = \left(\dfrac{d}{dt}\right)_r + \omega \times$ is the basic kinematical law upon which the dynamical equations of motion for a rigid body are founded. But its validity is not restricted solely to rigid body motion. It may be used whenever we wish to discuss the motion of a particle, or system of particles, relative to a rotating coordinate system. The above equation provides the needed modifications of the equations of motion relative to the non inertial system fixed in the rotating earth.

. When the particle is stationary in the moving system the centrifugal force is the only added term in the effective force. When the particle is moving, the *Coriolis* force comes into play

Key words : Transformation matrix, Euler's angles , infinitesimal rotations ,Coriolis force

Self Assessment Questions:
1. Write a note on Euler's angles
2. What are Infinitesimal rotations?
3. Discuss the properties of Transformation matrix.
4. Explain the rate of change of vector in a moving frame of reference.
5. Explain a) centripetal acceleration b) Coriolis force.

Reference books:
1. Classical Mechanics: H.Goldstein
2. Mechanics: Simon
3. Mechanics: Gupta, Kumar and Sharma.

Classical Mechanics Part
UNIT-I
Lesson 4

The Rigid body Equations of Motion

Objective:

To learn

1) about the angular momentum , angular velocity, moment of inertia, rotational kinetic energy of a rigid body, moment of inertia of a rigid body.

2) Euler's equation of motion and also torque free motion of a rigid body.

2) dynamics of spinning top.

Structure :

4.1 Angular velocity

4.2 Angular momentum

4.3 Moments and products of inertia

4.4 Rotational kinetic energy

4.5 Moment of inertia of a rigid body

4.6 Euler's equations of motion

4.7 Torque free motion of a rigid body

4.7.1 Conservation of kinetic energy

4.7.2 Conservation of angular momentum

4.7.3 Heavy symmetric top

4.7.4 First integrals of equation of motion

4.7.6 Notational motion

4.7.7 Fast top.

4.1: ANUGALR VELOCITY

A rigid body can possess simultaneously the translational and rotational motion, and the equations of motion, governing the two types of motion, may involve both the translational and rotational co-ordinates.

Consider the rotation of rigid body about an axis OQ within the body and take, for example, a point P of the body at a perpendicular distance PN from the axis. Then P moves in a circle with center at N and radius PN. If r is the position-vector of P relative to the origin O, (may be taken as center of mass) the $PN = r \sin \theta$, where θ is angle of r with axis. In a time dt, P would have moved a distance $r \sin \theta \, d\phi$, along an arc of angle $d\phi$.

$$\therefore \quad \text{Linear speed} = r \sin \theta \frac{d\phi}{dt}$$

$$\Rightarrow \quad \left| \dot{r} \right| = r \sin \theta \frac{d\phi}{dt} = r \sin \theta \, \omega.$$

But $\quad \dot{r} = v = \vec{\omega} \times r$ ———(1)

If either the axis of rotation (direction of $\vec{\omega}$) or the magnitude $\dfrac{d\phi}{dt}$ changes with time, $\vec{\omega}$ will be a function of time. We can likewise consider changes in components of $\vec{\omega}$, along the space set or body set of axes. If either set of components are represented by $\omega_x, \omega_y, \omega_z$, then

$$\vec{\omega} = \mathbf{i}\omega_x + \mathbf{j}\omega_y + \mathbf{k}\omega_z$$

Vector property of $\vec{\omega}$ suggests that rotation along a certain axis can be decomposed into rotation about orthogonal axes and vice-versa.

4.2 ANGULAR MOMENTUM :
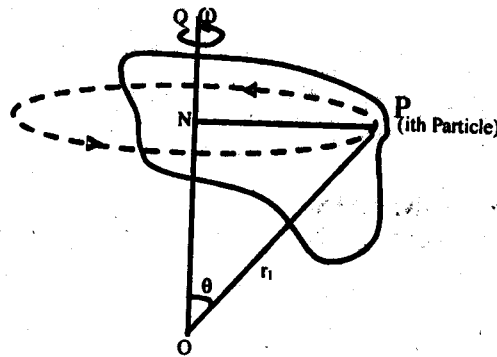
Fig 4.1:

If the rigid body is taken as a rigid collection of particles, then angular momentum is given by

$$L = \sum_i m(\mathbf{r}_i \times \mathbf{v}_i)$$

$$= \sum_i m_i \left(\mathbf{r}_i \times \dot{\mathbf{r}}_i \right)$$

$$= \sum_i m_i \left[\mathbf{r}_i \times \left(\vec{\omega} \times \mathbf{r}_i \right) \right]$$

$$\Rightarrow L = \sum_i m_i \left[\vec{\omega}(\mathbf{r}_i . \mathbf{r}_i) - \left(\mathbf{r}_i . \vec{\omega} \right)\mathbf{r}_i \right] \quad \text{(from (1))}$$

$$= \sum_i m_i \left[\vec{\omega} r_i^2 - \left(\mathbf{r}_i . \vec{\omega} \right)\mathbf{r}_i \right] \quad ...(2)$$

In terms of components of vectors involved in this equation, we write
$$L = \mathbf{i}L_x + \mathbf{j}L_y + \mathbf{k}L_z$$

$$= \sum_i m_i \left[\left(\mathbf{i}\omega_x + \mathbf{j}\omega_y + \mathbf{k}\omega_z \right) r_i^2 - \left(\mathbf{r}_i . \vec{\omega} \right)\left(\mathbf{i}x_i + \mathbf{j}y_i + \mathbf{k}z_i \right) \right]$$

since $\left(\mathbf{r}_i . \vec{\omega}\right)$ is scalar and will be a multiplying factor in every component.

$$\mathbf{i}L_x + \mathbf{j}L_y + \mathbf{k}L_z = \sum_i m_i \left[\left(\mathbf{i}\omega_x + \mathbf{j}\omega_y + \mathbf{k}\omega_z\right)r_i^2\right] - \left[\left(x_i\omega_x + y_i\omega_y + z_i\omega_z\right)\left(\mathbf{i}x_i + \mathbf{j}y_i + \mathbf{k}z_i\right)\right]$$

$$= \mathbf{i}\left[\omega_x \sum_i m_i \left(r_i^2 - x_i^2\right) - \omega_y \sum_i m_i x_i y_i - \omega_z \sum_i m_i x_i z_i\right] + \mathbf{j}\left[-\omega_x \sum_i m_i x_i y_i + \omega_y \sum_i m_i \left(r_i^2 - y_i^2\right) - \omega_z \sum_i m_i y_i z_i\right]$$

$$+ \mathbf{k}\left[-\omega_x \sum_i m_i x_i z_i - \omega_y \sum_i m_i y_i z_i + \omega_z \sum_i m_i \left(r_i^2 - z_i^2\right)\right]$$

$$\ldots(3)$$

Equating coefficients of $\mathbf{i}, \mathbf{j}, \mathbf{k}$.

$$L_x = \left[\omega_x \sum_i m_i \left(r_i^2 - x_i^2\right) - \omega_y \sum_i m_i x_i y_i - \omega_z \sum_i m_i x_i z_i\right] \qquad = I_{xx}\omega_x + I_{xy}\omega_y + I_{xz}\omega_z,$$

$$L_y = \left[-\omega_x \sum_i m_i x_i y_i + \omega_y \sum_i m_i \left(r_i^2 - y_i^2\right) - \omega_z \sum_i m_i y_i z_i\right] \qquad = I_{yx}\omega_x + I_{yy}\omega_y + I_{yz}\omega_z,$$

$$L_z = \left[-\omega_x \sum_i m_i x_i z_i - \omega_y \sum_i m_i y_i z_i + \omega_z \sum_i m_i \left(r_i^2 - z_i^2\right)\right] \qquad = I_{zx}\omega_x + I_{zy}\omega_y + I_{zz}\omega_z. \ldots(4)$$

4.3 MOMENTS AND PRODUCTS OF INERTIA

The nine coefficients as introduced above, can be written as a 3 x 3 arrangement of rows and columns as in a matrix nation.

$$\begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix} = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} \qquad \ldots(5)$$

$$\Rightarrow \quad \mathbf{L} = \mathbf{I}\vec{\omega} \qquad\qquad \ldots(6)$$

Equation (6) implies that when $\mathbf{I}$ operates on angular velocity vector $\vec{\omega}$, a physically different vector, the angular momentum, $\mathbf{L}$, results. $\mathbf{I}$ is therefore a different physical entity, termed as the moment of inertia tensor. The components of $\mathbf{I}$ have been arranged in matrix from in eq. (5) in a way to facilitate understanding of linear transformation : otherwise $\mathbf{I}$ is basically a tensor.

Now we come to physical meaning of the components of $\mathbf{I}$. The diagonal element of matrix form,

$$I_{xx} = \sum_i m_i \left(r_i^2 - x_i^2\right) = \sum_i m_i \left(y_i^2 + z_i^2\right)$$

$$I_{yy} = \sum_i m_i \left(r_i^2 - y_i^2\right) = \sum_i m_i \left(z_i^2 + x_i^2\right) \qquad \ldots(7)$$

$$I_{zz} = \sum_i m_i \left(r_i^2 - z_i^2\right) = \sum_i m_i \left(x_i^2 + y_i^2\right)$$

are called the moments of inertia coefficients. These are the moments of inertia of the body about the x, y, z – axes separately.

$$I_{xy} = I_{yx} = -\sum_i m_i x_i y_i$$

$$I_{yz} = I_{zy} = -\sum_i m_i y_i z_i \qquad \ldots(8)$$

$$I_{zx} = I_{xz} = -\sum_i m_i z_i x_i$$

are termed as the products of inertia associated with the corresponding co-ordinate planes. Obvious I is a symmetric tensor and is additive in nature; the moments of inertia of a body are the sums of those of its parts. A mathematical structure having nine components in three dimensions is termed a tensor of rank two. In this sense, I is a tensor of rank two and we shall write such an entity in type I to emphasize its tensorial character.

4.4 ROTATIONAL KINETIC ENERGY OF A RIGID BODY

The kinetic energy of motion of a system of particles is defined by

$$T = \frac{1}{2}\sum_i m_i v_i^2, \qquad \ldots(1)$$

where the summation extends to all the particles of the body and v_i is the linear velocity of the ith particle relative to O.
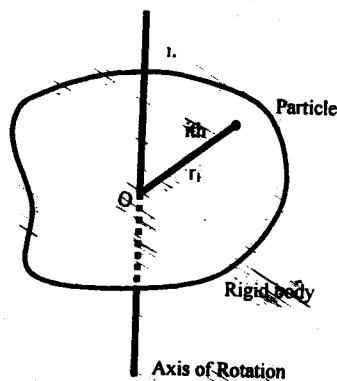
Fig 4.2.

But $\qquad \mathbf{v}_i = \vec{\omega} \times \mathbf{r}_i$

$(1) \Rightarrow \quad T = \frac{1}{2}\sum_i m_i \mathbf{v}_i \cdot \mathbf{v}_i$

$\qquad = \frac{1}{2}\sum_i m_i \mathbf{v}_i \cdot \left(\vec{\omega} \times \mathbf{r}_i\right)$

$\qquad = \frac{1}{2}\sum_i m_i \, \vec{\omega} \cdot (\mathbf{r}_i \times \mathbf{v}_i)$

$\Rightarrow \quad T = \frac{1}{2}\vec{\omega} \cdot \left[\sum_i m_i (\mathbf{r}_i \times \mathbf{v}_i)\right] = \frac{1}{2}\vec{\omega} \cdot \mathbf{L}, \qquad \ldots(2)$

In terms of moment of inertia tensor $\mathbf{I}$ and $\vec{\omega}$, it can be rewritten as

$$L = \sum_i m_i \left[\vec{\omega} r_i^2 - \left(\mathbf{r}_i \cdot \vec{\omega} \right) \mathbf{r}_i \right]$$

$$= \left\{ \sum_i m_i \left[r_i^2 - \mathbf{r}_i \left(\mathbf{r}_i \cdot \right) \right] \right\} \vec{\omega} = \mathbf{I} \vec{\omega} \quad \ldots(3)$$

from (3) & (1) $\quad T = \dfrac{1}{2} \vec{\omega} \cdot \mathbf{I} \cdot \vec{\omega} \quad \ldots(4)$

Let $\mathbf{n}$ be a unit vector in the direction of ω so that $\omega = \omega \mathbf{n}$ and the kinetic energy expression is,

$$T = \frac{\omega^2}{2} \mathbf{n} \cdot \mathbf{I} \cdot \mathbf{n} = \frac{1}{2} I \omega^2 \quad \ldots(5)$$

Such that the moment of I about the axis of rotation.

$$I = \mathbf{n} \cdot \mathbf{I} \cdot \mathbf{n} \quad \ldots(6)$$

But (2) $\Rightarrow \quad T = \dfrac{1}{2} \vec{\omega} \cdot \mathbf{L}$

And break up into components of $\vec{\omega}$ and $\mathbf{L}$:

$$T = \frac{1}{2} \left[\mathbf{i}\omega_x + \mathbf{j}\omega_y + \mathbf{k}\omega_z \right] \cdot \left[\mathbf{i}L_x + \mathbf{j}L_y + \mathbf{k}L_z \right]$$

$$= \frac{1}{2} \left[\omega_x L_x + \omega_y L_y + \omega_z L_z \right]$$

$$= \frac{1}{2} [\omega_x \left(I_{xx}\omega_x + I_{xy}\omega_y + I_{xz}\omega_z \right)$$

$$+ \omega_y \left(I_{yx}\omega_x + I_{yy}\omega_y + I_{yz}\omega_z \right)$$

$$+ \omega_z \left(I_{zx}\omega_x + I_{zy}\omega_y + I_{zz}\omega_z \right)]$$

$$= \frac{1}{2} [I_{xx}\omega_x^2 + I_{yy}\omega_y^2 + I_{zz}\omega_z^2$$

$$+ 2I_{xy}\omega_x\omega_y + 2I_{yz}\omega_y\omega_z + 2I_{zx}\omega_x\omega_z] \quad \ldots(7)$$

$\mathbf{L}$ has the simple form

$$\mathbf{L} = \mathbf{i}I_1\omega_x + \mathbf{j}I_2\omega_y + \mathbf{k}I_3\omega_z, \quad \ldots(8)$$

and therefore kinetic energy expression also assumes the simple form :

$$T = \frac{1}{2} \vec{\omega} \cdot \mathbf{L}$$

$$= \frac{1}{2} \left(\mathbf{i}\omega_x + \mathbf{j}\omega_y + \mathbf{k}\omega_z \right) \cdot \left(\mathbf{i}I_1\omega_x + \mathbf{j}I_2\omega_y + \mathbf{k}I_3\omega_z \right)$$

$$= \frac{1}{2} I_1\omega_x^2 + \frac{1}{2} I_2\omega_y^2 + \frac{1}{2} I_3\omega_z^2 \quad \ldots(9)$$

which is an expression involving the principal moments of inertia.

4.5 MOMENT OF INERTIA OF A RIGID BODY :

We shall show that $\mathbf{n} \cdot \mathbf{I} \cdot \mathbf{n}$ is the moment of inertia of the rigid body about the axis of rotation,

$$\mathbf{n} \cdot \mathbf{I} \cdot \mathbf{n} = \mathbf{n} \cdot \left\{ \sum_i m_i \left[r_i^2 - \mathbf{r}_i (\mathbf{r}_i \cdot) \right] \right\} \mathbf{n}$$

by putting up the operator $\mathbf{I}$. Simplifying, we get

$$\mathbf{n}.\mathbf{I}.\mathbf{n} = \mathbf{n}.\left\{\sum_i m_i \left[r_i^2 \mathbf{n} - \mathbf{r}_i (\mathbf{r}_i.\mathbf{n})\right]\right\}$$

$$= \sum_i m_i \left[r_i^2 (\mathbf{n}.\mathbf{n}) - (\mathbf{r}_i.\mathbf{n})(\mathbf{r}_i.\mathbf{n})\right]$$

$$= \sum_i m_i \left[(\mathbf{r}_i.\mathbf{r}_i) - (\mathbf{r}_i.\mathbf{n})(\mathbf{r}_i.\mathbf{n})\right]$$

since $\mathbf{n}.\mathbf{n} = 1$, $\mathbf{r}_i.\mathbf{r}_i = r_i^2$. Taking out common $\mathbf{r}_i$ outside we obtain

$$r.h.s. = \sum_i m_i \mathbf{r}_i.\left[\mathbf{r}_i - \mathbf{n}(\mathbf{r}_i.\mathbf{n})\right]$$

$$= \sum_i m_i \mathbf{r}_i.\left[\mathbf{n} \times (\mathbf{r}_i \times \mathbf{n})\right]$$

as $\mathbf{n} \times (\mathbf{r}_i \times \mathbf{n}) = (\mathbf{n}.\mathbf{n})\mathbf{r}_i - (\mathbf{r}_i.\mathbf{n})\mathbf{n}$. Now, by virtue of the property of scalar triple product,

$$r.h.s. = \sum_i m_i (\mathbf{r}_i \times \mathbf{n}).(\mathbf{r}_i \times \mathbf{n})$$

$$= \sum_i m_i \left| \mathbf{r}_i \times \mathbf{n} \right|^2 = I$$

4.6 EULER'S EQUATIONS OF MOTION OF A RIGID BODY :

The motion of a rigid body with one point fixed will take place under the action of a torque N in such a way that its total angular momentum varies at the rate equal to N :

$$\frac{d\mathbf{L}}{dt} = \mathbf{N} \qquad ...(1)$$

In a coordinate system rotating with body, we have the following relation between the two time derivatives

$$(d/dt)_{space} = (d/dt)_{body} + \vec{\omega} \times$$

Eq. (1) in terms of body-axes is therefore,

$$\left(\frac{d\mathbf{L}}{dt}\right)_{body} + \vec{\omega} \times \mathbf{L} = \mathbf{N} \qquad ...(2)$$

It is found most convenient to choose the principal axes for which.

$$\mathbf{L} = I_1\omega_x \mathbf{i} + I_2\omega_y \mathbf{j} + I_3\omega_z \mathbf{k} \qquad ...(3)$$

where ω_x, ω_y and ω_z are the components of the angular velocity vector along the principal axes.

Form eq. (3) and remembering that the principal moments of inertia and the body base vectors, $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are constant in the time with respect to the body co-ordinate system, we find that the time derivative of L, that is $d\mathbf{L}/dt$, in rotating system is

$$\left(\frac{d\mathbf{L}}{dt}\right)_{body} = I_1 \dot{\omega}_x \mathbf{i} + I_2\dot{\omega}_y \mathbf{j} + I_2 \dot{\omega}_z \mathbf{k}. \quad ...(4)$$

The x-component (2) is obtained as

$$N_x = \mathbf{N}.\mathbf{i} = I_1\dot{\omega}_x - \left(\vec{\omega} \times \mathbf{L}\right)_x$$

$$= I_1 \dot{\omega}_x - \left[\omega_y L_z - \omega_z L_y\right]$$

O

$$= I_1 \dot{\omega}_x - (I_2 - I_3)\omega_y \, \omega_z \quad ...(5)$$

from eq. (3). By cyclic permutation of x, y, z we obtain the remaining two equations, viz.,

$$N_y = I_2 \dot{\omega}_y - (I_3 - I_1)\omega_z \omega_x \quad ...(6)$$

$$N_z = I_3 \dot{\omega}_z - (I_1 - I_2)\omega_x \omega_y \quad ...(7)$$

These three equations are the Euler's equation for the motion of a rigid body with one point fixed.

4.7 TORQUE FREE MOTON OF A RIGID BODY :

When N = 0, Euler's eqs. Become

$$I_1 \dot{\omega}_x = \omega_y \omega_z \, (I_2 - I_3)$$

$$I_2 \dot{\omega}_y = \omega_z \omega_x \, (I_3 - I_1) \quad ...(1)$$

$$I_3 \dot{\omega}_z = \omega_x \omega_y \, (I_1 - I_2)$$

It is possible to integrate equations in terms of elliptic functions with initial conditions of kinetic energy and total angular momentum, which are the integral or constant of the motion; for, there are no net forces acting on the system. The system is conservative and Euler's dynamical equations furnish following two constants of motion :

4.7.1 Conservation of Kinetic Energy :

Multiplying eq. (1) by $\omega_x, \omega_y, \omega_z$ respectively and adding gives

$$I_1 \omega_x \dot{\omega}_x + I_2 \omega_y \dot{\omega}_y + I_3 \omega_z \dot{\omega}_z = \omega_x \omega_y \omega_z (I_2 - I_3 + I_3 - I_1 + I_1 - I_2) = 0$$

But $\quad \dfrac{d}{dt}\left(\dfrac{1}{2} I_1 \omega_x^2\right) = I_1 \omega_x \dot{\omega}_x \quad$ etc.

And therefore the left hand side is

$$\frac{d}{dt}\left(\frac{1}{2} I_1 \omega_x^2 + \frac{1}{2} I_2 \omega_y^2 + \frac{1}{2} I_3 \omega_z^2\right) = 0 .$$

The gives the conservation of kinetic energy, i.e.

$$T = \frac{1}{2} I_1 \omega_x^2 + \frac{1}{2} I_2 \omega_y^2 + \frac{1}{2} I_3 \omega_z^2 = \text{constant} \quad ...(2)$$

4.7.2 Conservation of angular momentum :

Another constant of motion is the angular momentum L, since we have presumed N = 0, whence

$$N = \frac{d\mathbf{L}}{dt} = 0 \text{ and therefore L is conserved:}$$

$$\mathbf{L} = \mathbf{i} I_1 \omega_x + \mathbf{j} I_2 \omega_y + \mathbf{k} I_3 \omega_z = \text{constant.} \quad ...(3)$$

We shall describe a very interesting solution of the problem due to Poinsot. This is a geometrical representation of motion of rigid body based on the integrals of motion, eqs. (2) and (3) developing into the motion of inertia-ellipsoid.

4.7.3 Heavy symmetric top:

The symmetric top is approximated to variety of rigid bodies, like the child's top and gyroscope. It pivots around a fixed point O on the symmetry axis and we take the distance of O from G, the center of mass, to be l.

Taking body z'-axis, the symmetry axis, as the principle axis, we put $I_1 = I_2$. Only force acting on the top is the force of gravity mg acting at G downward (Fig.3).
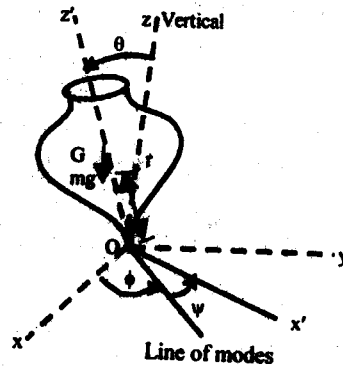

Fig 4.3:

The Euler's angles are the most convenient set of generalized co-ordinates to describe the motion in this case. We find the Lagrangian L for top.

$$L = T - V = \frac{1}{2} I_1 \left(\omega_{x'}^2 + \omega_{y'}^2 \right) + \frac{1}{2} I_3 \omega_{z'}^2 - mgl \cos\theta,$$

which, on substituting for angular velocities in terms of Euler's angles,

$$L = \frac{1}{2} I_1 \left(\dot{\theta}^2 + \dot{\phi}^2 + \sin^2\theta \right) + \frac{I_3}{2} \left(\dot{\psi} + \dot{\phi}\cos^2\theta \right) - mgl\cos\theta \quad \text{----(1)}$$

As we know that, apart from translational motion, the top has following three types of motion :

Precession : Which is rotation about space z-axis. Such a rotation corresponds to angle ϕ.

Nutation : Which is rotation about intermediate x_1 axis or line of nodes. Such a rotation corresponds to angle θ.

Spin : Which is rotation about z' axis. Such a rotation corresponds to angle φ. The spin velocity about z'-axis is ω_z' given by $\left(\dot{\phi}\cos\theta + \dot{\psi} \right)$.

4.7.4 First integrals of equations of motion :

Since the Lagrangian does not contain the Euler's angles φ and time t, i.e., they are cyclic, the corresponding momenta p_φ, p_ϕ and total energy E are constant in time.

These three integrals of motion are expressed by

$$p_\varphi = \frac{\partial L}{\partial \dot{\psi}} = I_3 \left(\dot{\psi} + \dot{\phi}\cos\theta \right) = I_3 \omega_3' = I_1 a. \qquad ...(2)$$

where a is a constant, and

$$p_\phi = \frac{\partial L}{\partial \dot{\phi}} = I_1 \dot{\phi}\sin^2\theta + I_3 \cos\theta \left(\dot{\psi} + \dot{\phi}\cos\theta \right) = I_1 b \quad ...(3)$$

where b is a constant, and

$$T + V = \frac{I_2}{2}\left(\dot{\theta}^2 + \dot{\phi}^2 \sin^2\theta\right) + \frac{I_3}{2}\omega_{z'}^2 + mgl\cos\theta = E \quad ...(4)$$

Equations (2,3) may be solved for ψ and ϕ, yielding $\phi = \dfrac{b - a\cos\theta}{\sin^2\theta}$ ------(5)

and $\dot{\psi} = \dfrac{I_1 a}{I_3} - \left(\dfrac{b - a\cos\theta}{\sin^2\theta}\right)\cos\theta.$ ------(6)

Substituting (5) and (6) in the energy equation, we obtain

$$E' = E - \frac{(I_1 a)^2}{2I_3} = \frac{1}{2}I_1\dot{\theta}^2 + \frac{1}{2}I_1\dot{\phi}^2 \sin^2\theta + mgl\cos\theta$$

$$= \frac{1}{2}I_1\dot{\theta}^2 + \frac{1}{2}I_1\left(\frac{b - a\cos\theta}{\sin\theta}\right)^2 + mgl\cos\theta \quad ...(7)$$

This equation, when solved for θ as a function of t, involves the use of elliptic integrals which tend to introduce complexity. Therefore some other ways must be found out which describe the phenomenon qualitatively. One such way is that of effective potential in equation (7) which, like the potential $V(\theta)$, is a function of θ alone :

$$V'(\theta) = mgl\cos\theta + \frac{1}{2}I_1\left(\frac{b - a\cos\theta}{\sin\theta}\right) \quad ......(8)$$

This is a one-dimensional problem in θ co-ordinate. In (fig.4a) we plot the effective potential as a function of θ. $V'(\theta)$ assumes infinite values for $\theta = 0$ or π, V' assumes a minimum value. In figure, θ_o at which V' is minimum must therefore be a solution of the equation:

$$\frac{dV'}{d\theta} = mgl\sin\theta + I_1 a\left(\frac{b - a\cos\theta}{\sin\theta}\right) - I_1\frac{(b - a\cos\theta)^2}{\sin\theta}\cos\theta$$

$$= -mgl\sin\theta + \frac{I_1(b - a\cos\theta)(a - b\cos\theta)}{\sin^3\theta} = 0 \quad ...(9)$$

For a particular value of energy $E' = E - \dfrac{(I_1 a)^2}{2I_3}$, θ motion is a bound motion, confined between the two values θ_1 and θ_2 of θ, which are the roots of the equation

$$E' = V'(\theta)$$

$$\Rightarrow E' = mgl\cos\theta + \frac{1}{2}I_1\left(\frac{b - a\cos\theta}{\sin\theta}\right)^2 \quad ...(10)$$

The variation in the angle θ is referred to the nutation of the symmetry axis of the top and is an up and down motion of the symmetry axis. If, however, the minimum effective potential equals the energy E', the angle θ keeps fixed at value θ_0 and the top precesses with the constant angular velocity given by

$$\dot{\phi} = \frac{b - a\cos\theta_0}{\sin^2\theta_0} \quad ....(11) \text{ about the vertical axis.}$$

4.7.5 Precession without Nutation

Consider the angular velocity of precession and spin when nutation is absent i.e., z' – axis remains fixed at θ_0. From eq. (9), it is obvious that, for a given θ_0, there will be two angular frequencies of precession $\dot{\phi}_0$ given by

$$\dot{\phi}_0 = \frac{\dot{\psi} I_3 \pm \sqrt{\left[\dot{\psi}^2 I_3^2 - 4mgl(I_1 - I_3)\cos\theta_0\right]}}{2(I_1 - I_3)\cos\theta_0} \quad ...(12)$$

Thus corresponding to positive sign of the radical, precession will be fast while that with negative sign is called a slow precession. Since $\dot{\phi}_0$ must be real, the quantity under the radical sign must be positive definite, i.e.,

$$I_3^2 \dot{\psi}^2 - 4mgl(I_1 - I_3)\cos\theta_0 \geq 0$$

$$\dot{\psi}^2 \geq \frac{4mgl(I_1 - I_3)\cos\theta_0}{I_3^2} \quad ...(13)$$

(a)   **(b)**
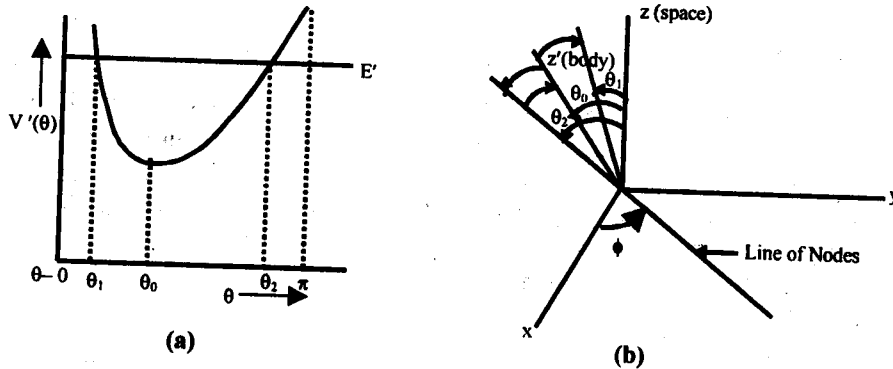
Fig 4.4: Illustrating the effect of two rotations performed in a given order

Thus this equation limits the values of $\dot{\psi}$ for which steady precession may occur at the angle θ_0. Further, we can also solve eq. (12) for $\dot{\phi}_0$ in terms of ω_z', the angular velocity about z'-axis called spin angular velocity and is given by

$$\dot{\phi}_0 = \frac{I_3\omega_z' \pm \sqrt{\left(I_3^2\omega_z'^2 - 4mgl\,I_1\cos\theta_0\right)}}{2I_1\cos\theta_0} \quad (13A)$$

Putting for $\dot{\phi}_0$ from eq. (5)

$$\therefore \quad (b - a\cos\theta_0) = \frac{1}{2}\frac{I_3\omega_z'}{I_1} - \frac{\sin^2\theta_0}{\cos\theta_0}\left[1 \pm \left(1 - \frac{4mgl I_1\cos\theta_0}{I_3^2\,\omega_z'^2}\right)^{1/2}\right]$$

$$= \frac{1}{2}\frac{I_3\sin^2\theta_0}{I_1\cos\theta_0}\left[\omega_z' \pm \left(\omega_z'^2 - \frac{4mgl\,I_1\cos\theta_0}{I_3^2}\right)^{1/2}\right] \quad ...(14)$$

We note that for $\theta_0 < \pi/2$, physical motion is possible and as stated previously, for uniform precession, quantity under radical must be positive, i.e.,

$$I_3^2\,\omega_z'^2 - 4mgl\,I_1\cos\theta \geq 0$$

$$\Rightarrow \quad (\omega_z')_{min} = \sqrt{\left(\frac{4mgl\,I_1\cos\theta}{I_3^2}\right)} \quad ...(15)$$

This is the minimum spin angular velocity ω'_z below which the top cannot precess uniformly at the angle θ_0 and is given by equation (15). When $\omega'_z > (\omega'_z)_{min}$ quantity under radical will not be zero and hence two values of $\dot{\phi}_0$ are given by

$$\dot{\phi}_0 = \frac{(b - a\cos\theta_0)_+}{\sin^2\theta_0} \approx \frac{I_3}{I_1}\frac{\omega'_z}{\cos\theta_0}$$

and

$$\dot{\phi}_0 = \frac{(b - a\cos\theta_0)_-}{\sin^2\theta_0} \approx \frac{mgl}{I_3\omega'_z} \qquad ...(16)$$

It is the slow precession which is ordinarily observed with a rapidly spinning top (large ω'_z).

4.7.6 Nutational Motion :

Let us now investigate the nutational motion that we have mentioned earlier. At the angle θ_0, the effective potential $V'(\theta)$ has a minimum value; it corresponds to a point of stable equilibrium, θ_1 and θ_2 being the turning angles corresponding to the effective energy E'. It follows, therefore, that situation exists for which the nutational motion of the axes of the top can be described as a frequency of these oscillations and given by

$$\omega_\theta = \sqrt{\left[\frac{1}{I_1}\left(\frac{d^2V'(\theta)}{d\theta^2}\right)\right]_{\theta=\theta_0}} \qquad ...(17)$$

We find that

$$\frac{d^2V'(\theta)}{d\theta^2}\bigg|_{\theta=\theta_0} = I_1\left[a^2 - 4a\dot{\phi}_0\cos\theta_0 + 3\dot{\phi}_0^2\cos^2\theta_0 + \dot{\phi}_0^2\right] \qquad ...(18)$$

The angular frequency of oscillation about θ_0 obviously depends on spin velocity ω'_z through a and on $\dot{\phi}_0$. This motion is delineated by the processional angular velocity $\dot{\phi}_0$ which is dictated by spinning velocity ω'_z. It is customary to depict the motion of the top by tracing the curve of the intersection of the figure axis (symmetry axis) on a sphere of unit radius. This curve is then the locus of the figure axis. The polar co-ordinates of a point on the locus are identical with the Euler angles θ, ϕ for the body system. The bounding angles θ_1 and θ_2 on the unit sphere satisfy the equation (10), when $E' = V'(\theta)$, in which p_ϕ, p_ψ and E are determined from the initial conditions. If we multiply this equation by $\sin^2\theta$, it becomes a cubic equation in $\cos\theta$. Referring to fig 4(a). there are two real roots $\cos\theta_1, \cos\theta_2$ between -1 and $+1$. The third root can, in fact, be shown to be greater than 1, on a consideration of eq. (10) and is not of physical interest. Initially, when the top is spinning fast, nutation motion is absent and $\cos\theta_1 = \cos\theta_0$. If initially $\dot{\theta} = 0$, the initial value $\cos\theta_1$ and θ satisfies equation (10). During nutation, the precessional velocity varies in accordance with eq. (5). If

$$|b| < |a| \quad \text{or} \quad |p_\phi| < |p_\psi|,$$

we can define angle θ_3 as follows : $\cos\theta_3 = \dfrac{p_\phi}{p_\psi} = \dfrac{b}{a}$ ...(19)

In view of equation (5), if $\theta > \theta_3, \dot{\phi}$ has the same sign as ω'_z and for $\theta < \theta_3$, it has an opposite sign. Also

$$\frac{dV'}{d\theta}\bigg|_{\theta=\theta_0}$$

is negative, i.e. the slope of the potential energy curve in fig.4(a) is negative, or, the angle between tangent to the same sing as ω'_z throughout the nutation, and the figure axis trace out a curve shown in fig. on the unit sphere. If $\theta_3 > \theta_1, \dot{\phi}$ changes sign during nutation. The corresponding motion is shown in fig.
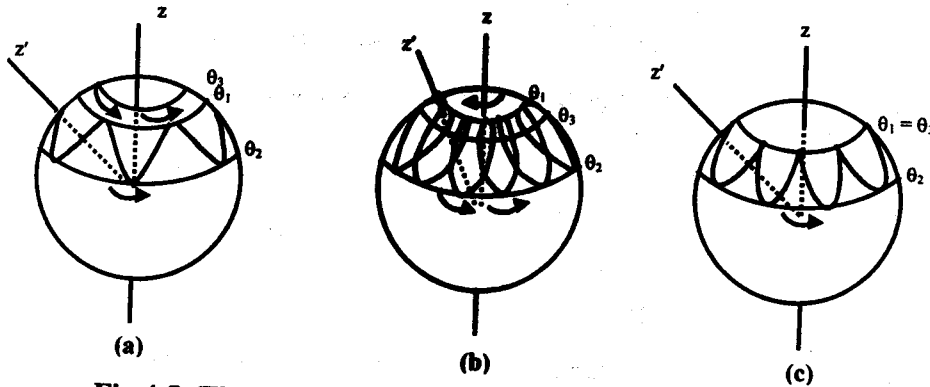
Fig 4.5: The possible shapes for the locus of the figure axis on the unit sphere

4.7.7 Fast top :

The fast top is a top spinning very rapidly about the figure axis with angular velocity ω'_z and is initially at rest at an angle θ_1 and then released. For this, we have the initial conditions :

i) initial angle, $\theta = \theta_1$, ii) initially at rest $\dot\theta = 0, \dot\phi = 0$, iii) spinning with ω'_z, $\dot\psi = \dot\psi_0 = \omega'_z$

So the equations (5, 6, 7) become.,

$$\left. \begin{array}{l} a = \dfrac{I_3 \dot\psi_0}{I_1} \\[3mm] b = \dfrac{I_3 \dot\psi_0 \cos\theta_1}{I_1} \\[3mm] E' = mgl\cos\theta_1 \end{array} \right\} \qquad \text{...(20)}$$

In motion, $\dot\theta$ begin to differ from their initial zero values so that first two terms of eq. (7) will now have some value. Consequently, to keep E constant, the term $mgl\cos\theta$ should decrease. In other words, energy can be conserved only by a decrease in potential energy that is, by an increase in θ. The initial θ_3 is, therefore, the same as θ_1, the minimum value θ can have (from eq. (20) $\frac{b}{a} = \cos\theta_1$). When released in this manner, the top always starts to fall, and continues to fall until the other bounding angle θ_2 is reached, precessing meanwhile. The figure axis then begins to rise again to θ_1, the complete motion being shown in fig 5(c).

If now we assume $\frac{1}{2} I_1 a^2 \geq mgl$, i.e., the rotational kinetic energy about the symmetry axis is much larger than the maximum allowable change in potential energy, then we find that ϕ is small and to a first approximation, we can set

$$\left. \frac{d^2 V(\theta)}{d\theta^2} \right|_{\theta=\theta_0} = I_1 a^2$$

$$\Rightarrow \quad \omega_\theta = a$$

$$= \frac{I_3 \dot\psi_0}{I_1} \qquad \qquad ...(21)$$

Let δ be the amplitude of this sinusoidal (nutational) motion in the unit of θ-angle and take place about the mean value θ_0, obviously,

$$\theta_0 = \theta_1 + \delta$$

and therefore, at any time angle θ is a function of time, varying sinusoidally about the value θ_0 :

$$\theta = \theta_0 - \delta \cos\omega t = \theta_1 + \delta - \delta \cos\omega t$$

so that $\dot\phi = \frac{a}{\sin\theta_1} \cdot (\delta - \delta \cos\omega t)$

obtained on using eq. (5)

The average of ϕ is given by

$$<\dot\phi> = \frac{\int_0^{2\pi} \frac{a}{\sin\theta_1}(\delta - \delta\cos\omega t)d(\omega t)}{\int_0^{2\pi} d(\omega t)}$$

$$= \frac{\frac{a}{\sin\theta_1}.\delta .2\pi - \left[\frac{a\delta}{\sin\theta_1}\sin\omega t\right]_0^{2\pi}}{2\pi}$$

Putting $a = \frac{I_3 \dot\psi_0}{I_1}$, we get

$$<\dot\phi> = \frac{I_3 \dot\psi_0}{I_1 \sin\theta_1}.\delta \qquad \qquad ...(22)$$

But $\delta^* = \frac{mgl \sin\theta_1}{I_1 a^2} \qquad \qquad ...(23)$

Therefore eq. (23), becomes

$$<\dot\phi> = \frac{I_3 \dot\psi_0}{I_1 \sin\theta_1} \cdot \frac{mgl \sin\theta_1}{I_1 a^2}$$

$$= \frac{I_3 \dot\psi_0}{I_1^2} \cdot mgl \cdot \frac{I_1}{I_3^2 \dot\psi^2}$$

(Putting for a)

$$= \frac{mgl}{I_3 \dot\psi_0} \cdot \qquad \text{----------(24)}$$

Thus we conclude about **Fast Top** that :

i) since $I_1 a^2 >> mgl$, δ, the amplitude of nutation, is small

ii) since $\theta = \theta_0 + \delta - \delta\cos\omega t$,   nutation is sinusoidal (simple harmonic).

iii) since $I_3\dot{\psi}_0 = I_1 a \geq \dfrac{mgl}{a}$

$<\dot{\phi}>$ is small as inferred from eq. (24). Thus the precession is slow.

iv) since $\omega_\theta = a$, frequency of nutation is large.

Thus the fast top released from rest, precesses slowly and notates simple harmonically with a large frequency and small amplitude.

Summary:

A rigid body can possess simultaneously the translational and rotational motion, and the equations of motion, governing the two types of motion, may involve both the translational and rotational co-ordinates. If the rigid body is taken as a rigid collection of particles, then angular momentum is given by

$$L = \sum_i m\left(r_i \times v_i\right)$$

In terms of the components of the vectorial quantities involved we can write the components of angular momentum as

$$L_x = \left[\omega_x \sum_i m_i\left(r_i^2 - x_i^2\right) - \omega_y \sum_i m_i x_i y_i - \omega_z \sum_i m_i x_i z_i\right]\qquad = I_{xx}\omega_x + I_{xy}\omega_y + I_{xz}\omega_z,$$

$$L_y = \left[-\omega_x \sum_i m_i x_i y_i + \omega_y \sum_i m_i\left(r_i^2 - y_i^2\right) - \omega_z \sum_i m_i y_i z_i\right]\qquad = I_{yx}\omega_x + I_{yy}\omega_y + I_{yz}\omega_z,$$

$$L_z = \left[-\omega_x \sum_i m_i x_i z_i - \omega_y \sum_i m_i y_i z_i + \omega_z \sum_i m_i\left(r_i^2 - z_i^2\right)\right]\qquad = I_{zx}\omega_x + I_{zy}\omega_y + I_{zz}\omega_z.$$

I is therefore a different physical entity, termed as the moment of inertia tensor. The components of I have been arranged in matrix from in a way to facilitate understanding of linear transformation : otherwise I is basically a tensor.

The diagonal element of matrix form,

$$I_{xx} = \sum_i m_i\left(r_i^2 - x_i^2\right) = \sum_i m_i\left(y_i^2 + z_i^2\right)$$

$$I_{yy} = \sum_i m_i\left(r_i^2 - y_i^2\right) = \sum_i m_i\left(z_i^2 + x_i^2\right)$$

$$I_{zz} = \sum_i m_i\left(r_i^2 - z_i^2\right) = \sum_i m_i\left(x_i^2 + y_i^2\right)$$

are called the moments of inertia coefficients. These are the moments of inertia of the body about the $x, y, z -$ axes separately.

$$I_{xy} = I_{yx} = -\sum_i m_i x_i y_i$$

$$I_{yz} = I_{zy} = -\sum_i m_i y_i z_i$$

$$I_{zx} = I_{xz} = -\sum_i m_i z_i x_i$$

are termed as the products of inertia associated with the corresponding co-ordinate planes. Obvious I is a symmetric tensor and is additive in nature; the moments of inertia of a body are the sums of those of its parts. A mathematical structure having nine components in three dimensions is termed a tensor of rank two. In this sense, I is a tensor of rank two and we shall write such an entity in type I to emphasize its tensorial character.

The kinetic energy of a rotating rigid body can be written as

$$T = \frac{1}{2}\vec{\omega}.\mathbf{L}$$

and the torque

$$\mathbf{N} = \left(\frac{d\mathbf{L}}{dt}\right)_{body} + \vec{\omega}\times\mathbf{L}$$

And the components of torque

$$N_x = I_1\dot{\omega}_x - (I_2 - I_3)\omega_y\,\omega_z$$

$$N_y = I_2\dot{\omega}_y - (I_3 - I_1)\omega_z\omega_x$$

$$N_z = I_3\dot{\omega}_z - (I_1 - I_2)\omega_x\omega_y$$

These three equations are the Euler's equation for the motion of a rigid body with one point fixed.
For torque free motion, kinetic energy

$$T = \frac{1}{2}I_1\omega_x^2 + \frac{1}{2}I_2\omega_y^2 + \frac{1}{2}I_3\omega_z^2 = \text{constant}$$

To represent the motion of a spinning top, the Euler's angles are the most convenient set of generalized co-ordinates
The Lagrangian L for a top is

$$L = T - V = \frac{1}{2}I_1\left(\omega_{x'}^2 + \omega_{y'}^2\right) + \frac{1}{2}I_3\omega_{z'}^2 - mgl\cos\theta,$$

which, on substituting for angular velocities in terms of Euler's angles,

$$L = \frac{1}{2}I_1\left(\dot{\theta}^2 + \dot{\phi}^2 + \sin^2\theta\right) + \frac{I_3}{2}\left(\dot{\phi} + \dot{\phi}\cos^2\theta\right) - mgl\cos\theta$$

Since the Lagrangian does not contain the Euler's angles φ and time t, i.e., they are cyclic, the corresponding momenta p_φ, p_ϕ and total energy E are constant in time.

These three integrals of motion are expressed by

$$p_\varphi = \frac{\partial L}{\partial\varphi} = I_3\left(\dot{\psi} + \dot{\phi}\cos\theta\right) = I_3\omega_3' = I_1a.$$

where a is a constant, and

$$p_\phi = \frac{\partial L}{\partial\phi} = I_1\dot{\phi}\sin^2\theta + I_3\cos\theta\left(\dot{\psi} + \dot{\phi}\cos\theta\right) = I_1b$$

where b is a constant, and

$$T + V = \frac{I_2}{2}\left(\dot{\theta}^2 + \dot{\phi}^2 \sin^2 \theta \right) + \frac{I_3}{2}\omega_z^2 + mgl\cos\theta = E$$

From the analysis of the motion of spinning top we learn that ,there is a minimum spin angular velocity ω_z' below which the top cannot precess uniformly. The fast top released from rest, precesses slowly and nutates simple harmonically with a large frequency and small amplitude.

Key words : Angular velocity, angular momentum ,moments and products of inertia, Rotational kinetic energy , moment of inertia of a rigid body , Euler's equations of motion, Torque free motion of a rigid body, Heavy symmetric top, First integrals of equation of motion, Nutational motion , Fast top.

Self-assessment questions
1 Define Euler angles and derive Euler's equations of motion in terms of Euler's angles.
2 What do you understand by the inertia tensor of a rigid body?
3. Deduce Euler's equations of motion of a rigid body with one point fixed. What is inertia tensor? Discuss its properties.
4. Explain the motion of heavy symmetric top.

Reference Books:
1. Classical Mechanics: Goldstein
2. Mechanics: Simon
3 Classical Mechanics: Gupta, Kumar and Sharma

Classical Mechanics Part
Unit II
Lesson 5

Special theory of Relativity

Objectives

To learn
1. About inertial and non-inertial frames of reference
2. Why Galelian transformations are not satisfactory
3. How Lorentz transformation satisfactorily explains the invariance of Physical laws
4. Consequences of Lorentz transformation
5. Principles of relativistic mechanics

Structure:

5.1 Basic postulates of special theory of relativity
5.2 Galilean transformation
5.3 Postulates of special relativity
5.4 Lorentz Transformation
5.5 The relativistic law of addition of velocities
5.6.Kinematic effects of Lorentz transformation
5.6.1 Lorentz – Fitzerald Contraction
5.6.2 Time delation
5.7 Covarient four-dimensional formulations
5.8 Force and energy relations in relativistic mechanics
5.9.Lagrangian formulation of relativistic Mechanics
5.10 Covarient Lagrangian formulation.

Introduction: Laws of classical mechanics hold good when the velocities of the particles of the system are small compared with velocity of light. When the velocities involved approach the velocity of light the true description consistent with the experimental facts is provided by special theory of relativity.

5.1 Basic postulates of special theory of relativity

A space system obeying Newton's law of motion $F = ma$ is called an inertial system. It appears that a system moving uniformly with respect to a 'space system' should itself be an inertial system. To show it, consider two co-ordinate systems. One inertial system S and the other S' moving uniformly with respect to S with velocity $\bar{v}$, A given point P is located by the radius vector r and r' with respect to the two systems which are related as

$$r' = r - vt \qquad ...(1)$$

Since the relative velocity is constant, the first time derivative of equation (1) is

$$\dot{r}' = \dot{r} - v \qquad ...(2)$$

and another differentiation gives $\ddot{r}' = \ddot{r}$ i.e.

$$a^1 = a \qquad\qquad ...(3)$$

So that the acceleration is same in both systems if Newton's law holds in one system it should hold in the other, if the force has same form in both systems. In many of the usual problems of mechanics the force is obviously unchanged between the two systems, as for example: with a constant force field such as $F = mg$.

The transformation represented by eq. (1)and eq (3)is called Galilean transformation, predicts that the velocity of light should be different in the two systems.

5.2 Galilean transformation

Now, suppose there is a source of light at the origin of the unprimed system emitting spherical waves traveling with the speed C. Let the radius vector r be the position vector of a point on some given wave surface.
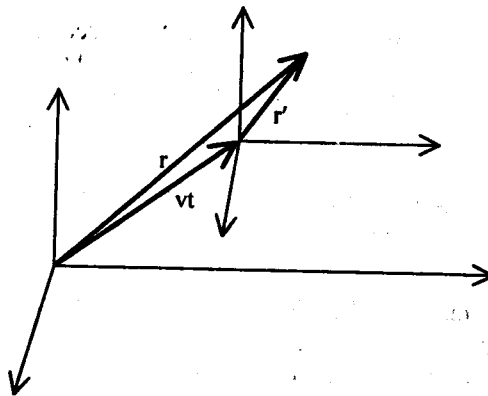

Fig 5.1:Illustrating the Galilean transformation

Then in the unprimed system the velocity of the point on the wave surface is $\dot{r} = Cn$, where n is a unit vector along r. According to equation (1), however, the corresponding wave velocity in the primed system is $\dot{r}' = Cn - v$. In the system moving with respect to the source of light the magnitude of the wave velocity will in general no longer be C; indeed, since it depends on direction, the waves will no longer be spherical.

New transformation needed:

A long series of investigations, especially the famous experiments of Michelson and Morley, have indicated that the velocity of light is always the same in all directions and is independent of the relative uniform motions of the observer, the transmitting medium and the source. Since the propagation of light in vacuum with the speed C is a consequence of Maxwell's equations, it must be concluded that the Galilean transformation does not preserve the form of Maxwell's equations. Now, it is a postulate of physics, that all phenomena of physics should appear the same in all systems moving uniformly relatively to each other. Measurements made entirely within a given system must be incapable of distinguishing that system from all others moving uniformly with respect to it. This postulate of equivalence requires that physical laws must be phrased in an identical manner for all uniformly moving systems, i.e., be covariant when subjected to a Galilean transformation. Experimentally both Newton's laws and Maxwell's equations seemed to satisfy the equivalence postulate but that theoretically, i.e., according to Galilean transformation, Maxwell's equations did not. Einstein concluded that it is the form of Maxwell's equations that must be kept invariant and therefore the Galilean transformation could not be correct. A new relationship between uniformly moving systems, the Lorentz transformation, must be found that preserves the speed of light in all uniformly moving

systems. Einstein showed that such a transformation requires revision of the usual concepts of time and simultaneity.

The program of the theory of special relativity is therefore two fold. First there must be obtained a transformation between two uniformly moving systems that will preserve the velocity of light. Second, the laws of physics must be examined as to their transformation properties under this Lorentz transformation. Those laws that do not keep their form invariant are to be generalized as to obey the equivalence postulate.

5.3 Postulates of special relativity

From above discussion, we arrive at the following two postulates of special relativity.

1. The principle of equivalence of physical laws:

This principle states that all laws of physics take the same identical form for all frames of reference in uniform relative motion, i.e., for all the inertial frames of reference. This is a direct consequence of the absence of an absolute or fixed frame of reference. For, if the laws of physics were to take on different forms in different frames of reference, it would be easily inferred from these differences as to which of them are at rest in space and which in motion. But, as we have seen, such distinction between the state of rest and of uniform motion is excluded by the absence of a universal frame of reference.

2. Principle of constancy of velocity of light:

This principle states that the velocity of light in free space is the same (C) relative to any inertial frame of reference, i.e., it is invariant to transformation from one inertial frame to another and has thus the same value $(3 \times 10^8 \, m/sec)$ for all observes irrespective of their state of motion.

This postulate is clearly a statement of the result of Michelson- Morley experiment.

The two implications of the above postulate are immediately obvious.

i) Velocity being not invariant to Galilean transformation , it follows that if C be the velocity of light in frame S, that in frame S' moving relative to S with velocity v, must be $C' = C - v$. In accordance with the second postulate of the special theory of relativity, however we have $C' = C$, since C must always have the same value irrespective of the state of motion of the frames of reference.
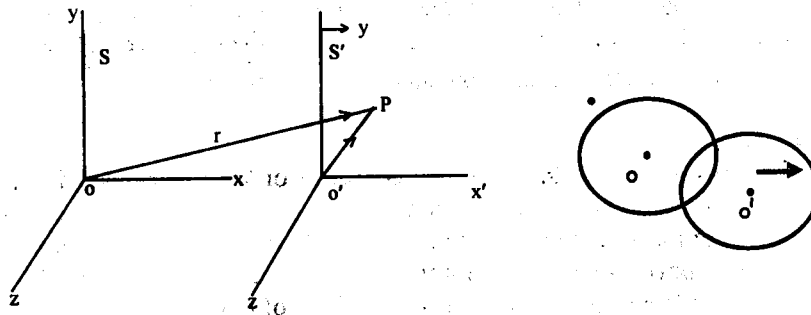
Fig 5.2 : Illustration of principle of constancy of velocity of light

ii) Again, suppose a flash of light is emitted from the origin when points O and O' of the two frames of reference S and S' just cross each other [see fig. (2)]. Suppose further that it is possible for the two observers in the two frames of reference to see the light or rather the wavefront spreading out into space. Then, in view of the constancy of the value of C, irrespective of the motion of the two frames of reference, each of the two observers at O and O' respectively will, even after drifting apart, claim to be at the center of

the spreading spherical wavefront of light as shown in fig. (2), thinking that the other has moved away from the center of the sphere.

5.4 Lorentz Transformation

Einstein first derived these transformation equations with the following objectives in mind. (However these transformation equations are called Lorentz transformation equations as Lorentz first derived these equations in his theory of electromagnetism.)

i) transformation must be linear   ii) It should preserve the speed of light, and

iii) It should approach Galilean transformation in the limit of low speeds compared to speed of light.

We consider two uniformly moving systems whose origins coincide at time $t = 0$. Further at $t = 0$, let a source of light, fixed at the origin of the unprimed system, emit a pulse of light. The pulse of light spreads out as a growing sphere. Thus an observer fixed in the unprimed system will see a spreading spherical wavefront propagating with speed C, the equation of which can be written as

$$x^2 + y^2 + z^2 = C^2 t^2 \qquad ...(1)$$

But since speed of light being invariant, observer in the primed system moving with respect to the source will also see the pulse of light propagating as the spherical wavefront from his our origin – equation being

$$x'^2 + y'^2 + z'^2 = C^2 t'^2 \qquad ...(2)$$

which shows that time is no more scalar invariant but is a function of a particular co-ordinate system in uniform motion relative to other systems.

Inspection of equations (1) and (2) reveals that the desired transformation must be such that

$$x^2 + y^2 + z^2 - C^2 t^2 = x'^2 + y'^2 + z'^2 - C^2 t'^2$$

$$...(3)$$

Writing x_1, x_2 and x_3 for x, y and z, the equation becomes

$$\sum_{i=1}^{4} x_i^2 - C^2 t^2 = \sum_{i=1}^{4} x_i'^2 - C^2 t'^2$$

If we put $x_4 = iCt$, a fourth imaginary co-ordinate, then

$$\sum_{\mu=1}^{4} x_\mu^2 = \sum_{\mu=1}^{4} x_\mu'^2 \qquad ...(4)$$

This means that the square of the radius vector is invariant under a transformation of co-ordinates in the four-dimensional x_1, x_2, x_3 and x_4 space.

We know that in three-dimensional case, if we rotate co-ordinate axes, then the components of a vector in new system are given by $x_\mu' = \sum_{v=1}^{3} a_{\mu v} x_v$ and the length of the vector remains unchanged, i.e., $\sum_{\mu=1}^{3} x_\mu'^2 = \sum_{\mu=1}^{3} x_\mu^2$

Thus basing on the analogy of above equation which represents spatial rotation of co-ordinate axes in three dimensional orthogonal space, we can say that equation (4) represents orthogonal transformation of a vector in four dimensional space or in other words Lorentz transformations are merely the orthogonal transformation of four dimensional space – the latter also being recognized as world space or Minkowski's space. Lorentz transformation can, therefore, be represented as

$$x_\mu' = \sum_{v=1}^{4} a_{\mu v} x_v ,$$

where $a_{\mu v}$ is the linear transformation matrix. We are here concerned only with a transformation that involves uniformly moving systems whose axes are parallel and is called pure Lorentz transformation. Such a

transformation does not involve any spatial rotation. Let us consider that primed system is moving with velocity v along x_3 axis. Now

$$x'_\mu = \sum_{v=1}^{4} a_{\mu v} x_v \qquad ...(5)$$

$$x_1^1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4$$

or
$$x_2^1 = a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 \qquad ...(6)$$

$$x_3^1 = a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{34}x_4$$

$$x_4^1 = a_{41}x_1 + a_{42}x_2 + a_{43}x_3 + a_{44}x_4$$

Obviously x_1^1, x_2^1 and x_3^1 being space co-ordinates will be real for which a_{14}, a_{24}, a_{34} should be imaginary (because $x_4 = iCt$ is imaginary). Similarly, for x_4^1 to be imaginary it is essential that a_{41}, a_{42} and a_{43} elements be imaginary while a_{44} real.

As the primed system is taken to be moving along x_3 axis, then x_1 and x_2, being in a direction perpendicular to that of system motion, will remain uneffected by the transformation. i.e.,

$$x_1^1 = x_1,$$

$$x_2^1 = x_2$$

As the components x_3 and x_4 undergo transformation and hence x_1 and x_2 should not appear in the transformation matrix of x_3^1 and x_4^1. Thus equation (6) become

$$x_1^1 = x_1$$

$$x_2^1 = x_2$$

$$x_3^1 = a_{33}x_3 + a_{34}x_4 \qquad ...(7)$$

$$x_4^1 = a_{43}x_3 + a_{44}x_4$$

Therefore the matrix of transformation is

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & a_{43} & a_{44} \end{bmatrix} \qquad ...(8)$$

Further these matrix elements have to obey some sort of orthogonality conditions as for spatial rotation in three-dimensional space, i.e.,

$$\sum_{v} a_{\mu v} a_{\lambda v} = \delta_{\mu \lambda} \qquad ...(9)$$

$$= 1 \text{ if } \mu = \lambda$$

$$= 0 \text{ if } \mu \neq x$$

Putting $\mu = \lambda = 3$ and first $v = 3$ and $v' = 4$, we get

$$a_{33}^2 + a_{34}^2 = 1 \qquad ...(10)$$

Also, with $\mu = \lambda = 4$ and $v = 3$ and 4,

$$a_{43}^2 + a_{44}^2 = 1 \qquad ...(11)$$

further $a_{33}a_{43} + a_{34}a_{44} = 0 \qquad ...(12)$

Thus orthogonality condition furnishes the three equations (10), (11) and (12) connecting the four matrix elements. The four unknown elements can be determined uniquely only when fourth relation between them is

provided. We know that the origin of primed system $\left(x_3^1 = 0\right)$ is moving uniformly along x_3 - axis, thus after time t its x_3 co-ordinate will be vt, i.e.,

$$x_3 = vt = v.\frac{x_4}{iC} = -i\frac{v}{C}x_4 = -i\beta x_4$$

which, with matrix relation for x_3^1, gives

$$x_3^1 = a_{33}x_3 + a_{34}x_4$$
$$= \left(a_{34} - i\beta a_{33}\right)x_4 = 0$$

or $a_{34} = i\beta a_{33}$,

which when substituted in equation (10), gives

$$a_{33} = \frac{1}{\sqrt{1-\beta^2}}$$

so that $a_{34} = \frac{i\beta}{\sqrt{1-\beta^2}}$

On solving equations (11) and (12), we can show that

$$a_{44} = \frac{1}{\sqrt{1-\beta^2}} \text{ and } a_{43} = \frac{-i\beta}{\sqrt{1-\beta^2}}$$

We can now write the Lorentz transformation as

$$\begin{bmatrix} x_1^1 \\ x_2^1 \\ x_3^1 \\ x_4^1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \dfrac{1}{\sqrt{1-\beta^2}} & \dfrac{i\beta}{\sqrt{1-\beta^2}} \\ 0 & 0 & \dfrac{-i\beta}{\sqrt{1-\beta^2}} & \dfrac{1}{\sqrt{1-\beta^2}} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

or

$$\begin{bmatrix} x^1 \\ y^1 \\ z^1 \\ iCt^1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \dfrac{1}{\sqrt{1-\beta^2}} & \dfrac{i\beta}{\sqrt{1-\beta^2}} \\ 0 & 0 & \dfrac{-i\beta}{\sqrt{1-\beta^2}} & \dfrac{1}{\sqrt{1-\beta^2}} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ iCt \end{bmatrix}$$

$$x^1 = x$$
$$y^1 = y$$

giving $z^1 = \dfrac{z-vt}{\sqrt{1-\beta^2}}$ ...(13)

$$iCt^1 = z.\left(\frac{-i\beta}{\sqrt{1-\beta^2}}\right) + \frac{iCt}{\sqrt{1-\beta^2}}$$

$$t^1 = \frac{t - \left(v/C^2\right)z}{\sqrt{1-\beta^2}}$$

Eq. (13) is known as Lorentz transformation equations. It is obvious that Lorentz transformation to real co-ordinate system is possible when $\beta < 1$ indicating that one cannot have a relative velocity greater than C, the speed of light.

5.5 The relativistic law of addition of velocities

Let us consider three frames of reference S, S^1 and S^{11}; S^1 moving with velocity V_1 relative to S and S^{11} with velocity V_2 relative to S^1. Now here we find out the relative velocity of S^{11} with respect to S. The transformation equations containing S^1 with S, and S^{11} with S^1 are respectively.

$$
\left.
\begin{aligned}
x^1 &= x \\
y^1 &= y \\
z^1 &= \frac{z - v_1 t}{\sqrt{1 - v_1^2/C^2}} \\
t^1 &= \frac{t - \left(v_1/C^2\right)z}{\sqrt{1 - v_1^2/C^2}}
\end{aligned}
\right\} \quad \ldots(1)
\qquad
\left.
\begin{aligned}
x^{11} &= x^1 \\
y^{11} &= y^1 \\
z^{11} &= \frac{z^1 - v_2 t}{\sqrt{1 - v_2^2/C^2}} \\
t^{11} &= \frac{t^1 - \left(v_2/C^2\right)z^1}{\sqrt{1 - v_2^2/C^2}}
\end{aligned}
\right\} \quad \ldots(2)
$$

Substituting eq. (1) and (2), we get

$$x^{11} = x$$

$$y^{11} = y$$

$$
z^{11} = \frac{\dfrac{z - v_1 t}{\sqrt{1 - v_1^2/C^2}} - v_2\left(\dfrac{t - \left(v_1/C^2\right)z}{\sqrt{1 - v_1^2/C^2}}\right)}{\sqrt{1 - v_2^2/C^2}}
$$

$$
= \frac{z\left(\dfrac{1 + v_1 v_2/C^2}{\sqrt{1 - v_1^2/C^2}}\right) - \dfrac{t(v_1 + v_2)}{\sqrt{1 - v_1^2/C^2}}}{\sqrt{1 - v_2^2/C^2}}
$$

$$
= \frac{\left(z - \dfrac{t(v_1 + v_2)}{\left(1 + v_1 v_2/C^2\right)}\right)\left(1 + \dfrac{v_1 v_2}{C^2}\right)}{\sqrt{1 - v_1^2/C^2}\ \sqrt{1 - v_2^2/C^2}}
$$

$$
= \frac{\left[z - t\left(\dfrac{v_1 + v_2}{1 + v_1 v_2/C^2}\right)\right]}{\dfrac{1}{\left(1 + \dfrac{v_1 v_2}{C}\right)}\left[\sqrt{\left(1 + \dfrac{v_1 v_2}{C^2}\right)^2 - \left(\dfrac{v_1 + v_2}{C}\right)^2}\right]}
\qquad = \frac{z - vt}{\sqrt{1 - \dfrac{v_2}{C^2}}}
$$

Where $v = \dfrac{v_1 + v_2}{\left(1 - \dfrac{v_1 v_2}{C^2}\right)}$

And similarly $t^{11} = \dfrac{t-\left(v/C^2\right)z}{\sqrt{1-v^2/C^2}}$

From these equations for z^{11} and t^{11} we infer that two Lorentz transformations carried out in succession are equivalent to one Lorentz transformation. we also note that the relative velocity v of S^1 with respect to S is not simply the sum of v_1 and v_2 but

$$v = \frac{v_1 + v_2}{1 + \dfrac{v_1 v_2}{C^2}} \qquad ...(3)$$

Further, the expression for v can be put in the form

$$v = C\left[1 - \frac{\left(1 - \dfrac{v_1}{C}\right)\left(1 - \dfrac{v_2}{C}\right)}{1 + \dfrac{v_1 v_2}{C^2}}\right],$$

Which shows that v cannot become equal to or greater than C, the speed of light as long as the factors $\dfrac{v_1}{C}$ and $\dfrac{v_2}{C}$ are small enough. From eq. (3) of $v_1 = C$ and $v_2 = C$, then

$$v = \frac{C+C}{1 + \dfrac{C^2}{C^2}} = C$$

that is, when velocity of light is add to the velocity of light, we obtain velocity of light.

5.6 Kinematic effects of Lorentz transformation

The two consequences of Lorentz transformation, that so called Lorentz-Fitzerald contraction of length and dilation of time scales, are explained as follows.

5.6.1 Lorentz – Fitzerald Contraction

Let us consider that a rigid rod at rest is lying along z-axis in the unprimed system. Its length in its own system is

$$l = z_2 - z_1$$

An observer connected with the primed system that is moving uniformly with respect to unprimed system measures the length of the rod by locating simultaneously the position of both end points as z_1^1 and z_2^1 in his system at the same time t^1. Thus from Lorentz equations, we have

$$z_1 = \frac{z_1^1 + vt^1}{\sqrt{1-\beta^2}} \qquad z_2 = \frac{z_2^1 + vt^1}{\sqrt{1-\beta^2}}$$

So the apparent length is

$$(z_1 - z_2)\sqrt{1-\beta^2} = \left(z_1^1 - z_2^1\right)$$

or $\quad l\sqrt{1-\beta^2} = l^1$

Thus, to the moving observer, rod appears to be contracted.

5.6.2 Time dilation

Suppose in unprimed system, at point z, a clock is located. At time t_2 on this clock an observer fixed in the moving primed system notes a time

$$t_1^1 = \frac{t_1 - vz/C^2}{\sqrt{1 - \beta^2}}$$

and at time t_2 be notes

$$t_2^1 = \frac{t_2 - vz/C^2}{\sqrt{1 - \beta^2}}$$

so that apparent time interval is

$$t_2^1 - t_1^1 = \frac{t_2 - t_1}{\sqrt{1 - \beta^2}}$$

since $\left(t_2^1 - t_1^1\right) > \left(t_2 - t_1\right)$, it appears to the moving observer that the stationary clock is moving at a slow rate. The effect is called "time dilation".

5.7 Covariant four – dimensional formulations

The task of verifying the laws of physics for invariance in form under Lorentz transformation is easily done by writing them interms of the four-dimensional world. Invariance of form under Lorentz transformation is not the only invariant property demanded of physical laws. Clearly the physical content of any given relation cannot be affected by the particular orientation chosen for the spatial axes, the laws of physics must also be invariant in form under rigid rotations, i.e., proper spatial orthogonal transformations.

Normally we do not worry about the invariance of our theories under spatial rotations. In constructing any equation it is always required that the terms of the equation be all scalars, or all vectors, in general all terms must be tensors of the same rank, and this requirement-automatically ensures the desired invariance under rotation. Thus a scalar relation will have the general form.

$$a = b,$$

and since both sides of the equation, being scalars, are invariant under spatial rotations the relation obviously holds in all coordinate systems. A vector relation of the form $\mathbf{F} = \mathbf{G}$, really stands for three separate relations between the components of the vector $F_i = G_i$

The values of these components, of course, are not invariant under the spatial rotation, rather they are transformed to new values F_i^1, G_i^1 that are the components of the transformed vectors, $\mathbf{F}^1, \mathbf{G}^1$. But because both sides of the component relations transform in identical fashion, the same relation must hold between the transformed components :

$$F_i^1 = G_i^1$$

The relationship between the two vectors is thus undisturbed by the spatial rotation; in the new system we still have $\mathbf{F}^1 = \mathbf{G}^1$

Similarly, an equality between two tensors of the second rank is $\mathbf{C} = \mathbf{D}$

Necessarily implies the same equality between the two transformed tensors $\mathbf{C}^1 = \mathbf{D}^1$, because the two tensors transform covariantly under a spatial rotation. On the other hand, an equation involving separately a component of a vector and say a component of a tensor obviously can not remain invariant in form under a three dimensional orthogonal transformation. Invariance of a physical law under rotation of the spatial coordinate system requires covariance of the terms of the equation under three-dimensional orthogonal transformation.

Now, the restricted Lorentz transformation has been identified as on orthogonal transformation in Minkowski, or world space we have already studied about scalars and vectors in this four dimensional space. Similarly one can set up tensors of other ranks in this space, with transformation properties that are obvious generalizations of the three dimensional transformations. These tensors of various ranks will be called as *world tensor*. The invariance of the form of any physical law under Lorentz transformation will then be

immediately evident once it is expressed in a covariant four-dimensional form, all terms being world tensors of the same rank. A law failing to meet the requirements of the equivalence principle cannot be put into a covariant form.

The simplest example of a four-vector is the position vector of a "point" in Minkowski space, with components (x_1, x_2, x_3, x_4). Since the four coordinates of a world point tell where in ordinary space something happened and when it happened, it is perhaps more descriptive to speak of a point in four-dimensional space as an *event*. As a particle move in ordinary space its corresponding point in four-dimensional space will describe a path known as *world line*. The four-vector dx_μ represents the change in the position four-vector for a differential motion along the world line. From the dot product of dx_μ with itself we can form a world scalar, denoted by $d\tau$ and defined by the equation

$$(d\tau)^2 = \frac{1}{C^2} dx_\mu dx_\mu \qquad ...(1)$$

The significance of $d\tau$ can be made clear by evaluating eq. (1) in the Lorentz system in which the particle is momentarily at rest. In such a system the components of the transformed vector dx'_μ are $(0,0,0, iCdt^1)$ and the invariant $d\tau$ is given by

$$(d\tau)^2 = -\frac{1}{C^2} dx'_\mu dx'_\mu = (dt^1)^2$$

Thus $d\tau$ is the time interval as measured on a clock traveling with the particle, hence it is referred to as an interval of the particle's *proper time* or *world time*.

The relation between $d\tau$ and an interval of time as measured in a given Lorentz system can be derived directly by expanding the eq. (1).

$$(d\tau)^2 = -\frac{1}{C^2}\left[(dx)^2 + (dy)^2 + (dz)^2 - C^2(dt)^2\right],$$

or $\qquad d\tau = dt \sqrt{1 - \frac{1}{C^2}\left[\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2\right]},$

which is equivalent to the relation

$$\frac{d\tau}{\sqrt{1-\beta^2}} = dt \qquad ...(2)$$

Here β being used to represent the velocity of a particle as observed in one inertial system. Eq. (2) says that a time interval measured in the rest system is always longer than the corresponding time interval observed in a system in which the particle is not at rest. This is an example of the familiar "time-dilation".

The square of the magnitude of the four-vector is not necessarily positive definite. Four-vectors for which the square of the magnitude is greater than or equal to zero are called *space-like*; when the squares of the magnitude are negative they are known as *time-like* vectors. As an example of these concepts it may be noted that the difference vector between two world points can be either space-like or time-like. Let X_μ be the difference vector, defined as

$$\overline{X}_\mu = x_{1\mu} - x_{2\mu},$$

the subscripts 1 and 2 denoting the two events. The magnitude of X_μ is given by

$$\overline{X}_\mu \overline{X}_\mu = |\mathbf{r}_1 - \mathbf{r}_2|^2 - C^2(t_1 - t_2)^2.$$

Thus X_μ is space-like if the two world points are separated such that

$$|\mathbf{r}_1 - \mathbf{r}_2|^2 \geq C^2(t_1 - t_2)^2,$$

while it is time-like if

$$|\mathbf{r}_1 - \mathbf{r}_2|^2 < C^2(t_1 - t_2)^2.$$

The condition for a time-like difference vector is equivalent to stating that it is possible to bridge the distance between the two events by a light signal, while if the points are separated by a space-like difference vector they cannot be connected by any wave traveling with the speed C.

The four-velocity u_v is defined as the rate of change of the position vector of a particle with respect to proper time :

$$u_v = \frac{dx_v}{d\tau} \qquad \qquad ...(3)$$

with space and time components

$$u_i = \frac{dx_i}{d\tau} = \frac{dx_i}{dt\sqrt{1-\beta^2}} = \frac{v_i}{\sqrt{1-\beta^2}}$$

$$\Rightarrow u_i = \frac{v_i}{\sqrt{1-\beta^2}} \text{ and } u_4 = \frac{iC}{\sqrt{1-\text{ }^{?}}} \quad ...(4)$$

$$(\because dx_4 = iC\iota.$$

The world velocity has a constant magnitude, for the sum $u_v u_v$ is give by

$$u_v u_v = \frac{v^2}{1-\beta^2} - \frac{C^2}{1-\beta^2} = -C^2 \quad ...(5)$$

and it is thus also time-like

Now consider the example of vector and scalar electromagnetic potentials, which together form a four-vector $A_\mu \rightarrow (\mathbf{A}, i\varphi)$. If the potentials satisfy the Lorentz gauge condition

$$\nabla \cdot \mathbf{A} + \frac{1}{C}\frac{\partial \phi}{\partial t} = 0 \qquad \qquad ...(6)$$

Then they separately satisfy wave equations of the form

$$\nabla^2 \mathbf{A} - \frac{1}{C^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} = -\frac{4\pi}{C}\mathbf{j}$$

$$\nabla^2 \phi - \frac{1}{C^2}\frac{\partial^2 \phi}{\partial t^2} = -4\pi\rho \qquad ...(7)$$

In the language of Minkowski space the Lorentz condition can be written in covariant form as

$$\frac{\partial A_\mu}{\partial x_\mu} = 0 \qquad \qquad ...(8)$$

In obvious generalization of the three-dimensional del operator ∇, the four dimensional gradient-operator may be denoted by the symbol $\square$. $\square^2$ (known as the D'Alembertion) is therefore a world scalar differential operator :

$$\square^2 = \frac{\partial^2}{\partial x_\mu \partial x_\mu} = \nabla^2 - \frac{1}{C^2}\frac{\partial^2}{\partial t^2}$$

Hence the set of wave equations eq. (6), can be written as a clearly covariant world vector equation.

$$\square^2 A_\mu = -\frac{4\pi j_\mu}{C} \qquad \qquad ...(9)$$

5. Force and energy equations in relativistic mechanics

Newton's equations of motion, being invariant under a Galilean transformation, cannot be invariant under Lorentz transformation, they must be suitably generalized to provide a law of force satisfying the covariance requirements of special relativity. For velocities small compared to C we have

$$\frac{d}{dt}(mv_i) = F_i \qquad \text{...(1)}$$

Eq. (1) is not Lorentz-invariant, therefore, its relativistic generalization will be a four-vector equation, whose spatial component reduces to (1) in the limit as $\beta \to 0$. The only four-vector whose space part reduces to v for small velocities is the world velocity u_v. Further, while m can be taken as an invariant property of the particle, we know that time t is not a Lorentz-invariant but it can obviously replaced by the scalar proper time τ, which approaches t as $\beta \to 0$. The generalized Newton's equations of motion for a single particle must therefore have the form

$$\frac{d}{d\tau}(mu_v) = \bar{k}_v \qquad \text{...(2)}$$

Where $\bar{k}_v$ is some four-vector, known as the Minkowski force. K_i reduces to F_i in the limit of small velocities. Thus K_i may be equal to the product of F_i with any function of β that reduces to unity as $\beta \to 0$; the exact relation clearly depends upon the Lorentz transformation properties of Force components. Two types of approach have been used to determine the behaviour of F under Lorentz transformation.

One procedure begins by pointing out that, fundamentally, forces arise from only a few physical sources-forces are either gravitational, electromagnetic, or possibly nuclear. The transformation properties must be the same for all forces no matter what their origin.

The electromagnetic force on a particle is given by

$$F_i = q \cdot \left(\frac{\partial}{\partial x_i}\left(\phi - \frac{1}{C}v.A\right) + \frac{1}{C}\frac{dA_i}{dt} \right) \quad \text{...(3)}$$

In terms of the four-vector potential $\bar{A}_\mu$, the expression $\phi - \frac{1}{C}V.A$ can be written covariantly as .

$$\phi - \frac{1}{C}V.A. = -\frac{1}{C}\sqrt{1-\beta^2}\, \bar{u}_v \bar{A}_v \qquad \text{...(4)}$$

and the force components F_i become

$$F_i = -\frac{q}{C}\sqrt{1-\beta^2}\left(-\frac{\partial}{\partial x_i}(u_v A_v) + \frac{d\bar{A}_i}{d\tau} \right) \quad \text{...(5)}$$

The expression in the parentheses transforms as the spatial component of a four-vector, so that F_i is equal to the produce $\sqrt{1-\beta^2}$ and the spatial component of a four-vector, which is to be identified as the Minkowski force $\bar{k}_\mu$. Hence the connection between the ordinary and Minkowski forces must be

$$F_i = K_i \sqrt{1-\beta^2} \qquad \text{...(6)}$$

Irrespective of the origin of the forces A by-product of this derivation is the particular form of the Minkowski force on charged particles :

$$K_\mu = \frac{q}{C}\left(\frac{\partial}{\partial x_\mu}(u_v A_v) - \frac{dA_\mu}{d\tau} \right). \qquad \text{...(7)}$$

The alterative procedure attempts to avoid the necessity of using a physical theory beyond mechanics itself, it simply defines force as being the time rate of change of momentum, in all Lorentz systems.

$$\frac{dp_i}{dt} = F_i \qquad \ldots(8)$$

The momentum indicated in (8), however, is not mv_i, but rather some relativistic generalization that reduces to it in the limit of small velocities.

Now from the relation between τ and t, and the definition of world velocity, we can write the spatial components of eq. (2) as

$$\frac{d}{dt\sqrt{1-\beta^2}}\left(\frac{mv_i}{\sqrt{1-\beta^2}}\right) = K_i$$

$$\Rightarrow \frac{d}{dt}\left(\frac{mv_i}{\sqrt{1-\beta^2}}\right) = K_i\sqrt{1-\beta^2}$$

Comparison with (8) shows that the conservation of momentum theorem is invariant providing the momentum is defined by

$$P_i = \frac{mv_i}{\sqrt{1-\beta^2}} \qquad \ldots(9)$$

Which reduces to mv_i as $\beta \rightarrow 0$. Comparison of (9) with $u_i = \dfrac{v_i}{\sqrt{1-\beta^2}}$ and $u_4 = \dfrac{iC}{\sqrt{1-\beta^2}}$ shows that P_i

forms the spatial part of a four-vector that is called the four-momentum :

$$P_\nu = mu_\nu \qquad \ldots(10)$$

The generalized equation of motion for a single particle thus can also be written as

$$\frac{dP_\nu}{d\tau} = \bar{k}_\nu \qquad \ldots(11)$$

So far only the space part of the four-vector eq. (11) has been discussed; nothing has been said about the physical significance of fourth equation. The time-like part of the four-vector k_μ can be obtained directly from the dot product of (11) with the world velocity :

$$u_\nu \frac{d}{d\tau}(mu_\nu) = \frac{d}{d\tau}\left(\frac{m}{2}u_\nu u_\nu\right) = K_\nu u_\nu \quad \ldots(12)$$

Since the square of the magnitude u is a constant, $-C^2$, and m is here likewise a constant, the left hand side of (12) vanishes, leaving

$$K_\nu u_\nu = \frac{\mathbf{F}.\mathbf{v}}{1-\beta^2} + \frac{iCK_4}{\sqrt{1-\beta^2}} = 0 \qquad \ldots(13)$$

$$\left(\because k_\nu u_\nu = k_i \frac{v_i}{\sqrt{1-\beta^2}} + k_4 \frac{iC}{\sqrt{1-\beta^2}} \right.$$

$$\left. = \frac{F_i V_i}{\sqrt{1-\beta^2}} + \frac{iCk_4}{\sqrt{1-\beta^2}} = \frac{\mathbf{F}.\bar{v}}{\sqrt{1-\beta^2}} + \frac{iCk_4}{\sqrt{1-\beta^2}} \right)$$

The fourth – component of the Minkowski force is therefore

$$K_4 = \frac{i}{C}\frac{\mathbf{F}}{\sqrt{1-\beta^2}} \qquad \ldots(14)$$

and the corresponding fourth component of equation (2) appears as

$$\frac{d}{dt}\frac{mC^2}{\sqrt{1-\beta^2}} = \mathbf{F}.v \qquad ...(15)$$

Now, the kinetic energy T is defined in general to be such that $\mathbf{F}.V$, the rate at which the force does work on the particle, is the time rate of increase of T

$$\frac{dT}{dt} = \mathbf{F}.v \qquad ...(16)$$

$$\left(\because \mathbf{F}.v = m\frac{d^2r}{dt^2}.\frac{dr}{dt} = \frac{d}{dt}\left(\frac{1}{2}m\dot{r}^2\right) = \frac{dT}{dt} \right)$$

This is a definition of kinetic energy that agrees with the form $\frac{1}{2}mv^2$.

Thus form (15) & (16) $T = \dfrac{mC^2}{\sqrt{1-\beta^2}} \qquad ...(17)$

In the limit as β^2 becomes much less than eq. (17) can be expanded as

$$T \to mC^2\left(1+\frac{\beta^2}{2}\right) = mC^2 + \frac{mv^2}{2} \quad ...(18)$$

Comparison of eq. (10) with eq. (17) shows that $\dfrac{iT}{C}$ is the fourth component of the four-momentum. In order to preserve the correct transition to non-relativistic values, it is preferable to transfer the name "Kinetic Energy" to a separate quantity, K_1 defined by

$$K = T - mC^2 = mC^2(\gamma - 1)$$

T will often called as the energy. Any situation that conserves the spatial linear momentum must also conserve T. To verify this theorem we need only note that the statement "spatial linear momentum is conserved" must be invariant under a Lorentz transformation.

To express kinetic energy in terms of four momentum, we write the Lorentz invariant of four momentum, i.e., its magnitude

$$P_\mu P_\mu = \left(\frac{mv_i}{\sqrt{1-\beta^2}}+\frac{iT}{C}\right)\left(\frac{mv_i}{\sqrt{1-\beta^2}}+\frac{iT}{C}\right)$$

$$= \frac{m^2v^2}{\left(1-\beta^2\right)}-\frac{T^2}{C^2} \qquad = P^2 - \frac{T^2}{C^2}$$

Also $P_\mu P_\mu = m^2 u_\mu^2 = -m^2 C^2$

Therefore, $-m^2 C^2 = P^2 - \dfrac{T^2}{C^2}$

$$T^2 = P^2 C^2 + m^2 C^4 \qquad ...(20)$$

Which is the relativistic analogue of the relation $T = \dfrac{P^2}{2m}$ in nonrelativistic mechanics except that T here includes the rest energy.

5.9 Lagrangian formulation of relativistic Mechanics

Now we want to set up a Lagrangian that will give the correct relativistic equations of motion. The concept of Lagrangian was introduced in two ways, namely D'Alembert's principle and Hamilton's principle. Since the D'Alembert's principle involves the momentum P_i which is now different from non-relativistic momentum $m_i v_i$, approach to this new Lagrangian formulation is not possible by this way. Hamilton's principle is still helpful in finding a function for L for which the Euler-Lagrange equations, obtained from this variational principle, agree with known relativistic equations of motion. For example, we take the case of a single particle acted on by conservative forces independent of velocity and suppose a suitable relativistic Lagrangian for such a case is found to be

$$L = -mC^2 \sqrt{1-\beta^2} - V, \qquad ...(1)$$

Here V purely depends upon position, then introduction of the Lagrangian in Euler-Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v_i}\right) - \frac{\partial T}{\partial x_i} = 0 \qquad ...(2)$$

should yield a relativistic equation. This in fact is true since from equation (1) we have

$$\frac{\partial L}{\partial v_i} = -mC^2 \times \frac{-\frac{1}{2}.2\frac{v_i}{C^2}}{\sqrt{1-\beta^2}}$$

$$= \frac{mv_i}{\sqrt{1-\beta^2}}$$

Which when substituted in eq. (2) gives

$$\frac{d}{dt}\left(\frac{mv_i}{\sqrt{1-\beta^2}}\right) = -\frac{\partial V}{\partial x_i}$$

$$= F_i$$

$$\frac{d}{d\tau}\left(\frac{mv_i}{\sqrt{1-\beta^2}}\right) = \frac{F_i}{\sqrt{1-\beta^2}} = K_i$$

and is the relativistic equation of motion.
We have thus shown that the choice of L represented by eq. (1) is correct. We note that L is no longer $(T-V)$, but

$$\frac{\partial L}{\partial v_i} = \frac{mv_i}{\sqrt{1-\beta^2}} = p_i,$$

i.e., partial derivative of L with velocity is still the momentum. Taking any desired set of generalized co-ordinates q_i, canonical momenta can still be defined as

$$p_i = \frac{\partial L}{\partial q_i}$$

Which shows that if a coordinate is cyclic, conjugate momentum will be conserved. Since the definition of canonical momentum and the general form of Lagrange's equations are the same as in non-relativistic case, the function

$$H = \sum_i \dot{q}_i p_i - L$$

Would still represent a constant of motion provided L does not contain time explicitly – a condition also imposed in non-relativistic case. As L is no longer $(T-V)$ nor is $\sum \dot{r_i} p_i$ equal to $2T$, we are to prove in another way that H is also the total energy. In the case of a single particle, H is given by

$$H = \sum_i \frac{mv_i^2}{\sqrt{1-\beta^2}} + mc^2 \sqrt{1-\beta^2} + V$$

which, on solving reduces to

$$H = \frac{mC^2}{\sqrt{1-\beta^2}} + V = T + V = E$$

For velocity dependent potential

We consider Lagrangian for a single particle in an electro-magnetic field. The force acting on a particle with charge q is given by

$$F = q\left(E + \frac{1}{C}(v \times B) \right)$$

which can be put in the form

$$F_i = \left[\frac{d}{dt}\left(\frac{\partial}{\partial v_i} \right) - \frac{\partial}{\partial x_i} \right]\left(q\varphi - \frac{q}{C} v.A \right)$$

putting $p_i = \left(\dfrac{mv_i}{\sqrt{1-\beta^2}} \right)$

for relativistic case, we write

$$\frac{d}{dt}\left(\frac{mv_i}{\sqrt{1-\beta^2}} \right) = \left[\frac{d}{dt}\left(\frac{\partial}{\partial v_i} \right) - \frac{\partial}{\partial x_i} \right]\left(q\phi - \frac{q}{C} v.A \right) \quad ...(3)$$

Which can be derived from Hamilton's principle provided L is given by

$$L = - mC^2 \sqrt{1-\beta^2} - q\phi + \frac{q}{C} v.A \quad ...(4)$$

The canonical momentum will have additional terms due to the dependence of potential on velocity along with mu_i, i.e.,

$$p_i = \frac{\partial L}{\partial v_i} = \frac{\partial}{\partial v_i}\left[1 - mC^2 \sqrt{1-\beta^2} - q\phi + \frac{q}{C} V.A \right]$$

$$= \frac{mv_i}{\sqrt{1-\beta^2}} + \frac{q}{C} A_i \quad = mu_i + \frac{q}{C} A_i$$

This additional term is also present in the non-relativistic case.

5.10 Covariant Lagrangian Formulations

In the lagrangian procedure no effect has been made to keep the ideal of a covariant four-dimensional form for all laws of mechanics. Thus the time t has been treated as a parameter entirely distinct from the spatial coordinates, while a covariant formulation would require that space and time be considered as entirely similar coordinates in worldspace. Now some invariant parameter should be used, instead of t, to trace the progress of the system point in configuration space. If the parameter of integration is a Lorentz invariant, then the Lagrangian function itself must be a worldscalar in any covariant formulation. Finally, instead of being a

function of x_i and $\dot{x}_i$, the Lagrangian should be a function of the coordinates in Miskowski space and of their derivatives with respect to the invariant parameter.

We consider a system of only one particle. The natural choice of the invariant parameter in such a system would seem to be the particle's proper time τ. But the various components of the generalized velocity u_v, must then obey the relation

$$-u_v u_v = C^2 \qquad \ldots(1)$$

Which shows they are not independent. Therefore we shall instead assume the choice of some Lorentz invariant quantity θ with no further specification than that it be a monotone function of the progress of the world point along the particle's world line. Now we represent the differentiation with respect to θ by

$$x_v^1 = \frac{dx_v}{d\theta}$$

A suitably covariant Hamilton's principle must therefore appear as

$$\delta I = \delta \int_{\theta_1}^{\theta_2} \wedge\left(x_\mu, x_\mu^1\right) d\theta, \qquad \ldots(2)$$

Where the Lagrangian function $\wedge$ must be a world scalar.
The Euler-Lagrange equations corresponding to eq. (2) are

$$\frac{d}{d\theta}\left(\frac{\partial \wedge}{\partial x_\mu^1}\right) - \frac{\partial \wedge}{\partial x_\mu} = 0 \qquad \ldots(3)$$

The problem is to find the form of $\wedge$ such that eq. (3) is equivalent to the equations of motion, $\frac{d}{d\tau}\left(m u_v\right) = k_v$

Now $\quad \dfrac{dx_i}{dt} = \dfrac{dx_i}{d\theta} \cdot \dfrac{d\theta}{dt} = iC\dfrac{x_i^1}{x_4^1} \qquad \ldots(4)$

$$I = \int_{t_1}^{t_2} L\left(x_j, t_1 \dot{x}_j\right) dt = \frac{-i}{C}\int_{\theta_1}^{\theta_2} L\left(x_\mu, iC\frac{x_j^1}{x_4^1}\right) x_4^1 \, d\theta$$

Now a suitable $\wedge$ is given by the relation $\quad \wedge\left(x_\mu, x_\mu^1\right) = -\dfrac{i x_4^1}{C} L\left(x_\mu, iC\dfrac{x_j^1}{x_4^1}\right) \quad \ldots(5)$

The new Lagrangian $\wedge$ is a homogeneous function of the generalized velocities in the first degree :
$$\wedge = \left(x_\mu, \alpha x_\mu^1\right) = a \wedge\left(x_\mu, x_\mu^1\right) \qquad \ldots(6)$$

A Lagrangian obeying eq. (6) is often called a homogeneous Lagrangian. From Euler's theorem on homogeneous functions that if $\wedge$ is homogeneous to first degree in x_μ^1, then

$$\wedge = x_\mu^1 \frac{\partial \wedge}{\partial x_\mu^1}$$

One can then show that as a result the function $\wedge$ identically satisfies the relation

$$\left(\frac{d}{d\theta}\left(\frac{\partial \wedge}{\partial x_\mu^1}\right) - \frac{\partial \wedge}{\partial x_\mu}\right) x_\mu^1 = 0 \qquad \ldots(7)$$

Let us carry out the transformation for a free particle. Now the 'non covariant' Lagrangian for the free particle is $\quad L = -mC\sqrt{C^2 - \dot{x}_C \dot{x}_i}$

By the transformation of eq. (5) a possible covariant Lagrangian is then

6)

$$\wedge = -mC\sqrt{-x_\mu^1 x_\mu^1} \qquad ...(8)$$

With this Lagrangian the Euler-Lagrange equations are equivalent to

$$\frac{d}{d\theta}\left(\frac{mC x_v^1}{\sqrt{-x_\mu^1 x_\mu}}\right) = 0$$

The parameter θ must be a monotone function of the proper time τ so that derivatives with respect to θ are related to those interms of τ according to

$$x_v^1 = \frac{dx_v}{d\theta} = \frac{d\tau}{d\theta} u_v$$

Hence the Lagrangian equations correspond to

$$\frac{d}{d\tau}\left(\frac{mCu_v}{\sqrt{-u_v u_v}}\right) = \frac{d(mu_v)}{d\tau} = 0$$

Which are equations of motion for a single particle.

In the other method, the covariant Lagrange equations can therefore be written directly in terms of τ :

$$\frac{d}{d\tau}\left(\frac{\partial\wedge}{\partial u_v}\right) - \frac{\partial\wedge}{\partial x_v} = 0 \qquad ...(9)$$

Now it is required that $\wedge$ be a world scalar (or function of a world scalar) that leads to the correct equations of motion For example, a suitable $\wedge$ for a free particle would clearly be the quadratic expression

$$\wedge = \frac{1}{2}mu_v u_v \qquad ...(10)$$

We use eq. (8) for the 'kinetic energy' part of the Lagrangian in all subsequent discussions.

If the particle is not free, but is acted on y external forces, then interaction terms have to be added to the Lagrangian of eq. (10) that would lead to the corresponding Minkowski forces. Now we consider the example of a particle in electromagnetic field.

A suitable Lagrangian can easily be seen to be

$$\wedge(x_\mu, u_\mu) = \frac{1}{2}m u_\mu u_\mu + \frac{q}{C}u_\mu A_\mu(x_\lambda) \qquad ...(11)$$

The corresponding Lagrange's equations are then

$$\frac{d}{d\tau}(mu_v) = -\frac{q}{C}\frac{dA_v}{d\tau} + \frac{\partial}{\partial x_v}\left(\frac{q}{C}u_\mu A_\mu\right),$$

Which are exactly the generalized equations of motion $\frac{d}{d\tau}(mu_v') = k_v$, with the Minkowski force k_j on a charged particle,

$$k_\mu = \frac{q}{C}\left(\frac{\partial}{\partial x_\mu}(u_v A_v) - \frac{dA_\mu}{d\tau}\right)$$

Now $P_\mu = \frac{\partial\wedge}{\partial u_\mu} = mu_\mu + \frac{q}{C}\bar{A}_\mu = P_\mu + \frac{q}{C}\bar{A}_\mu \qquad ...(12)$

Thus again we note that the 'mechanical momentum' four-vector p_μ differs from the canonical momentum p_μ by a term linear in the electromagnetic potential. The canonical momentum conjugate to x_μ is now

$$p_\mu = \frac{iT}{C} + \frac{iq\phi}{C} = \frac{iE}{C}$$

Where E is the total energy of the particle, $T + q\phi$. The connection between the magnitude of the spatial 'mechanical' momentum and the energy T is given by $T^2 = P^2C^2 + m^2C^4$. From eq. (12) it is seen that the canonical momenta conjugate to x_i from the components of a spatial Cartesian vector f related to P by

$$f = P + \frac{q}{C} A \qquad \qquad ...(13)$$

Interms of f, eq. $T^2 = p^2C^2 + m^2C^4$ can be rewritten as

$$T^2 = \left(f - \frac{q}{C} A\right)^2 + m^2C^4 \qquad ...(14)$$

Which is a useful relation between the energy T and the canonical momentum vector f.

Summary : When the velocities of particles become comparable to the velocity of light the Newtonian mechanics fails and Galelian transformation do not holds good. We need a new transformation and a new mechanics. Such transformations are the Lorentz transformation and the new mechanics is called relativistic mechanics.

Using the basic postulates of special theory of relativity we derived the form of Lorentz Transformation. Depending on them we learned about the relativistic law of addition of velocities and kinematic effects of Lorentz transformation like Lorentz – Fitzerald Contraction, Time dilation. We came across the Covarient four-dimensional formulations, Force and energy relations in relativistic mechanics, Lagrangian formulation of relativistic Mechanics, and Covarient Lagrangian formulation.

Key words : Basic postulates of special theory of relativity. Galilean transformation . Postulates of special relativity . Lorentz Transformation . Kinematic effects of Lorentz transformation . Lorentz – Fitzerald Contraction. Time dilation. Covarient four-dimensional formulations Lagrangian formulation of relativistic Mechanics. Covarient Lagrangian formulation.

Self-assessment questions:
1. State the basic postulates of special theory of relativity and derive Lorentz transformation from them.
2. State and explain basic postulates of special theory of relativity. Deduce the formula for addition of velocities in relativisitic mechanics.
3. Discuss about co-variant four-dimensional formulation. Derive equations for force and energy in relativistic mechanics.
4. Explain Lagranian formulation of relativistic mechanics and also explain covariant Lagrangian formulations.

Reference Books :
1.Mechanics: Simon
2.Classical mechanics: H. Goldstein.
3. Classical mechanics: Gupta, Kumar and Sharma.

Classical Mechanics Part
UNIT-II
Lesson 6
Canonical (or) Contact Transformations

Objective:
1) To introduce the concept of Canonical transformations
2) To define the various generating functions
3) Explaining the Poisson brackets.
4) Proving the canonical invariance of poisson brackets.
5) Learning about lagrange brackets and their invariance under canonical transformation.

Structure of the lesson :

6.1 Transformation : A given system can be described by more than one set of generalized coordinates. We can choose a set, which is most convenient for the solution of the problem under consideration. For example to discuss the motion of a particle in a plane, we may use as generalized coordinates.

The Cartesian coordinates $q_1 = x$, $q_2 = y$

(or)   The plane polar coordinates $q_1 = r, q_2 = \theta$

We note that second set is more convenient because for central force θ is a cyclic coordinate, while in first set neither x nor y is cyclic. By choosing second set, solution of the problem becomes easy as only one variable r is to appear in the Hamiltonian. Thus we want to discuss here a specific procedure for transforming one set of variables in to other set which may be more convenient. If a problem has been formulated in the form of Hamilton's canonical equations. The contact transformations can be aimed to put these equations into a more easily soluble form i.e. to make integration of the equation of motion simpler.

Suppose we transform from Cartesian to plane polar coordinates, then transformation equations are

$$r = \sqrt{x^2 + y^2} = r(x, y) \qquad \theta = \tan^{-1}\left(\frac{y}{x}\right) = \theta(x, y)$$

This is an example of coordinate transformation. We shall discuss the transformation between generalized coordinates that are canonical.

6.1.1 Point Transformation

Generalising the form transformation from one set of coordinates q_j to a new set Q_j can be expressed as $Q_j = Q_j(q_j, t)$. Such transformations are called point transformations.

We know that configuration space is adequate only in providing the information about the position coordinates q_j and not about the velocities $\dot{q}_j$. Therefore transformations of the type

$Q_j = Q_j(q_j, t)$ can be referred to as the transformations of the configuration space (or) in other words point transformations are the transformations of configuration space.

6.1.2 Canonical Transformation

The transformations of the type represented by forthcoming equation in which Q_j and P_j are canonical are called contact (or) canonical transformations. It is essential to provide information of coordinates as well as of momenta. We know that phase space adequately provides the specification of coordinates q_j as well as of momenta P_j for the system under consideration. We can say that canonical transformations are the transformations of phase space. They are characterized by the property that they leave the form of Hamilton's equations of motion invariant.

We know that while deducing Lagrangian equations, no stress was given to any particular choice of coordinate system. Lagrangian equations of motion are invariant in form with respect to the choice of the set of any generalized coordinates. Therefore in neco set Q_j, Lagrange's equations will be

$$\left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{Q}_j} \right) - \frac{\partial L}{\partial Q_j} \right) = 0$$

i.e. Lagrange's equations are covariant with respect to point transformations and if we define P_j as

$$P_j = \frac{\partial L}{\partial \dot{Q}_j}\left(Q, \dot{Q}_j \right)$$

The Hamilton's canonical equations will also be covariant i.e.

$$\dot{Q}_j = \frac{\partial H}{\partial P_j}\left(Q_j, P_j \right) \qquad \dot{P}_j = \frac{-\partial H}{\partial Q_j}\left(Q_j, P_j \right)$$

Therefore this transformation is extended to Hamiltonian formulation. In Hamiltonian formulation we admit the existence of one more independent variable called momentum and above general form must be widened to accommodate the new variable. Consequently the simultaneous transformation of the independent coordinates and momenta q_j, p_j to a new set Q_j, P_j can be represented in the form

$$Q_j = Q_j(q, p, t)$$
$$P_j = P_j(q, p, t)$$

...(1)

For Q_j, P_j the new set of coordinates to be canonical like q_j and p_j of old set as is demanded in Hamiltonian formulation, they should be able to be expressed in Hamiltonian form of equations of motion i.e.,

$$\dot{Q}_j = \frac{\partial k}{\partial P_j}$$

$$\dot{P}_j = -\frac{\partial k}{\partial Q_j}$$

Where k is a function of (Q, P, t) i.e. $k(Q, P, t)$ and is substitute for the Hamiltonian H of old set in new set of coordinates. More over if Q_j, P_j are to be canonical coordinates they must also satisfy the modified Hamilton's principle of the form

$$\delta \int_{t_1}^{t_2} \left[\sum_j P_j \dot{Q}_j - k(Q, P, t) \right] dt = 0 \quad \dots(2)$$

The old coordinates p_j, q_j are already canonical.

$$\therefore \quad \delta \int_{t_1}^{t_2} \left[\sum_j p_j \dot{q}_j - H(q, p, t) \right] dt = 0 \quad \dots(3)$$

The simultaneous validity of above relations (2) & (3) as their right hand side is zero does not mean that the integrands of the two integrals are equal. We can write

$$\delta \int_{t_1}^{t_2} \left\{ \left(\sum_j p_j \dot{q}_j - H \right) - \sum_i \left(P_j \dot{Q}_j - k \right) \right\} dt = 0 \quad \dots(4)$$

in which the integrand to a certain extent is unknown. Eq. (4) will not be affected if we add to (or) subtract from it a total time derivative of a function $F = F(q, p, t)$ because

$$\delta \int_{t_1}^{t_2} \frac{dF}{dt} dt = \delta \left[F(q, p, t) \right]_{t_1}^{t_2}$$

$$= \frac{\partial F}{\partial q_j} \delta q_j \Big|_{t_1}^{t_2} + \frac{\partial F}{\partial P_j} \delta p_j \Big|_{t_1}^{t_2} = 0$$

Since at end points the variation in q_j and p_j vanishes. Therefore we can write eqn.

(4) as $\quad \delta \int_{t_1}^{t_2} \left\{ \left(\sum_j p_j \dot{q}_j - H \right) - \sum \left(P_j \dot{Q}_j - k \right) - \frac{dF}{dt} \right\} dt = 0$

Thus, it follows that $\quad \sum_j \left(p_j \dot{q}_j - H \right) - \left(\sum_j P_j \dot{Q}_j - k \right) = \frac{dF}{dt}$

6.2 Generating Function

The first bracket of $\sum_j \left(p_j \dot{q}_j - H \right) - \left(\sum_j P_j \dot{Q}_j - k \right) = \frac{dF}{dt} \quad \dots(1)$

is regarded as a function of q_j, p_j and t, the second as a function of Q_j, P_j and t. F is thus in general a function of $(4n+1)$ variables q_j, p_j, Q_j, P_j and t. But the two sets of variables (coordinates) are connected by the 2n transformation equations and out of 4n variables besides t only 2n are independent. Now F is a function of both old and new set of coordinates and therefore out of 2n variables. n should be taken from new and n from old set i.e., one variable should be out of p_j and q_j and other should be from Q_j, P_j set. Thus following four forms of function F are possible.

$$F_1(q,Q,t),\ F_2(q,P,t),\ F_3(p,Q,t),\ F_4(p,P,t)$$

As F is a function of old and new set of coordinates, it can affect the transformation from old set to new set, i.e. transformation relations can be derived by the knowledge of the function F. It is thus termed as the generating function. Out of these above four forms choice of one particular will depend upon the problem for example if we are dealing with a point transformation, then because of the relation $Q_j = Q_j(q,t)$, q_j and Q_j can not be treated as independent variables and hence first form $F_1(q,Q,t)$ of the generation function is to be excluded.

Now we shall consider the individual cases concerning these forms.

6.2.1. First form $F_1(q,Q,t)$: For this case we can write eqn. (1)

$$\sum_j p_j \dot{q}_j - H = \sum_j P_j \dot{Q}_j - k + \frac{dF_1}{dt}(q,Q,t) \quad ...(2)$$

Now $F_1 = F_1(q,Q,t)$ so that it's total time derivative is

$$\frac{dF_1}{dt} = \sum_j \frac{\partial F_1}{\partial q_j} \dot{q}_j + \sum_j \frac{\partial F_1}{\partial Q_j} \dot{Q}_j + \frac{\partial F_1}{\partial t}$$

Putting this in eq. (2) we get

$$\sum_j p_j \dot{q}_j - H = \sum_j P_j \dot{Q}_j - k + \sum_j \frac{\partial F_1}{\partial q_j} \dot{q}_j + \sum_j \frac{\partial F_1}{\partial Q_j} \dot{Q}_j + \frac{\partial F_1}{\partial t}$$

$$\sum_j \left(\frac{\partial F_1}{\partial q_j} - p_j\right) \dot{q}_j + \sum_j \left(P_j + \frac{\partial F_1}{\partial Q_j}\right) \dot{Q}_j \quad -k + H + \frac{\partial F_1}{\partial t} = 0 \quad ...(3)$$

since q_j and Q_j are to be treated as independent variables. Eq. (3) can hold only if the coefficient of q_j and Q_j separately vanish i.e.

$$p_j = \frac{\partial F_1}{\partial q_j}(q,Q,t) \qquad ...(4)$$

$$P_j = -\frac{\partial F_1}{\partial Q_j}(q,Q,t) \qquad ...(5)$$

$$k = H + \frac{\partial F_1}{\partial t}(q,Q,t) \qquad ...(6)$$

on solving eqn. (4) we find $\quad Q_j = Q_j(q_j, P_j, t) \qquad ...(7)$

which when substituted in eq. (5) gives

$$P_j = P_j(q_j, P_j, t) \qquad ...(8)$$

Eqns. (7) & (8) are the desired transformation equations and relation (6) gives connection between old and new Hamiltonian.

6.2.2. Second form $F_2(q,P,t)$: To apply eq. (2) for this case, change of basis of description from q,Q to q,P should be effected for which we use Legendre transformation discussed in brief as follows.

Consider a function of only two variables i.e. $f(x,y)$ so that

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

$$= u\,dx + v\,dy \qquad \qquad ...(A)$$

Suppose we want to change the basis of description from x,y to independent variables u and y then let g be the function of u and y defined by the equation

$$g(u,y) = f(x,y) - ux$$
$$dg = df - u\,dx - x\,du$$
$$= u\,dx + v\,dy - u\,dx - x\,du$$
$$= v\,dy - x\,du$$

which is exactly in the form desired so that we can write

$$= \frac{\partial g}{\partial y}dy + \frac{\partial g}{\partial u}du$$

where $v = \frac{\partial g}{\partial y}$ and $x = -\frac{\partial g}{\partial u}$

and is quite identical to equation (A)

thus if $u = \frac{\partial f}{\partial x}$, then the relation

$$g(u,y) = f(x,y) - ux \qquad \qquad ...(B)$$

Would be appropriate to effect a change from the basis x,y to u,y

Now we apply it to the present case.

Here $P_j = -\dfrac{\partial F_1}{\partial Q_j}$

Then putting $u = -P_j$, $x = Q_j$, $y = q_j$, $g = F_2$, $f = F_1$ and one more variable t in eq. (B), we get

$$F_2(q_j, P_j, t) = F_1(q_j, Q_j, t) + \sum_j P_j Q_j \qquad ...(9)$$

putting from eq. (9) the value of F_1 in eq. (2) we get

$$\sum_j p_j \dot{q}_j - H = \sum_j P_j \dot{Q}_j - k$$

$$+ \frac{d}{dt}\Box\left\{F_2(q_j, P_j, t) - \sum_j P_j Q_j\right\}$$

$$= \sum_j P_j \dot{Q}_j - k + \sum_j \frac{\partial F_2}{\partial q_j}\dot{q}_j + \sum_j \frac{\partial F_2}{\partial p_j}\dot{P}_j + \frac{\partial F_2}{\partial t} - \sum_j \dot{P}_j Q_j - \sum_j P_j \dot{Q}_j$$

$$= \sum_j P_j \dot{Q}_j - k + \frac{\partial F_2}{\partial t} + \sum_j \frac{\partial F_2}{\partial q_j}\dot{q}_j + \sum_j \frac{\partial F_2}{\partial P_j}\dot{P}_j \sum_j\left(\frac{\partial F_2}{\partial q_j} - P_j\right)\dot{q}_j + \sum_j\left(\frac{\partial F_2}{\partial p_j} - Q_j\right)\dot{P}_j$$

$$+ H + \frac{\partial F_2}{\partial t} - k = 0 \qquad ...(10)$$

Since q_j and P_j are independent variables.

Eq. (10) can be satisfied only when

$$p_j = \frac{\partial F_2}{\partial q_j}(q_j, P_j, t) \qquad \ldots(11)$$

$$Q_j = \frac{\partial F_2}{\partial P_j}(q, P_j, t) \qquad \ldots(12)$$

$$k = H + \frac{\partial F_2}{\partial t}(q_j, P_j, t) \qquad \ldots(13)$$

eq. (11) can be solved to give

$$P_j = P_j(q_j, p_j, t)$$

which when substituted in (12) will yield on solving

$$Q_j = Q_j(q_j, P_j, t)$$

which are desired transformation equations.

6.2.3. Third form $F_3(P, Q, t)$: To connect F_3 with F_1 we again apply Legendre transformation. For this case we see that

$$p_j = \frac{\partial F_1}{\partial q_j}$$

Also $g(u, y) = f(x, y) - ux$, where $u = \frac{\partial f}{\partial x}$ then taking $u = P_j, x = q_j, y = Q_j$, $g = F_3, f = F_1$ and one more

variable t, we write

$$F_3(Q, p, t) = F_1(Q, q, t) - \sum_j p_j q_j$$

$$F_1(Q, q, t) = F_3(Q, p, t) - \sum_j p_j q_j$$

equation (2) therefore appears as $\qquad -\sum_j q_j p_j - H = \sum_j P_j Q_j - k + \frac{d}{dt} F_3(p, Q, t)$

which on further simplification and on equating the coefficients yields the transformation equation.

$$q_j = -\frac{\partial F_3}{\partial P_j}(p_j, Q_j, t) \qquad \ldots(14)$$

$$P_j = -\frac{\partial F_3}{\partial Q_j}(p_j, Q_j, t) \qquad \ldots(15)$$

$$k = H + \frac{\partial F_3}{\partial t}(p_j, Q_j, t) \qquad \ldots(16)$$

eq. (14) gives $\qquad Q_j = Q_j(q, p, t)$

and eq. (15) gives $\qquad P_j = P_j(q, p, t)$

Which are the desired transformation equations.

6.2.4. Fourth form $Fu(p, P, t)$: F_4 and F_1 can be connected by Legendre transformation to give

$$F_4(p, P, t) = F_1(q, Q, t) + \sum_j P_j Q_j - \sum_j p_j q_j$$

and eq. (2) reduces to

$$-\sum_j p_j \dot{q}_j - H = -\sum_j Q_j \dot{P}_j - k + \frac{d}{dt} F_4(p,P,t)$$

which on following same procedure gives the following equations

$$q_j = -\frac{\partial F_4}{\partial P_j}(p_j, P_j, t) \qquad ...(17)$$

$$Q_j = \frac{\partial F_4}{\partial P_j}(p_j, P_j, t) \qquad ...(18)$$

$$k = H + \frac{\partial F_4}{\partial t}(p_j, P_j, t) \qquad ...(19)$$

Which lead to the transformations.

6.3 Examples of Canonical Transformation

Let us first consider few special cases of the forms of function when they do not involve time

(A) $\quad F = \sum_j q_j P_j$

The function is a special case of generating function F_2 and hence applying eqs. (11), (12) and (13) we get

$$P_j = \frac{\partial F_2}{\partial q_j} = \frac{\partial}{\partial q_j} \sum_j q_j P_j = P_j \qquad ...$$

$$Q_j = \frac{\partial F_2}{\partial P_j} = \frac{\partial}{\partial P_j} \sum_j q_j P_j = q_j$$

$$k = H + \frac{\partial F_2}{\partial t} = H, \text{ since } F_2 \neq f(t)$$

Thus the function F generates the identity transformation. Further if $F = \sum_j q_j P_j$ then this will give

$Q_j = -q_j, p_j = -P_j$ and $k = H$. This shows the fact that space inversion constitutes a special case of a canonical transformation.

(B) $\quad F = \sum_j q_j Q_j$

This corresponds to F_1, generating function when it is not a function of t and thus gives

$$p_j = Q_j \qquad P_j = -q_j \qquad H = k$$

(C) $\quad F = \sum_i f_i(q_j) P_i$ where f_i is arbitrary $\quad = F_2$

$$\therefore \quad P_j = \sum_i P_i \frac{\partial f_i}{\partial q_j} \qquad Q_j = f_j(q_k), k = H$$

This demonstrates the type of generating function required to generate a point transformation.

6.4 Examples 1 : Simple Harmonic oscillator.

Suppose Hamiltonian for a particle is given by $\quad H = \frac{1}{2}\left(kq^2 + \frac{p^2}{m}\right)$ and we want to determine its motion.

The problem is given in canonical coordinates p_j and q_j. We shall transform this into new set of canonical coordinates P_j and Q_j in which all Q_j are cyclic and corresponding momenta are constant. Then we shall integrate the equation of motion, so obtained to get the result. Finally we shall answer interms of the same coordinates p_j and q_j.

Consider that the contact transformations are generated by the function.

$$F_1 = \mu q^2 \cot Q \qquad \qquad ...(1)$$

so that for this transformation

$$p = \frac{\partial F_1}{\partial q} = 2\mu q \cot Q$$

$$P = -\frac{\partial F_1}{\partial Q} = \mu q^2 \cos ec^2 Q \qquad ...(2)$$

$$H = k \; [\sin ce \; F_1 = f(t)]$$

from eq. (2) we arrive at

$$\dot{q} = \sqrt{P/u} \; \sin Q$$

$$p = \sqrt{4\mu P} \cos Q \qquad \qquad ...(3)$$

so that new Hamiltonian, k is

$$\therefore \quad k = \frac{1}{2}\left(kq^2 + \frac{P^2}{M}\right)$$

$$= \frac{1}{2}\left(k\frac{P}{\mu}\sin^2 Q + \frac{4\mu P}{M}\cos^2 Q\right)$$

$$= \frac{kP}{2\mu}\left(\sin^2 Q + \frac{4\mu^2}{Mk}\cos^2 Q\right) \quad ...(4)$$

Thus we have obtained Hamiltonian in the set of coordinates P and Q.

If $\mu = \frac{1}{2}\sqrt{mk}$ this reduces to

$$k = \frac{kP}{2\mu} = P\sqrt{\frac{k}{M}} \qquad \qquad ...(5)$$

Which is in quite simple form since Q does not appear in k and so is an ignorable coordinate, its corresponding momentum will be constant, i.e.,

$$\dot{P} = -\frac{\partial k}{\partial Q} = 0$$

So that $P = \alpha = $ constant

Also $\dot{Q} = \frac{\partial k}{\partial P} = \sqrt{\frac{k}{m}} = a$ constant

So that $Q = \sqrt{\dfrac{k}{M}} \cdot t + \beta$ ... (6)

Where β is a constant of integration fixed by the initial conditions. This illustrates that how easy it becomes, in new set of coordinates, to integrate the Hamilton's equations of motion. Finally the result is to be given in the same set in which question is asked. We take the following steps.

Putting this value of Q in the expression for q i.e., eq. (3) we get

$$q = \sqrt{\frac{P}{\mu}} \sin\left\{ \sqrt{\frac{k}{M}} \cdot t + \beta \right\}$$

$$= \sqrt{\frac{\alpha}{\mu}} \sin\left\{ \sqrt{\frac{k}{M}} \cdot t + \beta \right\} \qquad ...(7)$$

Which is the equation for simple harmonic motion executed by simple harmonic oscillator.

Example 2 : Show that the transformation $P = \dfrac{1}{2}\left(p^2 + q^2\right)$

$Q = \tan^{-1}\left(q/p\right)$ is canonical.

It is quite obvious that for such a case $H(q, p) = k(Q, P)$ since the generating function does not involve time. Since old set of coordinates q, p are canonical.

$$\dot{P} = -\frac{\partial H}{\partial q} \qquad \dot{q} = +\frac{\partial H}{\partial P} \qquad ...(1)$$

But $\qquad \dot{P} = -\dfrac{\partial p}{\partial P}.\dot{P} + \dfrac{\partial P}{\partial Q}.\dot{Q},$

$$\dot{q} = -\frac{\partial q}{\partial P}.\dot{P} + \frac{\partial q}{\partial Q}.\dot{Q} \qquad ...(2)$$

$$\frac{\partial H}{\partial q} = \frac{\partial k}{\partial P}.\frac{\partial P}{\partial q} + \frac{\partial k}{\partial Q}\frac{\partial Q}{\partial q}$$

$$\frac{\partial H}{\partial P} = \frac{\partial k}{\partial P}.\frac{\partial P}{\partial p} + \frac{\partial k}{\partial Q}\frac{\partial Q}{\partial p} \qquad ...(3)$$

Form the given transformation equations, we have

$$\frac{\partial P}{\partial p} = P, \frac{\partial P}{\partial q} = q, \frac{\partial Q}{\partial p} = \frac{-q}{p^2 + q^2}, \frac{\partial Q}{\partial q} = \frac{p}{p^2 + q^2}$$

Also differentiating transformation equations with respect to P and Q respectively we find

$$1 = P\frac{\partial p}{\partial P} + q\frac{\partial q}{\partial P} \qquad 0 = \left(P\frac{\partial q}{\partial P} - q\frac{\partial p}{\partial P}\right)\Big/ p^2 + q^2$$

$$0 = p\frac{\partial q}{\partial Q} + q\frac{\partial p}{\partial Q} \qquad 1 = \left(P\frac{\partial q}{\partial Q} - q\frac{\partial p}{\partial Q}\right)\Big/ p^2 + q^2$$

Solving simultaneously, we find

$$\frac{\partial p}{\partial P} = \frac{p}{p^2 + q^2}, \frac{\partial q}{\partial P} = \frac{q}{p^2 + q^2}, \frac{\partial p}{\partial Q} = -q, \frac{\partial q}{\partial Q} = p$$

Then equations (1) and (2) become

$$\dot{p} = \frac{p}{\left(p^2 + q^2\right)}\dot{P} - q\dot{Q} \ , \quad \dot{q} = \frac{q}{\left(p^2 + q^2\right)}\dot{P} + p\dot{Q} \quad ...(4)$$

$$\frac{\partial H}{\partial q} = q \frac{\partial k}{\partial P} + \frac{p}{\left(p^2 + q^2\right)} \frac{\partial k}{\partial Q} \quad , \quad \frac{\partial H}{\partial p} = p \frac{\partial k}{\partial P} - \frac{q}{\left(p^2 + q^2\right)} \frac{\partial k}{\partial Q}$$

$$...(5)$$

Thus from equations (1), (4) and (5) we have

$$\frac{p}{\left(p^2 + q^2\right)} \dot{P} - q \dot{Q} = -q \frac{\partial k}{\partial P} - \frac{p}{\left(p^2 + q^2\right)} \frac{\partial k}{\partial Q}$$

$$\frac{q}{\left(p^2 + q^2\right)} \dot{P} - p \dot{Q} = p \frac{\partial k}{\partial P} - \frac{q}{\left(p^2 + q^2\right)} \frac{\partial k}{\partial Q}$$

Solving these simultaneously, we find

$$\dot{P} = -\frac{\partial k}{\partial Q} \qquad \dot{Q} = \frac{\partial k}{\partial P}$$

Form which we infer that P and Q are canonical and hence the transformation is canonical.

Example : 3 For a certain canonical transformation it is known that

$$Q = \sqrt{q^2 + p^2}$$

$$F = \frac{1}{2}\left(q^2 + p^2\right)\tan^{-1}\left(\frac{q}{p}\right) + \frac{1}{2}qp.$$

Find $p(q, p)$ and $F(q, Q)$

Form the expression for Q. We find p and put it in F to get $F(q, Q)$ that is

$$p = \sqrt{Q^2 - q^2}$$

So that $F = \frac{1}{2}Q^2 \tan^{-1}\frac{q}{\sqrt{Q^2 - q^2}} + \frac{1}{2}q\sqrt{Q^2 - q^2}$

$$= \frac{1}{2}Q^2 \sin^{-1}\frac{q}{Q} + \frac{1}{2}q\sqrt{Q^2 - q^2}$$

We now use equation of transformation

$$P = \frac{-\partial F}{\partial Q}$$

$$= \frac{\partial}{\partial Q}\left[\frac{1}{2}Q^2 \sin^{-1}\frac{q}{Q} + \frac{1}{2}q\sqrt{Q^2 - q^2}\right]$$

$$= -Q\sin^{-1}\frac{q}{Q} + \frac{1}{2}Q^2 \frac{-q}{Q\sqrt{Q^2 - q^2}}$$

$$+ \frac{1}{2}q \times \frac{1}{2}\left(Q^2 - q^2 - \frac{1}{2}Q\right)$$

$$= -\left[Q\sin^{-1}\frac{q}{Q} - \frac{1}{2}\frac{qQ}{\sqrt{Q^2 - q^2}} + \frac{1}{2}\frac{qQ}{\sqrt{Q^2 - q^2}}\right] \quad P(q, Q) = -Q\sin^{-1}\frac{q}{Q}$$

As we want $P(q, p)$, change Q into $\sqrt{q^2 + p^2}$ to get

$$P(q, p) = -\sqrt{q^2 + p^2} \; \sin^{-1} \frac{q}{\sqrt{q^2 + p^2}}$$

$$= -\sqrt{q^2 + p^2} \; \tan^{-1(q/p)}$$

6.5 POISSON BRACKETS

Let F be any dynamical variable of a system. Suppose F is function of conjugate variables q_j and p_j and t then

$$\frac{dF}{dt} = \frac{dF}{dt}(q_j, p_j, t) = \sum_j \frac{\partial F}{\partial q_j} \cdot \dot{q}_j + \sum_j \frac{\partial F}{\partial p_j} \cdot \dot{p}_j + \frac{\partial F}{\partial t}$$

$$= \sum_j \left(\frac{\partial F}{\partial q_j} \frac{\partial H}{\partial p_j} - \frac{\partial F}{\partial p_j} \frac{\partial H}{\partial q_j} \right) + \frac{\partial F}{\partial t}$$

On using Hamilton's canonical equations of motion.
The first bracketed term is called Poisson Bracket of F with H. In general if x and y are two dynamical variables then their Poisson bracket is defined as

$$[x, y]_{q,p} = \sum_j \left[\frac{\partial x}{\partial q_j} \frac{\partial y}{\partial p_j} - \frac{\partial x}{\partial p_j} \frac{\partial y}{\partial q_j} \right] \qquad ...(1)$$

from which it is quite easy to arrive at the following properties of Poisson brackets.

Properties :
$$[x, y] = -[y, x]$$
$$[x, x] = 0$$
$$[x, y + z] = [x, y] + [x, z] \qquad ...(2)$$
$$[x, y, z] = y[x, z] + [x, y]z$$

$$[q_i, q_j]_{q,p} = 0 = [p_i, p_j]_{q,p}$$

Also $$[q_i, p_j]_{q,p} = 0, \text{ if } i \neq j \qquad ...(3)$$
$$= 1, \text{ if } i = j$$

The equation (3) is known as the fundamental basic Poisson brackets.
Property III of eq. (2) can be obtained as follows.

$$[x, y + z] = \sum_j \left(\frac{\partial x}{\partial q_j} \frac{\partial (y + z)}{\partial p_j} - \frac{\partial x}{\partial p_j} \frac{\partial (y + z)}{\partial q_j} \right)$$

$$= \sum_j \left(\frac{\partial x}{\partial q_j} \frac{\partial y}{\partial p_j} + \frac{\partial x}{\partial q_j} \frac{\partial z}{\partial p_j} \right) - \sum_j \left(\frac{\partial x}{\partial p_j} \frac{\partial y}{\partial q_j} + \frac{\partial x}{\partial p_j} \frac{\partial z}{\partial q_j} \right)$$

$$= \sum_j \left(\frac{\partial x}{\partial q_j} \frac{\partial y}{\partial p_j} - \frac{\partial x}{\partial p_j} \frac{\partial y}{\partial q_j} \right) + \sum_j \left(\frac{\partial x}{\partial q_j} \frac{\partial z}{\partial p_j} - \frac{\partial x}{\partial p_j} \frac{\partial z}{\partial q_j} \right)$$

$$= [x, y] + [x, z]$$

Properties of eqn. (3) can be obtained as follows.

$$[q_i, q_j]_{q,p} = \sum_k \left(\frac{\partial q_i}{\partial q_k} \frac{\partial q_j}{\partial p_k} - \frac{\partial q_i}{\partial p_k} \frac{\partial q_j}{\partial q_k} \right)$$

Because q_i or q_j is not the function of p_k. So

$$\frac{\partial q_j}{\partial p_k} = \frac{\partial q_i}{\partial p_k} = 0$$

giving $[q_i, q_j]_{q,p} = 0 = [p_i \ p_j]_{q,p}$

Further $[q_i, p_j]_{q,p} = \sum_k \left(\frac{\partial q_i}{\partial q_k} \frac{\partial p_j}{\partial p_k} - \frac{\partial q_i}{\partial p_k} \frac{\partial p_j}{\partial q_k} \right) = \sum_k \frac{\partial q_i}{\partial q_k} \frac{\partial p_j}{\partial p_k}$

the second term vanishes because $\frac{\partial q_i}{\partial p_k} = 0$ and $\frac{\partial p_j}{\partial q_k} = 0$

Also $\frac{\partial q_i}{\partial q_k} = \delta_{ik}$ and $\frac{\partial p_j}{\partial p_k} = \delta_{jk}$

So that $[q_i, p_j]_{q,p} = \sum \delta_{ik} \, \delta_{jk}$

$$= \delta_{ij}$$
$$= 0 \quad \text{if } i \ne j$$
$$= 1 \quad \text{if } i = j$$

Example : If $[\phi, \psi]$ is the Poisson bracket of ϕ and ψ then prove that

a) $\frac{\partial}{\partial t}[\phi, \psi] = \left[\frac{\partial \phi}{\partial t}, \psi \right] + \left[\phi, \frac{\partial \psi}{\partial t} \right]$

b) $\frac{d}{dt}[\phi, \psi] = \left[\frac{d\phi}{dt}, \psi \right] + \left[\phi, \frac{d\psi}{dt} \right]$

We have from the definition of poisson bracket that $[\phi, \psi] = \frac{\partial \phi}{\partial q_i} \frac{\partial \psi}{\partial p_j} - \frac{\partial \psi}{\partial q_i} \frac{\partial \phi}{\partial p_i}$

$$\frac{\partial}{\partial t}[\phi, \psi] = \frac{\partial}{\partial t} \left(\frac{\partial \phi}{\partial q_i} \frac{\partial \psi}{\partial p_j} - \frac{\partial \psi}{\partial q_i} \frac{\partial \phi}{\partial p_i} \right)$$

$$= \frac{\partial}{\partial q_i} \left(\frac{\partial \phi}{\partial t} \right) \frac{\partial \psi}{\partial p_i} + \frac{\partial \phi}{\partial q_i} \frac{\partial}{\partial p_i} \left(\frac{\partial \psi}{\partial t} \right) - \frac{\partial}{\partial q_i} \left(\frac{\partial \psi}{\partial t} \right) \frac{\partial \phi}{\partial p_i} - \frac{\partial \psi}{\partial q_i} \frac{\partial}{\partial p_i} \left(\frac{\partial \phi}{\partial t} \right)$$

$$= \left\{ \frac{\partial}{\partial q_i} \left(\frac{\partial \phi}{\partial t} \right) \frac{\partial \psi}{\partial p_i} - \frac{\partial \psi}{\partial q_i} \frac{\partial}{\partial p_i} \left(\frac{\partial \phi}{\partial t} \right) \right\} + \left\{ \frac{\partial \phi}{\partial q_i} \frac{\partial}{\partial p_i} \left(\frac{\partial \psi}{\partial t} \right) - \frac{\partial}{\partial q_i} \left(\frac{\partial \psi}{\partial t} \right) \frac{\partial \phi}{\partial p_i} \right\}$$

$$= \left[\frac{\partial \phi}{\partial t}, \psi \right] + \left[\phi, \frac{\partial \psi}{\partial t} \right]$$

6.6 Invariance of Poisson Brackets with respect to canonical Transformation

Poisson brackets are invariant under a canonical transformation. We can express this mathematically as

$$[x, y]_{q,p} = [x, y]_{Q,P} \quad \text{-------(1)}$$

The equivalence ensures the equality in value and not of form i.e. x and y do not necessarily have the same form in terms of the two sets of coordinates.

We consider first the general case

$$[x, y]_{Q,P} = \sum_i \left(\frac{\partial x}{\partial Q_i} \frac{\partial y}{\partial P_i} - \frac{\partial x}{\partial P_i} \frac{\partial y}{\partial Q_i} \right) = \sum_{i,j} \left\{ \frac{\partial x}{\partial Q_i} \left(\frac{\partial y}{\partial q_j} \frac{\partial q_j}{\partial P_i} + \frac{\partial y}{\partial p_j} \frac{\partial p_j}{\partial P_i} \right) - \frac{\partial x}{\partial P_i} \left(\frac{\partial y}{\partial q_j} \frac{\partial q_j}{\partial Q_i} + \frac{\partial y}{\partial p_j} \frac{\partial p_j}{\partial Q_i} \right) \right\}$$

$$= \sum_j \left\{ \frac{\partial y}{\partial q_j} [x, q_j]_{Q,P} + \frac{\partial y}{\partial p_j} [x, p_j]_{Q,P} \right\} \quad \ldots(2)$$

On arranging and recollecting the terms.

From eq. (2) we can write $(q_j, x)_{Q,P}$ simply by putting q_j for x and x for y (or) alternatively

$$[x, q_j]_{Q,P} = -[q_j, x]_{Q,P}$$

$$= -\sum_m \left\{ \frac{\partial q_j}{\partial Q_m} \frac{\partial x}{\partial P_m} - \frac{\partial q_j}{\partial P_m} \frac{\partial x}{\partial Q_m} \right\}$$

$$= -\sum_m \left\{ \frac{\partial q_j}{\partial Q_m} \left(\frac{\partial x}{\partial q_k} \frac{\partial q_k}{\partial Pm} + \frac{\partial x}{\partial p_k} \frac{\partial p_k}{\partial P_m} \right) - \frac{\partial q_j}{\partial P_m} \left(\frac{\partial x}{\partial q_k} \frac{\partial q_k}{\partial Q_m} + \frac{\partial x}{\partial p_k} \frac{\partial p_k}{\partial Q_m} \right) \right\}$$

$$= -\sum_k \left\{ \frac{\partial x}{\partial q_k} \sum_m \left(\frac{\partial q_j}{\partial Q_m} \frac{\partial q_k}{\partial P_m} - \frac{\partial q_j}{\partial P_m} \frac{\partial q_k}{\partial Q_m} \right) + \frac{\partial x}{\partial p_k} \sum_m \left(\frac{\partial q_j}{\partial Q_M} \frac{\partial p_k}{\partial P_{mm}} - \frac{\partial q_j}{\partial P_m} \frac{\partial p_k}{\partial Q_m} \right) \right\}$$

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

$$= -\sum_k \frac{\partial x}{\partial p_k} \delta_{jk} = -\frac{\partial x}{\partial p_j} \quad \ldots(3)$$

similarly $[x, P_j]_{Q,P} = \frac{\partial x}{\partial q_j}$

putting equations (3) and (4) in equation (2), we get

$$[x, y]_{Q,P} = \sum_j \left\{ -\frac{\partial y}{\partial q_j} \frac{\partial x}{\partial p_j} + \frac{\partial y}{\partial p_j} \frac{\partial x}{\partial q_j} \right\} = [x, y]_{q,p} \quad \ldots(4)$$

Which is the desired result. Since with either of the set of conjugate variables i.e. with (q, p) or with (Q, P) the value of the Poisson bracket remains the same, it is of no use to write the subscript on the bracket and we shall now omit the same.

We have shown that a canonical transformation can be generated from functions $F_1(q_j, Q_j, t)$, $F_2(q_j, P_j, t)$, $F_3(p_j, Q_j, t)$ and $F_4(p_j, P_j, t)$. In the case of F_1 generating function, we have derived the relations

$$p_j = \frac{\partial F_1}{\partial q_j} \qquad P_j = -\frac{\partial F_1}{\partial Q_j}$$

From these relations it is quite easy to see that

$$\frac{\partial p_j}{\partial Q_i} = \frac{\partial^2 F_1}{\partial Q_i \partial q_j} = -\frac{\partial P_i}{\partial q_j} \quad \ldots(5)$$

Similarly in the case of $F_2(q_j, P_j, t)$ generating function the relations are

$$p_j = \frac{\partial F_2}{\partial q_j} \qquad Q_j = \frac{\partial F_2}{\partial P_j}$$

From these relations we find that

$$\frac{\partial p_i}{\partial P_i} = \frac{\partial^2 F_2}{\partial P_i \partial q_j} = \frac{\partial Q_i}{\partial q_j} \qquad \dots(6)$$

Similarly from F_3 and F_4 generating functions we can find that

$$\frac{\partial q_j}{\partial Q_i} = \frac{\partial P_i}{\partial p_j} \qquad \dots(7) \qquad\qquad \frac{\partial q_j}{\partial P_i} = -\frac{\partial Q_i}{\partial p_j} \qquad \dots(8)$$

Now we write the fundamental Poisson bracket

$$[Q_i, P_j]_{q,p} = \sum_k \left(\frac{\partial Q_i}{\partial q_k} \frac{\partial P_j}{\partial p_k} - \frac{\partial Q_i}{\partial p_k} \frac{\partial P_j}{\partial q_k} \right)$$

Applying the results of equations (7) and (5) we find

$$[Q_i, P_j]_{q,p} = \sum_k \left(\frac{\partial Q_i}{\partial q_k} \frac{\partial q_k}{\partial Q_j} + \frac{\partial Q_i}{\partial p_k} \frac{\partial p_k}{\partial Q_j} \right)$$

$$= \frac{\partial Q_i}{\partial Q_j} = \delta_{ij} \quad = [Q_i, P_j]_{Q,P}$$

and similarly $\quad [Q_i, Q_j]_{q,p} = 0 = [Q_i, Q_j]_{Q,P}$

$$[P_i, P_j]_{q,p} = 0 \; [P_i, P_j]_{Q,P} \quad \dots(9)$$

Thus we have proved the assertion for the fundamental brackets.

6.7 LAGRANGE'S BRACKETS

Lagrange's bracket of (u, v) with respect to the basis (q_j, P_j) is defined as

$$(u, v)_{q,p} = \sum_j \left(\frac{\partial q_j}{\partial u} \frac{\partial p_j}{\partial v} - \frac{\partial p_j}{\partial u} \frac{\partial q_j}{\partial v} \right) \quad \dots(1)$$

We note that
a) Lagrange bracket is invariant under canonical transformation
Poincares theorem states that the integral

$$J_1 = \iint_S \sum_J dq_j \, dp_j \qquad \dots(2)$$

Taken over an arbitrary two-dimensional surface S of the 2n dimensional (q, p) phase space is invariant under canonical transformation.
Position of a point on any two dimensional surface is specified completely by two parameters e.g. : u, v expressed as

$$\left. \begin{array}{l} q_j = q_j (u, v) \\ p_j = p_j (u, v) \end{array} \right\} \qquad \dots(3)$$

Transforming the integral (2) in terms of new variables (u, v) by means of Jacobian.

$$dq_j \, dp_j = \frac{\partial(q_j, p_j)}{\partial(u, v)} \, du \, .dv \qquad \dots(4)$$

$$\frac{\partial(q_j, p_j)}{\partial(u, v)} = \begin{vmatrix} \dfrac{\partial q_j}{\partial u} & \dfrac{\partial p_j}{\partial u} \\ \dfrac{\partial q_j}{\partial v} & \dfrac{\partial p_j}{\partial v} \end{vmatrix} \qquad \dots(5)$$

as the Jacobian

Further the statement that the integral J_1 remains invariant implies that

$$\iint_S \sum_j dq_j \, dp_j = \iint_S \sum_j dQ_j \, dP_j \quad ...(6)$$

Where (Q_j, P_j) is another set of canonical coordinates to which the set (q_j, p_j) has been transformed.

From equation (5) & (6)

$$\iint_S \sum_j \frac{\partial(q_j, p_j)}{\partial(u, v)} \, du \, dv = \iint_S \sum_j \frac{\partial(Q_j, P_j)}{\partial(u, v)} \, du \, dv$$

Since the surface S is arbitrary area $du dv$ is arbitrary and therefore expressions on both the sides will be equal only when the integrands are identical, i.e.,

if $$\sum_j \frac{\partial(q_j, p_j)}{\partial(u, v)} = \sum_j \frac{\partial(Q_j, P_j)}{\partial(u, v)}$$

$$\sum_j \begin{vmatrix} \frac{\partial q_j}{\partial u} & \frac{\partial p_j}{\partial u} \\ \frac{\partial q_j}{\partial v} & \frac{\partial p_j}{\partial v} \end{vmatrix} = \sum_j \begin{vmatrix} \frac{\partial Q_j}{\partial u} & \frac{\partial P_j}{\partial u} \\ \frac{\partial Q_j}{\partial v} & \frac{\partial P_j}{\partial v} \end{vmatrix}$$

$$\sum_j \left(\frac{\partial q_j}{\partial u} \frac{\partial p_j}{\partial v} - \frac{\partial p_j}{\partial u} \frac{\partial q_j}{\partial v} \right) = \sum_j \left(\frac{\partial Q_j}{\partial u} \frac{\partial P_j}{\partial v} - \frac{\partial P_j}{\partial u} \frac{\partial Q_j}{\partial v} \right)$$

$$\{u, v\}_{q,p} = \{u, v\}_{Q,P} \quad ...(7)$$

in Lagrange's brackets notation.

Thus Lagrange's bracket is invariant under canonical transformation and it is worthless to designate any basis to the bracket i.e. hence drop the subscripts q, p or Q, P.

Lagrange brackets do not obey the commutative law :

$$\{u, v\} = \sum_j \left(\frac{\partial q_j}{\partial u} \frac{\partial p_j}{\partial v} - \frac{\partial p_j}{\partial u} \frac{\partial q_j}{\partial v} \right) \quad = -\sum_j \left(\frac{\partial p_j}{\partial u} \frac{\partial q_j}{\partial v} - \frac{\partial q_j}{\partial u} \frac{\partial p_j}{\partial v} \right)$$

$$= -\sum_j \left(\frac{\partial q_j}{\partial v} \frac{\partial p_j}{\partial u} - \frac{\partial p_j}{\partial v} \frac{\partial q_j}{\partial u} \right) \quad = -\{v, u\} \quad ...(8)$$

$$\{q_i, q_j\} = 0$$

we find that $$\{p_j, p_j\} = 0$$

$$\{q_i, p_j\} = \delta_{ij}$$

The relations at once follow from the general expression for Lagrange's bracket because of independent nature of coordinates (q, p).

$$\{q_i, q_j\} = \sum_k \left(\frac{\partial q_k}{\partial q_i} \frac{\partial p_k}{\partial q_j} - \frac{\partial q_k}{\partial q_j} \frac{\partial p_k}{\partial q_i} \right) \quad ...(9)$$

But since $q's$ and $p's$ are independent.

$$\frac{\partial p_k}{\partial q_j} = 0 \text{ and } \frac{\partial p_k}{\partial q_i} = 0$$

From equation (9) $\{q_i, q_j\} = 0$ Similarly $\{p_i, p_j\} = 0$ ...(10)

$$\{q_i, p_j\} = \sum_k \left(\frac{\partial q_k}{\partial q_i} \frac{\partial p_k}{\partial p_j} - \frac{\partial q_k}{\partial p_j} \frac{\partial q_k}{\partial q_i} \right) \quad = \sum_k \frac{\partial q_k}{\partial q_i} \frac{\partial p_k}{\partial P_j} \quad \dots(11)$$

The second term vanishes because of the same reason. The partial derivatives have the values

$$\frac{\partial q_k}{\partial q_i} = \delta_{ki} \quad \text{and} \quad \frac{\partial p_k}{\partial p_j} = \delta_{kj}$$

$$\{q_i, p_j\} = \sum_k \delta_{ki} \, \delta_{kj} = \delta_{ij} \quad \dots(12)$$

Summary: We have explained canonical transformations and the four forms of generating functions . The concept of Poisson brackets and their invariance under canonical transformations are discussed. Lagrange brackets and their invariance under canonical transformations are explained.

Key words : Transformation – Point transformation – canonical transformation - Generating function – Simple harmonic oscillator – Poisson bracket –Legrange bracket

Self- Assessment questions
1. Derive the canonical transformation equations?
2. What is a generating function? Obtain the four forms of generating function.
3. Define Poisson bracket? Show that Poisson brackets are canonically invariant.
4. Define Lagrange's bracket? Show that Lagrange's brackets are invariant under canonical transformation.
5. Show that the transformation $P = \frac{1}{2}\left(p^2 + q^2\right)$ $Q = \tan^{-1}(q/p)$ is canonical.

Reference Books :
1.Mechanics: Simon
2.Classical mechanics: H. Goldstein.
3. Classical mechanics: Gupta, Kumar and Sharma

Classical Mechanics Part

UNIT-II

Lesson 7

HAMILTON – JACOBI METHOD

Objectives:
1. Deriving the Hamilton – Jacobi equation.
2. Solving Hamilton – Jacobi equation
3. Applying Hamilton – Jacobi theory to Kepler's problem.
4. Knowing about Action and angle variables
5. Solving Harmonic oscillator problem using action angle variable method

Structure:
7.1 Hamilton – Jacobi partial differential equation
7.2 Solution of Hamilton – Jacobi equation
7.3 Hamilton – Jacobi equation for Hamilton 's characteristic function.
7.4 Physical significance of Hamilton 's characteristic function
7.5 Solution of Kepler's problem by Hamilton – Jacobi method
7.6 Action and angle variables
7.7 Solution of harmonic oscillator equation by action angle variable method
7.8 Separation of variables in Hamilton – Jacobi equation.

Introduction:
Canonical transformations provide a general procedure for easy solution of mechanical problems. There are two ways of effecting such transformation.

i) One way to obtain the solution of a mechanical problem is to transform old set of coordinates into new set of coordinates that are all cyclic and consequently all momenta are constants. In this way the new equations of motion can be integrated to give a solution and is adopted with Hamiltonian H is conserved. This way was only a sort of coordinate transformation and not a method in it self. We switch $H(q_j, p_j)$ to $k(Q_j, P_j)$ and then apply Hamilton's equations of motion, which are then integrated to yield the result.

ii) Another way to obtain the solution is to seek canonical transformation from coordinates and momenta (q, p) at time t to a new set of constant quantities which may be $2n$ initial values (q_0, p_0) at $t = 0$.

The transformation equations are

$$q = q(q_0, p_0, t) \qquad p = p(q_0, p_0, t)$$

Since these equations of transformation give the values of q, p in terms of their initial values $(q_0 \ p_0)$ at time $t = 0$. They are the desired solution of a mechanical problem. Thus greatest advantage of such a transformation is that we are doubly benefited i.e. in obtaining transformation equations we arrive at the solution as well. The way is due to Jacobi. Thus Jacobi's way is a transformation as well as a method in itself. In it we do not first transform and then apply Hamilton's equations of motion but while performing transformation we are arriving at the result as well.

This procedure is more general because it can be applied to the cases for which Hamiltonian involves the time.

7.1 Hamilton – Jacobi Partial Differential Equation

Under the second kind transformation, as the new set is of constant coordinates (initial values q_0, p_0). If we require that the transformed Hamiltonian k is zero, then the new equations of motion (involving coordinates of transformed set P_j, Q_j) are

$$\left. \begin{array}{l} \dot{Q}_j = \dfrac{\partial k}{\partial P_j} = 0 \\[2mm] \dot{P}_j = -\dfrac{\partial k}{\partial Q_j} = 0 \end{array} \right\} \qquad ...(1)$$

Which obviously ensures that the new coordinates P_j, Q_j are constant in time i.e. both the coordinates are rendered cyclic in this technique.

The new Hamiltonian k is related to old Hamiltonian H and to the generating function F by the equation

$$K = H + \frac{\partial F}{\partial t}$$

Which will be zero only when

$$H(q, p, t) + \frac{\partial F}{\partial t} = 0 \qquad ...(2)$$

The generating function F is to be the function of one old coordinate and one new coordinate because it has only 2n independent variables and performs the transformation between old and new set. Let us take F to be a function of $(q, p, t)^*$. With such a generating function, we can write the equation of transformation as

$$p_j = \frac{\partial F}{\partial q_j}(q, p, t)$$

With this substitution Hamiltonian $H(q_j, p_j, t)$ becomes $H(q_j, \dfrac{\partial F}{\partial q_j}, t)$ and the equation

$$H(q_1, q_2, ...q_n : P_1, P_2, ...P_n : t) + \frac{\partial F}{\partial t}(q, P, t) = 0$$

$$H\left(q_1, q_2, ...q_n : \frac{\partial F}{\partial q_1}, ... \frac{\partial F}{\partial q_n} : t \right) + \frac{\partial F}{\partial t}(q, P, t) = 0 \quad ...(3)$$

Equation (3) is a partial differential equation in (n+1) variables (n for q's and one for t and p is constant) while eqn. (2) is in $(2n+1)$ variables. This shows that the above substitution has reduced the number of variables by n. Eqn. (3) is called Hamiltonian Jacobi's equation, the solution of which is called Hamilton's principal function and is denoted by S. Thus equation (3) is also termed as H-J equation for the Hamilton's principal function.

7.2 Solution of Hamilton Jacobi equation

Solution of Hamilton Jacobi equation is to obtain transformation equation in which coordinate of new set be constant quantities. Since these coordinates, being constant quantities can be taken as initial values the solution of Hamilton – Jacobi equation is the desired solution of the mechanical problem. We can express the transformation equation and hence the solution as

$$q = q(q_0, p_0, t) \qquad\qquad q = q(\beta_j, \alpha_j, t)$$

Where β_j and α_j are related to initial values of q and p at the time t.

To obtain solution of Hamilton Jacobi equation (3), we integrate equation (3) to obtain its solution S, the Hamilton's principal function, then it will merely provide dependence on the old coordinates and time. It will not be apparent from the solution as to how the new momenta P_j appear in S.

Equation (3) has the form of a first order partial differential equation in $(n+1)$ variables. Consequently a complete solution must involve $(n+1)$ independent constants of integration $\alpha_1, \alpha_2, ..., \alpha_{n+1}$. As eqn. (3) whose solution is S, is a partial differential equation, only derivatives of S with respect to q_j and t will appear in the equation. This means $(S + \alpha)$ will also be the solution of eq. (3) where α is any additive constant. Out of $(n+1)$ constants of integration one should therefore be an additive constant to S but since this additive constant will have no effect in transformation, we can take S to be involving only n constants $(\alpha_1, \alpha_2, ...\alpha_n)^*$. Hence a complete solution of eq. (3) can be written as

$$S = S(q_1, q_2, ...q_n, \alpha_1, \alpha_2, ...\alpha_n, t) \quad ...(4)$$

Where none of n constants is solely additive. We can take these n constants as the new momenta P_j i.e.

$$\alpha_j = P_j$$

Therefore n transformation equations can be written as

$$p_j = \frac{\partial S}{\partial q_j}(q_j, \alpha_j, t) \qquad ...(5)$$

Which are n equations and at $t = t_0$ give the n values of α_j in terms of the initial values of q and p. Thus constants of integration are evaluated in terms of the specific initial conditions of the problem i.e.

$$\begin{aligned} \alpha_j &= \alpha_j(q_j, p_j, t) \\ P_j &= P_j(q_j, p_j, t) \end{aligned} \qquad ...(6)$$

Which is one of the transformation equations.
The second transformation equations which provide the new constant coordinates appear as

$$Q_j = \beta_j = \frac{\partial S(q_j, \alpha_j, t)}{\partial P_j}$$

$$= \frac{\partial S(q_j, \alpha_j, t)}{\partial \alpha_j} \qquad ...(7)$$

Which are n equations and at $t = t_0$ give the n values of β_j interms of the known initial values of q_j. Eq. (7) can be 'turned inside out' to furnish q_j interms of α_j, β_j and t.

$$q_j = q_j(\beta_j, \alpha_j, t) \qquad ...(8)$$

Which solves the problem by giving the coordinates as function of time and initial conditions.

7.3 Discuss about the Hamilton's principal Function which is the solution of Hamilton-Jacobi Equation

From the solution of $H - J$ equation, we recognize S, the Hamilton's principal function as the generating function which gives rise to a canonical transformation involving constant momenta and constant coordinates. As we have explained in earlier steps of this method. With the achievement of transformation equation, solution of mechanical problem is obtained.

Choice of α_j as the new momenta is arbitrary to some extent, we can choose n quantities γ_j which are independent functions of α_j, the constants of integration $\gamma_j = \gamma_j(\alpha_1, \alpha_2, ...\alpha_n)$. This choice expresses Hamilton's principal function $S(q_j, \gamma_j, t)$.

To have more information about the significance of the function S, we consider its total time derivative. i.e.

$$S = S(q_j, \alpha_j, t)$$

$$\frac{dS}{dt} = \sum_j \frac{\partial S}{\partial q_j} \cdot \dot{q}_j + \sum_j \frac{\partial S}{\partial \alpha_j} \cdot \dot{\alpha}_j + \frac{\partial S}{\partial t}$$

$$...(9)$$

From eq. (5) $\quad \frac{\partial S}{\partial q_j} = p_j$

From eq. (3) $H + \frac{\partial S}{\partial t} = 0$

$$\frac{\partial S}{\partial t} = -H$$

Also $\dot{\alpha}_j = 0$, since α_j is constant.

Putting these in eq. (9)

$$\frac{dS}{dt} = \sum_i p_j \dot{q}_j - H$$

$$= L$$

$$= \int L \, dt + a \quad \text{constant} \quad ...(10)$$

The expression differs from Hamilton's principle in a constant showing that this time integral is of indefinite form. Thus the same integral when in definite form shapes the Hamilton's principle which results in the solution of mechanical problems, while in indefinite form it shapes the Hamilton's principle function that provides an alternative way of solving the problems.

7.4 Hamilton – Jacobi Equation for Hamilton's Characteristic Function

Conservative Systems :

For conservative systems in which H does not depend on time t explicitly, we have H as constant, α, is the total energy E of the system. Hamilton Jacobi equation for Hamilton's principal function $S(q, p, t)$ becomes

$$H\left(q_j \cdot \frac{\partial S}{\partial q_j}\right) + \frac{\partial S}{\partial t} = 0 \qquad ...(1)$$

Since the explicit dependence of S on t is involved in the last term, we can separate the variables. Let us assume the solution of the form

$$S(q_j, \alpha_j, t) = \omega(q_j, \alpha_j) - \alpha_1 t \qquad ...(2)$$

from which it follows that

$$\frac{\partial S}{\partial q_j} = \frac{\partial \omega}{\partial q_j} \quad \text{and} \quad \frac{\partial S}{\partial t} = -\alpha_1.$$

Putting these in eq. (1), we find that

$$H\left(q_j \cdot \frac{\partial \omega}{\partial q_j}\right) = \alpha_1 \qquad ...(3)$$

Which is independent of time. Out of α_j integration constants, one constant α_1 is a constant of motion being equal to H.

7.5 Physical significance of Hamilton's Characteristic function ω

We write it as $\omega = \omega(q_j, \alpha_j)$

/

$$\frac{d\omega}{dt} = \sum_j \frac{\partial \omega}{\partial q_j} \dot{q}_j + \sum_j \frac{\partial \omega}{\partial \alpha_j} \frac{d\alpha_j}{dt}$$

$$= \sum_j \frac{\partial \omega}{\partial q_j} \dot{q}_j$$

Since α_j's are constant.

$$\frac{\partial \omega}{\partial q_j} = \frac{\partial S}{\partial q_j} = p_j$$

$$\frac{d\omega}{dt} = \sum_j p_j \dot{q}_j$$

$$\omega = \int \left(\sum_j p_j \dot{q}_j \right) dt \quad = \text{action } A$$

Thus ω is identified as action A. Function ω is called Hamilton's characteristic function.

7.6 Kepler's problem (Solution by Hamilton – Jacobi Method)

The problem in general consists in finding the path of a charged particle under the action of a central force. To be more close to the Kepler's problem we choose the motion in an inverse square field. Let us suppose that a particle with charge e is moving about a stationary nucleus with positive charge ze, where z is atomic number. Denoting the conjugate momentum along γ and θ directions respectively by P_γ and P_θ, Hamiltonian function becomes

$$H = \frac{1}{2m}\left(P_\gamma^2 + \frac{P_\theta^2}{r^2} \right) - \frac{e^2 z}{r} \quad \dots(1)$$

As the system is conservative, Hamiltonian will represent the total energy α_1 of the system

For such system $H = \alpha_1 = E$ (say) $\quad \dots(1a)$

$$P_\gamma^2 + \frac{P_\theta^2}{\gamma^2} = 2ME + \frac{2me^2 z}{r} \quad \dots(2)$$

Since no explicit dependence on t is involved in H we can separate the variables. Further for such cases we shall proceed with Hamilton's characteristic function ω taking as the generating function for the transformation.

$$P_i = \frac{\partial S}{\partial q_i} = \frac{\partial \omega}{\partial q_i}$$

$$P_\gamma = \frac{\partial S}{\partial r} = \frac{\partial \omega}{\partial r}$$

$$P_\theta = \frac{\partial S}{\partial \theta} = \frac{\partial \omega}{\partial \theta}$$

Eq. (2) is $\left(\frac{\partial \omega}{\partial r} \right)^2 + \frac{1}{r^2}\left(\frac{\partial \omega}{\partial \theta} \right)^2 = 2ME + \frac{2me^2 z}{r}$

Separating the variables $\omega = \omega_\gamma(\gamma) + \omega_\theta(\theta)$

$$\left(\frac{\partial \omega_\gamma}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial \omega_\theta}{\partial \theta}\right)^2 = 2mE + \frac{2me^2 z}{r} \quad ...(3)$$

Since in a central force motion angular momentum is conserved. $P_\theta = $ constant.

$$P_\theta = \frac{\partial \omega_\theta}{\partial \theta} = \text{constant}$$

$\omega_\theta = \alpha_2 \theta + $ constant of integration   ...(4)

By substituting eq. (3) becomes

$$\frac{\partial \omega_\gamma}{\partial r} = \sqrt{\left(2ME + \frac{2me^2 z}{r} - \frac{\alpha_2^2}{r^2}\right)}$$

$$\omega_\gamma = \int \sqrt{\left(2ME + \frac{2me^2 z}{r} - \frac{\alpha_2^2}{r^2}\right)} \, dr + a \quad \text{(where a is constant of integration)}$$

This gives the transformation function as

$$\omega = \omega_\theta + \omega_\gamma = \alpha_2 \theta + \int \left(2ME + \frac{2me^2 z}{r} - \frac{\alpha_2^2}{r^2}\right)^{1/2} dr + \text{constant of integration} \quad ...(5)$$

We know that $Q_1 = \frac{\partial \omega}{\partial \alpha_1} = t + \beta_1$

$$\frac{\partial \omega}{\partial \alpha_i} = \beta_i \text{ when } i \neq 1 \text{ hence } \frac{\partial \omega}{\partial \alpha_2} = \beta_2$$

From equation (5), on calculating β_2

$$\beta_2 = \frac{\partial \omega}{\partial \alpha_2} = \theta - \int \frac{\alpha_2 \, d\gamma}{r^2 \sqrt{\left(2ME + \frac{2me^2 z}{r} - \frac{\alpha_2^2}{r^2}\right)}} \quad ...(5a)$$

The integral is evaluated to give

$$\frac{1}{r} = \frac{Me^2 z}{\alpha_2^2} - \left(\frac{M^2 e^4 z^2}{\alpha_2^4} + \frac{2ME}{\alpha_2^2}\right)^{1/2} \cos(\theta - \beta_2^1) \quad ...(6)$$

Where $\beta_2^1 = \beta_2 + \pi/2$

$$\frac{1}{\gamma} = \frac{me^2 z}{\alpha_2^2}\left[1 - \left(1 + \frac{2E\alpha_2^2}{Me^4 z^2}\right)^{1/2} \cos(\theta - \beta_2^1)\right] \quad ...(6)$$

Eq. (6) is of the form

$$\frac{1}{\gamma} = \frac{1}{\epsilon l}\left[1 - \epsilon \cos(\theta - \beta_2^1)\right] \quad ...(7)$$

Which is the equation of a conic with eccentricity ϵ,

Where $\epsilon l = \frac{\alpha_2^2}{Me^2 z}$

$$\in = \sqrt{\left(1 + \frac{2E\alpha_2^2}{Me^4 z^2}\right)} \qquad \qquad ...(8)$$

If $E < 0$, $\in < 1$, path of the particle is an ellipse.

$E = 0, \in = 1$, path is parabola.

$E > 0, \in > 1$, path is a hyperbola.

In Kepler problem we are interested in elliptical path. From equation (7) the relation between the major axis of the conic and the energy when it is an ellipse can be expressed as

$$2a = \frac{2 \in l}{1 - \in^2}$$

Which on using relation (8) gives

$$E = \frac{-ze^2}{2a} \qquad \qquad ...(9)$$

7.7 Action (phase Integrals) and Angle variables:

We extend Hamilton − Jacobi procedure so as to adopt it particularly to the solution of problems in periodic motions. In this technique we do not choose α_j as the new momenta, but instead we use suitably defined constants J_i which form a set of n independent functions of the α_j's and are called action variables.

Suppose we have separated the variables in Hamilton-Jacobi partial differential equation obtaining

$$p_j = \frac{\partial \omega}{\partial q_j}\left(q_j, \alpha_1, \alpha_2, ..., \alpha_n\right) \qquad ...(1)$$

Which gives $p_j = p_j\left(q_j, \alpha_1, \alpha_2, ... \alpha_n\right)$

We shall define the phase integral (or) action variable conjugate to the coordinate q_j by the integral

$$J_i = \oint P_j \, dq_j \qquad \qquad ...(2)$$

Where the integration is to be carried even a complete period of oscillation or rotation cycle of q_j. In case q_j is a cyclic coordinate, its conjugate momentum p_j is constant and this integral is to be taken from 0 to 2π

$$J_i = 2\pi P_j \qquad ...(3) \quad \text{for all cyclic coordinates.}$$

If we compare eq. (2) with the action A

$$A = \int \Sigma p_j \, \dot{q}_j \, dt = \int \Sigma p_i \, dq_j$$

It becomes quite obvious why J_i is designated as action variable.

Eq. (2) can also be written as

$$J_i = \oint \frac{\partial \omega \left(q_j, \alpha_1, ... \alpha_n\right)}{dq_j} \, dq_j \qquad ...(3a)$$

since the definite integral evaluated here is not a function of q_j i.e. q_j simply appears here as a variable of integration. Each action variable J_i is function of $\alpha_1, ... \alpha_n$ constants of integration.

$$J_i = J_i\left(\alpha_1, ..., \alpha_n\right) \qquad \qquad ...(4)$$

We can solve above equations for α's getting $\alpha_j = \alpha_j\left(J_1, ... J_n\right)$

By means of these relations we can express the Hamilton's characteristic function ω as

$$\omega = \omega\left(q_1, ... q_n, J_1, ... J_n\right) \qquad ...(5)$$

While Hamiltonian appears as function of J_i's are called angle variables ω_j and are given by the expression

$$\omega_j = \frac{\partial \omega(q_j, \alpha_j)}{\partial J_i} \qquad ...(6)$$

Since J_i has the dimensions of angular momentum coordinate conjugate to it should be an angle and hence the name angle variable.

Now the equations of motion for the angle variables are

$$\dot{\omega}_j = \frac{\partial H}{\partial J_i}(J_1, ...J_n) = v_j(J_1, ...J_n) \quad ...(6a)$$

having a solution $\omega_j = v_j t + \beta_j \qquad ...(7)$

Where v_j's are a set of constant functions of the action variables. Eq. (7) shows that ω_j's are linear functions of time.

Eq. (6) and (7) can be combined and solved for q's to give

$$q_j = q_j(J_1, ...J_n, \beta_1, ...\beta_n, t) \quad ...(8)$$

The greatest advantage of action and angle variables lies in the fact that one can obtain the frequencies of periodic motion without finding a complete solution for the motion of the system as can be seen from the following.

Let us denote change in an angle variable ω_j with the completion of one cycle by q_i one of the coordinates by $\Delta \omega_j$ given as

$$\Delta \omega_j = \oint \frac{\partial \omega_j}{\partial q_i} dq_i \qquad ...(9)$$

Putting $\omega_j = \frac{\partial \omega}{\partial J_i}$

$$\Delta \omega_j = \oint \frac{\partial^2 \omega}{\partial q_i \, \partial J_i} dq_i$$

$$= \frac{\partial}{\partial J_i} \oint \frac{\partial \omega}{\partial q_i} dq_i = \frac{\partial}{\partial J_i} \oint P_i \, dq_i \quad ...(10)$$

On using transformation equation

Thus $\Delta \omega_j = \frac{\partial}{\partial J_i} J_i = \delta_{ij}$

$\qquad = 1$ when $i = j$

$\qquad = 0$ when $i \neq j$

The above expressions shows that only when q_j (as $i = j$) goes through a complete period ω_j changes by unity. Then from equation (7)

$$\Delta \omega_j = v_j \Delta t$$

As we have just stated $\Delta \omega_j = 1$ when q_j goes through a complete period i.e. when $\Delta t = \tau_j$ where τ_j is the period associated with q_j.

Thus $1 = v_j \tau_j$

$$v_j = \frac{1}{\tau_j} = \omega_j$$

From above expression v_j is to be identified as the frequency associated with the periodic motion of q_j and hence we calculate the frequency of periodic motion without finding a complete solution for the motion of the system.

7.3 Solution of Harmonic oscillator problem by Action Angle variable method

For such a problem there is only one action variable

$$J = \oint p\, dq = \oint \frac{\partial \omega}{\partial q}(q, \alpha)\, dq \qquad \ldots(1)$$

We know that $\dfrac{\partial \omega(q, \alpha)}{\partial q} = \sqrt{mk}\sqrt{\dfrac{2\alpha}{k} - q^2}$

Then from eq. (1)

$$J = \sqrt{mk}\oint\sqrt{\left(\frac{2\alpha}{k} - q^2\right)}\, dq \qquad \ldots(2)$$

Substituting in eq. (2) for the periodic variation of coordinate

$$q = \sqrt{\left(\frac{2\alpha}{k}\right)}\sin\theta$$

$$J = 2\alpha\sqrt{\frac{m}{k}}\int_0^{2\pi}\cos^2\theta\, d\theta$$

Where limits are taken over a complete cycle in q. Solving the above integral

$$J = 2\pi\alpha\sqrt{\frac{m}{k}} \qquad \ldots(3)$$

$$\alpha = \frac{J}{2\pi}\sqrt{\frac{k}{m}}$$

Since the system is conservative. H is constant. Let us choose it to be α. Then

$$H = \alpha = \frac{J}{2\pi}\sqrt{\frac{k}{m}}$$

The frequency of oscillation is

$$\omega = \nu = \frac{1}{\tau} = \frac{\partial H}{\partial J} = \frac{1}{2\pi}\sqrt{\frac{k}{m}}$$

Which is the formula for the frequency of harmonic oscillator.

7.9 Separation of variables in the Hamilton-Jacobi Equation

Under certain conditions to solve the partial differential Hamilton-Jacobi equation, it is possible to separate the variables in the Hamilton-Jacobi equation. $H - J$ technique is a useful tool only when such a separation can be effected.

A coordinate q_j is said to be separable in the $H - J$ equation when Hamilton's principal function can be split into two additive parts, one of which depends only on the coordinate q_j and the other is entirely independent of q_j. If q_1 is taken as a separable coordinate, then the Hamiltonian must be such that

$$S(q_1,\ldots\ldots,q_n : \alpha_1,\ldots\ldots\alpha_n : t) = S_1(q_1 : \alpha_1,\ldots\ldots\alpha_n : t) + S^1(q_2,\ldots\ldots,q_n,\alpha_1,\ldots\ldots\alpha_n, t) \ldots(1)$$

and the $H - J$ equation can be split into two equations one separately for S_1 and the other for S^1. Similarly the $H - J$ equation is described as completely separable.

A solution for Hamilton's principal function of the form

$$S = \sum_i Si(q_i : \alpha_1,\ldots\ldots\alpha_n : t) \qquad \ldots(2)$$

will then split the $H - J$ equation into n equations of the type

$$H_i\left(q_i : \frac{\partial S_i}{\partial q_i} : \alpha_1, ..., \alpha_n\right) = \alpha_i \quad ...(3)$$

The constants α_i are referred to now as the separation constants. In eq. (3) involves only one of the coordinates q_i and the corresponding partial derivative of S_i with respect to q_i. It is possible to reduce them to quadratures. One has to solve for the partial derivative of S_i with respect to q_i and then integrate over q_i. The transition from Hamilton's principal function S to the characteristic function for conservative mechanical systems can be treated as an instance where t is a separable variable in $H - J$ equation.

$$S(q, \alpha, t) = S_0(\alpha, t) + \omega(q, \alpha) \quad ...(4)$$

Under the assumption, H is not an explicit function of time, the $H - J$ equation with this trial solution

$$H\left(q, \frac{\partial \omega}{\partial q}\right) + \frac{\partial S_0}{\partial t} = 0 \quad ...(5)$$

The first term is independent of t and the second term can be a function of t. Hence the equations can hold only if the two terms are both constant with equal and opposite values

$$\frac{\partial S_0}{\partial t} = -\alpha_1 \quad ...(6) \qquad H\left(q, \frac{\partial \omega}{\partial q}\right) = \alpha_1 \quad ...(7)$$

The first equation is solved by $S_0 = -\alpha_1 t$, the second eq. is the $H - J$ equation for ω.

$$H\left(q_2, ..., q_n : \gamma : \frac{\partial \omega}{\partial q_2}, ..., \frac{\partial \omega}{\partial q_n}\right) = \alpha_1 \quad ...(8)$$

$$\omega = \omega_1(q_1, \alpha) + \omega^1(q_2, ..., q_n, \alpha) \quad ...(9)$$

$$P_1 = \gamma = \frac{\partial \omega_1}{\partial q_1} \quad ...(10)$$

The constant γ is thus the separation constant and the obvious solution for ω_1 is

$$\omega_1 = \gamma q_1$$

$$\omega = \omega^1 + \gamma q_1 \quad ...(11)$$

The separated form for ω then appears as

$$\omega = \sum_{i=1}^{n} \omega_i(q_i, \alpha) = \omega_1(q_1, \alpha) + \sum_{i=2}^{n} \alpha_i q_i \quad ...(12)$$

Here ω_1 is the solution of the reduced $H - J$ equation

$$H\left(q_1 : \frac{\partial \omega_1}{\partial q_1} : \alpha_2, ..., \alpha_n\right) = \alpha_1 \quad ...(13)$$

This is an ordinary first order differential equation in the independent variable q_1, it can be immediately reduced to quadratures, and the complete solution for ω can be obtained

$$\omega = \omega_j(q_j, \alpha) + \omega^1(q_i, \alpha)$$

Where q_i represents the set of all q's except q_j then $H - J$ eq. appears

$$H\left(q_i, \frac{\partial \omega^1}{\partial q_i}, f\left(q_j, \frac{\partial \omega_j}{\partial q_j}\right)\right) = \alpha_1 \quad ...(14)$$

$$f\left(q_j, \frac{\partial \omega_j}{\partial q_j}\right) = g\left(q_i, \frac{\partial \omega^1}{\partial q_i}, \alpha_1\right) \quad ...(15)$$

$$\left. \begin{array}{c} f\left(q_j, \dfrac{\partial \omega_j}{\partial q_j}\right) = \alpha_j \\[3mm] g\left(q_i, \dfrac{\partial \omega^1}{\partial q_i}\right) = \alpha_j \end{array} \right\} \qquad \ldots(16)$$

and the separation of the variable has been accomplished.

The separability of the $H - J$ equation depends not only on the physical, problem involved but also the; system of generalized coordinates employed. Thus the one body central force problem is separable in polar coordinates, but not in cartesian coordinates.

Summary: We have explained Hamilton – Jacobi partial differential equation and derived the solution of Hamilton – Jacobi equation. Hamilton – Jacobi equation for Hamilton 's characteristic function and the physical significance of Hamilton 's characteristic function are explained The solution of Kepler's problem was derived by Hamilton – Jacobi method .The concept of Action and angle variables was explained Solution of harmonic oscillator equation was derived by action angle variable method

Key words: Hamilton – Jacobi partial differential equation – Hamilton 's characteristic function. – Kepler's problem - Action and angle variables – Separation of variables in Hamilton – Jacobi equation.

Self- Assessment questions

1. Write down the Hamilton Jacobi equation for Hamilton's principal function.
2. Give an account of the Hamilton-Jacobi theory and illustrate it by applying it to the Kepler's problem
3. What do you understand by Hamilton's characteristic function?
4. Explain Action angle variables.

Reference Books:
1. Mechanics: Symon
2. Classical mechanics: H. Goldstein.
3. Classical mechanics: Gupta, Kumar and Sharma

Classical Mechanics Part
UNIT-II
Lesson 8

Small Oscillations

Objective:

1. To obtain Lagrange's equation of motion for small oscillations of a system.
2. Expressing vibrational problems in terms of normal coordinates
3. Obtaining normal frequencies and illustrating how the theory can be applied to explain the free vibrations of a linear tri atomic molecule.
4. Vibrations in systems with external forces.

Structure:

8.1 Principle axes transformation
8.2 Normal coordinates and normal frequencies
8.3 Free vibrations of a linear tri atomic molecule (I case, II case)
8.4 Forced vibrations and the effect of dissipative forces

Introduction : The theory of small oscillations finds wide spread physical applications in acoustics, molecular spectra, vibrations of mechanisms, and coupled electrical circuits. If the deviations of the system from stable equilibrium conditions are small enough, the motion can generally be described as that of a system of coupled linear harmonic oscillators.

Equation of motion for a vibrating system :

Let us consider a conservative system in which the potential energy is a function of position only.

A system of particles is said to be in stable equilibrium if all the particles are and remain at rest. In the conservative force field, therefore, generalized forces acting on each particle must vanish.

$$Q_j = -\left(\frac{\partial V}{\partial q_j}\right)_0 \qquad \ldots(1)$$

This equation yields values q_{0j} of the generalized coordinates that the particles have in equilibrium configuration. If the equilibrium to be stable, the potential energy must be a minimum when evaluated at these q_{0j}. The displacements of the generalized co-ordinates from their equilibrium value will be denoted by u_j and so.

$$q_j = q_{0j} + u_j \qquad \ldots(2)$$

Since q_{0j} are fixed, we used the 'u_j' as generalized co-ordinates for the system.

The potential energy about the equilibrium by Taylor's series,

$$V(q_1, q_2, \ldots, q_n) = V(u_1, u_2, \ldots, u_n)$$

$$= V(q_{01}, q_{02}, \ldots, q_{0n}) + \sum_j \left(\frac{\partial V}{\partial q_j}\right)_0 u_j + \frac{1}{2}\sum_{j,k}\left(\frac{\partial^2 V}{\partial q_j\, \partial q_k}\right)_0 u_j u_k + \ldots\ldots\ldots(3)$$

$\therefore$ Eqn. (13) may also be written as

$$\left(Va - \omega^2 Ta\right) = 0 \qquad ...(17)$$

where 'a' is now a column matrix of 'n' components. $a = \begin{bmatrix} a_1 \\ a_2 \\ .... \\ a_n \end{bmatrix}$

8.1 PRINCIPAL AXES TRANSFORMATION

The symmetric nature of 'I' bears an important advantage. We can choose a certain system of body axes with respect to which the off-diagonal elements should disappear and only the diagonal elements remain in the expression for I. Such axes are called the principal axes of the body and the corresponding moments of Inertia as the principal moments of Inertia.

If we denote this form of Moment of Inertia by I^1 and I_1, I_2 and I_3 stand for principal values

$$I^1 = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \qquad ...(1)$$

But we know that $\vec{L} = I\vec{\omega}$

$$\begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix} = \begin{pmatrix} I_{xx} & I_{xy} & Ixz \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix}$$

$\therefore$ from eqn. (1) $\begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix} = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} = \begin{pmatrix} I_1\omega_x \\ I_2\omega_y \\ I_3\omega_z \end{pmatrix}$

or $L_x = I_1\omega_x, L_y = I_2\omega_y, L_z = I_3\omega_z \qquad ...(2)$

i.e. each of angular momentum component along a principal axes is a function of corresponding angular velocity only related to it via the principal moment of Inertia about that direction.

If 'I' is not symmetric such a procedure leads us into the realm of diagonalisation of matrix, a process which only means that by a suitable transformation of axes we shall kick out the off-diagonal elements leaving only the diagonal ones. The diagonal form of inertia tensor in which I_1, I_2, I_3 appear as elements is very easy to obtain.

Now we solve the following determinant equation which will be cubic in 'I' and therefore will furnish three values for I, Viz, I_1, I_2, I_3 which are desired principal moments of Inertia.

$$\begin{vmatrix} I_{xx} - I & I_{xy} & I_{xz} \\ I_{xy} & I_{yy} - I & I_{yz} \\ I_{xz} & I_{yz} & I_{zz} - I \end{vmatrix} = 0 \qquad ...(3)$$

This eqn. (3) is called the secular equation of Inertia tensor and its solutions, the secular values (or) eigen values.

If the symmetry axis of the body is taken as axes of rotation and the origin of body axes lies on this, then the principal axes are the symmetry axis and any two perpendicular axes in the plane normal to the symmetry axis. In such cases, we shall have two equal roots of eqn. (3).

In case of a sphere, every axis through center is symmetry axis and therefore any three orthogonal axes through the center are principal axes.

8.2 NORMAL COORDINATES AND NORMAL FREQUENCIES OF VIBRATION

When a solution of equations of motion contain only one single frequency (normal) the coordinate appearing in that solution is called the Normal coordinate. Therefore, the generalized coordinates, each one of them executing oscillations of one single frequency, are called Normal coordinates.

Consider $A^T M A = (q_1 q_2) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \end{pmatrix}$

$\therefore \quad A^T M A = q q_1^2 + 2b q_1 q_2 + c q_2^2$

where $A = \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} \rightarrow$ a column matrix

$M = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \rightarrow$ Matrix

$A^T = (q_1 \quad q_2) \rightarrow$ Transpose Matrix of 'A'

The transformation of equations of motion into such coordinates (normal coordinates), denoted by η and therefore of Lagrangian is effected by means of linear transformation.

We define the new set of coordinates η_k related to the original co-ordinates u_j by the defining equations

$$u_j = \sum_k a_{jk} \eta_k \qquad \qquad ...(1)$$

If we express $u_j's$ and $\eta_k's$ as the elements of the single column matrix (u) and 'n' respectively then, we can write eqn. (1) as $\qquad u = an \qquad ...(2)$

In order to find out Lagrangian in new-coordinates, we shall first express kinetic and potential energy in terms of η.

The potential energy is $V = \frac{1}{2} \sum_{j,k} V_{jk} u_j u_k$

$\therefore \qquad V = \frac{1}{2} \sum_{j,k} u_j V_{jk} u_k$

Which is quadratic in 'u' and therefore can be expressed as

$V = \frac{1}{2} u^T V u$

Putting for 'u' from eqn. (2), we get

$V = \frac{1}{2} (a\eta)^T V a\eta$

$= \frac{1}{2} \eta^T a^T V a \eta$

$= \frac{1}{2} \eta^T \wedge \eta \qquad (\because a^T V a = \wedge)$

Further above equation is quadratic in 'η' so that

$V = \frac{1}{2} \sum_i \lambda_i \eta_i^2$

$V = \frac{1}{2} \sum_i \omega_i^2 \eta_i^2 \qquad (\because \lambda_i = \omega_i^2)$

The kinetic energy is given by

$$T = \frac{1}{2} \sum_j \sum_k T_{jk} \dot{u}_j \dot{u}_k \qquad \ldots(3)$$

$$= \frac{1}{2} \sum_j \sum_k \dot{u}_j T_{jk} \dot{u}_k \qquad = \frac{1}{2} \dot{u}^T T \dot{u}$$

$$= \frac{1}{2} \dot{\eta}^T a^T T a \dot{\eta}$$

$$= \frac{1}{2} \dot{\eta}^T \dot{\eta} \qquad \left(\because a^T T a = 1 \right)$$

$$= \frac{1}{2} \sum_l \dot{\eta}_l^2$$

Therefore, the Lagrangian in New coordinate system will be

$$L = \frac{1}{2} \sum_{l=1}^{n} \dot{\eta}_l^2 - \frac{1}{2} \sum_{l=1}^{n} \omega_l^2 \eta_l^2 \qquad \ldots(4)$$

giving $\dfrac{\partial L}{\partial \dot{\eta}_l} = \sum_{l=1}^{n} \dot{\eta}_l \quad \dfrac{\partial L}{\partial \eta_l} = -\sum_{l=1}^{n} \omega_l^2 \eta_l$

which when substituted in

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\eta}_i} \right) - \frac{\partial L}{\partial \eta_l} = 0$$

gives $\sum_{l=1} \left(\ddot{\eta}_l + \omega_l^2 \eta_l \right) = 0$

Therefore the equations of motion in new coordinates are

$$\left. \begin{array}{l} \ddot{\eta}_1 + \omega_1^2 \eta_1 = 0 \\ \ddot{\eta}_2 + \omega_2^2 \eta_2 = 0 \\ \ldots\ldots \ldots\ldots \ldots\ldots \ldots\ldots \\ \ldots\ldots \ldots\ldots \ldots\ldots \ldots\ldots \\ \ddot{\eta}_l + \omega_l^2 \eta_l = 0 \end{array} \right\} \qquad \ldots(5)$$

In the above eqn. (5), the coordinate η_1 corresponds to ω_1, η_2 to ω_2 etc. Thus each coordinate executes only one single frequency oscillation and therefore η_1, η_2 etc are called as Normal coordinates.

The solution of equation

$$\ddot{\eta}_l + \omega_l^2 \eta_l = 0$$

will be

$$\left. \begin{array}{ll} \eta_l = A_l \cos\omega_l t + B_l \sin\omega_l t & \text{if} \quad \omega_l^2 > 0 \\ \eta_l = A_l t + B_l & \text{if} \quad \omega_l^2 = 0 \\ \eta_l = A_l e^{\omega_l t} + B_l e^{-\omega_l t} & \text{if} \quad \omega_l^2 < 0 \end{array} \right\} \qquad \ldots(6)$$

In the first case since ω_l^2 are real and positive. All coordinates always finite for any time 't' and we then say that the equilibrium is stable for the case $\omega_l^2 > 0$. But for the rest of the two cases, we find that the coordinates become infinite as the time advances and consequently such solution refers to unstable equilibrium.

Normal modes of vibration:

$$\omega = 2\pi \upsilon$$

The solution become

$$\eta_1 = A_1 \cos 2\pi \upsilon_1 t + B_1 \sin 2\pi \upsilon_1 t$$

$$\eta_2 = A_2 \cos 2\pi \upsilon_2 t + B_2 \sin 2\pi \upsilon_2 t$$

....(7)

....

$$\eta_n = A_n \cos 2\pi \upsilon_n t + B_n \sin 2\pi \upsilon_n t$$

where A's and B's are 2n arbitrary constants to be determined from the initial conditions.

Suppose we choose these coordinates such that all except A_1 and B_1 are zero. Then only coordinate η_1 will vary periodically with time, while the other coordinates will remain zero for all times. Such a situation corresponds to a normal mode of vibration i.e. the system is vibrating in a normal mode. Obviously, any of the coordinates can be given liberty to vary with time by such a choice of coordinates and therefore, there will be 'n' normal modes of vibration and 'n' normal frequencies $\upsilon_1, \upsilon_2, \upsilon_3 ...... \upsilon_n$ corresponding to each normal coordinate $\eta_1, \eta_2, \eta_3 ..... \eta_n$

We write eqn. (7) as

$$\eta_l = A_l \cos(\omega_l t + \delta_l)$$

where δ_l is the phase factor. Then from eqn. (1) the old coordinates are given by

$$u_j = \sum_k a_{jk} A_k \cos(\omega_k t + \delta_k) \quad ...(8)$$

8.3 FREE VIBRATIONS OF A LINEAR TRIATOMIC MOLECULE

Let us consider a linear triatomic symmetric molecule of the type yx_2 (e.g. CO_2) shown in fig1. where 'y' is the central atom and we shall neglect the interaction between this and end atoms x. We further assume that there exists on elastic bond between the central atom and the end atoms of force constant 'k'. Let the mass of each end atom be 'm' and that of central atom be 'M'.
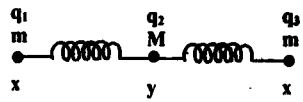
Fig 8.1: YX₂ molecule.

Let us denote the displacements of atoms from equilibrium configuration by the generalized coordinates q_1, q_2 and q_3. Then

Kinetic energy $T = \frac{1}{2} m(\dot{q}_1^2 + \dot{q}_3^2) + \frac{1}{2} M \dot{q}_2^2$

$$2T = (\dot{q}_1 \ \dot{q}_2 \ \dot{q}_3) \begin{pmatrix} m & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & m \end{pmatrix} \begin{pmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \end{pmatrix}$$

given $T = (T_{ij}) = \begin{pmatrix} m & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & m \end{pmatrix}$ ...(1)

Potential energy $V = \frac{1}{2} k(q_2 - q_1)^2 + \frac{1}{2} k(q_3 - q_2)^2$

$\therefore \quad V = \frac{1}{2} k \left(q_1^2 + 2q_2^2 + q_3^3 - 2q_1 q_2 - 2q_2 q_3 \right)$

or $\quad 2V = (q_1 \quad q_2 \quad q_3) \begin{pmatrix} k & -k & 0 \\ -k & 2k & -k \\ 0 & -k & k \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix}$

$V = (V_{ij}) = \begin{pmatrix} k & -k & 0 \\ -k & 2k & -k \\ 0 & -k & k \end{pmatrix} \qquad ...(2)$

Then we write secular equation

$\qquad | V - \omega^2 T | = 0.$

$\begin{vmatrix} k - \omega^2 m & -k & 0 \\ -k & k - \omega^2 M & -k \\ 0 & -k & k - \omega^2 m \end{vmatrix} = 0 \ ...(3)$

on developing this equation, we get

$\left(k - \omega^2 m \right) \left[\left(2k - \omega^2 M \right)\left(k - \omega^2 m \right) - k^2 \right] - k.k\left(k - \omega^2 m \right) - 0 = 0$

$\left(k - \omega^2 m \right)\left[\left(2k - \omega^2 M \right)\left(k - \omega^2 m \right) - 2k^2 \right] = 0$

$\omega^2 \left(\omega^2 m - k \right)\left[k\left(M + 2m \right) - \omega^2 mM \right] = 0$

The cubic equation gives three values of 'ω'.

Normal frequencies:

$\begin{cases} \omega_1 = 0 \\[2mm] \omega_2 = \sqrt{\dfrac{k}{m}} \\[3mm] \omega_3 = \left\{ \dfrac{k}{m}\left(1 + \dfrac{2m}{M} \right) \right\}^{\frac{1}{2}} \end{cases} \qquad ...(4)$

The first case refers to translatory motion $(\omega_1 = 0)$ of the atoms and the rest two to the oscillatory motion. In order to calculate the corresponding normal coordinates η_1, η_2 and η_3 we shall proceed with

$\qquad q_j = \sum_k a_{jk} \, \eta_k$

$\therefore \quad \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix} \quad ...(5)$

Therefore our aim is to calculate the components of eigen vectors a_1, a_2 and a_3.

First Case : $\omega_1 = 0$ and calculation of the components a_{11}, a_{21} and a_{31} of eigen vector a_1 :

For this purpose, we apply the relation

$\qquad \sum_j \left(V_{ij} - \omega^2 T_{ij} \right) a_j = 0 \qquad\qquad (i = 1,2,3)$

From eqn. (3), we write for a_1 as

$$\begin{pmatrix} k-\omega_1^2 m & -k & 0 \\ -k & 2k-\omega_1^2 M & -k \\ 0 & -k & k-\omega_1^2 m \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} = 0$$

substituting $\omega_1^2 = 0$

$$\begin{pmatrix} k & -k & 0 \\ -k & 2k & -k \\ 0 & -k & k \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} = 0$$

giving $ka_{11} - ka_{21} = 0$

$-ka_{11} + 2ka_{21} + ka_{31} = 0$

$-ka_{21} + ka_{31} = 0$

$a_{11} = a_{21} = a_{31} = \alpha$ (say)   ...(6)

$$\therefore \quad a_1 = \begin{pmatrix} \alpha \\ \alpha \\ \alpha \end{pmatrix}$$

which shows that displacements of all the atoms are in the same direction and equal in amount.

Second case : $\omega_2 = \sqrt{\dfrac{k}{m}}$ and calculate the components a_{12}, a_{22} and a_{32} of eigen vector a :

For this purpose, we can write

$$\begin{pmatrix} k-\omega_2^2 m & -k & 0 \\ -k & 2k-\omega_2^2 M & -k \\ 0 & -k & k-\omega_2^2 m \end{pmatrix} \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} = 0$$

Substituting $\omega_2 = \sqrt{\dfrac{k}{m}}$, we get

$a_{22} = 0$

$a_{12} = -a_{32} = \beta$ (say)

So that component of a_2 are

$$a_2 = \begin{pmatrix} \beta \\ 0 \\ -\beta \end{pmatrix} \qquad \qquad ...(7)$$

similarly for the third case $\omega_3 = \sqrt{\dfrac{k}{m}\left(1+\dfrac{2m}{M}\right)}$ and components of eigen vector a_3 are to be calculated putting

$\omega_3^2 = \dfrac{k}{m}\left(1+\dfrac{2m}{M}\right)$.

We get

$$\begin{pmatrix} -\dfrac{2mk}{M} & -k & 0 \\ -k & -k\dfrac{M}{m} & -k \\ 0 & -k & -\dfrac{2mk}{M} \end{pmatrix} \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix} = 0$$

Solving, $-\dfrac{2mk}{M}a_{13} - ka_{23} = 0$

$$-ka_{13} - \dfrac{kM}{m}a_{23} - ka_{33} = 0$$

$$-ka_{23} + \dfrac{2m}{M}ka_{33} = 0$$

on solving, these equations we get

$$a_{13} = a_{33} = \gamma$$

$$a_{23} = -\dfrac{2m}{M}a_{13} = -\dfrac{2m}{M}\gamma,$$

So that the components of a_3 will be

$$a_3 = \begin{pmatrix} \gamma \\ -\dfrac{2m}{M}\gamma \\ \gamma \end{pmatrix} \qquad \qquad ...(8)$$

Eqn. (6) (7) and (8) define the required components of eigen vectors. If α, β, γ are known for which we apply the following orthogonality.

$$a^T Ta = 1 \qquad \qquad ...(9)$$

Because $(a_{ij}) = \begin{pmatrix} \alpha & \beta & \gamma \\ \alpha & 0 & -\dfrac{2m}{M}\gamma \\ \alpha & -\beta & \gamma \end{pmatrix}$ and $T_{ij} = \begin{pmatrix} m & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & m \end{pmatrix}$

we write eqn. (9) as

$$\begin{pmatrix} \alpha & \alpha & \alpha \\ \beta & 0 & -\beta \\ \gamma & -\dfrac{2m}{M}\gamma & \gamma \end{pmatrix} \begin{pmatrix} m & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & m \end{pmatrix} \begin{pmatrix} \alpha & \beta & \gamma \\ \alpha & 0 & -\dfrac{2m}{M}\gamma \\ \alpha & -\beta & \gamma \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} \alpha & \alpha & \alpha \\ \beta & 0 & -\beta \\ \gamma & -\dfrac{2m}{M}\gamma & \gamma \end{pmatrix} \begin{pmatrix} \alpha m & \beta m & \gamma m \\ \alpha M & \beta M & -2\gamma m \\ \alpha m & -\beta m & \gamma m \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

(or) $\begin{pmatrix} \alpha^2(2m+M) & 0 & 0 \\ 0 & 2\beta^2 m & 0 \\ 0 & 0 & \gamma^2 2m\left(1+\dfrac{2m}{M}\right) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

$$\Rightarrow \quad \left.\begin{array}{l} \alpha = \dfrac{1}{(2m + M)^{\frac{1}{2}}} \\[4mm] \beta = \dfrac{1}{(2m)^{\frac{1}{2}}} \\[4mm] \gamma = \dfrac{1}{\left\{2m\left(1 + \dfrac{2m}{M}\right)\right\}^{\frac{1}{2}}} \end{array}\right\} \qquad \ldots(10)$$

Now using eqn. (6), (7) (8) and (10) with eq. (5), we can write for the normal coordinates η_1, η_2 and η_3 associated with normal frequencies ω_1, ω_2 and ω_3 respectively. i.e.

$$\begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} = \begin{pmatrix} \dfrac{1}{\sqrt{2m+M}} & \dfrac{1}{\sqrt{2m}} & \dfrac{1}{\sqrt{2m\left(1+\dfrac{2m}{M}\right)}} \\[6mm] \dfrac{1}{\sqrt{2m+M}} & 0 & \dfrac{-2m/M}{\sqrt{2m\left(1+\dfrac{2m}{M}\right)}} \\[6mm] \dfrac{1}{\sqrt{2m+M}} & \dfrac{-1}{\sqrt{2m}} & \dfrac{1}{\sqrt{2m\left(1+\dfrac{2m}{M}\right)}} \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix}$$
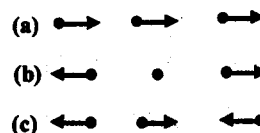
Which completes the discussion.

(a)

(b)

(c)

Fig 8.2: Vibrations of triatomic (linear) molecule

It can be seen that in case 2, with ω_2, we find that $a_{22} = 0$

$$a_{12} = -a_{32}$$

Which indicates that the central atom does not take part in motion and end atoms oscillate with equal amplitude but opposite in phase (see fig.2 (b)).

In case (3)

$$a_{13} = a_{33} = \gamma$$

$$a_{23} = -\frac{2m}{M}\gamma$$

Which indicates that end atoms vibrate in phase with equal amplitude, where as the central atom vibrates in opposite phase with a different amplitude. [see fig. 2c]

8.4 FORCED VIBRATIONS AND THE EFFECT OF DISSIPATIVE FORCES

Force vibrations occur when the system is displaced initially from the equilibrium configuration and is then allowed to oscillate by itself. The system is set into oscillation by an external driving force that continues to act on the system after $t = 0$. The frequency of such a forced oscillation is then determined by the frequency of the driving force and not by the resonant frequencies.

If F_j is the generalized force corresponding to the coordinate η_j, then the generalized force Q_i for the normal coordinate ζ_i is

$$Q_i = a_{ji} F_j \qquad \ldots(1)$$

The equations of motion when expressed in normal coordinates now become

$$\ddot{\zeta}_i + \omega_i^2 \zeta_i = Q_i \qquad \ldots(2)$$

Eq. (2) are a set of 'n' homogeneous differential equations that can be solved only when we know the dependence of Q_i on time. While the solution will not be as simple as in the free case, the normal coordinates preserve their advantage of separating the variables, and each equation involves only a single coordinates.

Frequently the driving force varies sinusoidally with time. Since eq. (2) are linear equations, its solutions for particular frequencies can be superposed to find the complete solution for given Q_i. It is therefore of general importance to study the nature of the oscillations when the force Q_i can be written as

$$Q_i = Q_{0i} \cos(\omega t + \delta_i) \qquad \ldots(3)$$

where 'ω' is the angular frequency of an external force.

The equations of motion now appear as

$$\ddot{\zeta}_i + \omega_i^2 \zeta_i = Q_{0i} \cos(\omega t + \delta_i) \qquad \ldots(4)$$

A complete solution of eq. (4) consists of the general solution to the homogeneous equation (i.e. the free modes of vibrations) place a particular solution to the inhomogeneous equation.

The solution of eq. (4) have the form

$$\zeta_i = B_i \cos(\omega t + \delta_i) \qquad \ldots(5)$$

Here the amplitudes B_i are determined by substituting the solution in equations.

$$B_i = \frac{Q_{ji}}{\omega_i^2 - \omega^2} \qquad \ldots(6)$$

The complete motion is then

$$\eta_j = a_{ji}\zeta_i = \frac{a_{ji}Q_{ji}\cos(\omega t + \delta_i)}{\omega_i^2 - \omega^2} \qquad \ldots(7)$$

Thus the vibration of each particle is again composed of linear combinations of the normal modes, but now each normal oscillation occurs at the frequency of the driving force.

Two factors determine the extent to which each normal mode is excited. One is the amplitude of the generalized driving force Q_{0i}. If the force on each particle has no component in the direction of vibration of some particulars normal mode, then obviously the generalized force corresponding to the mode will vanish and Q_{0i} will be zero. an external force can excite a normal mode only if it tends to move the particles in the same direction as in the given mode. The second factor is the closeness of the driving frequency to the free frequency of the mode.

In consequence of the denominators in eq. (6), the closer 'ω' approaches to any ω_i the stronger will that mode be excited relative to the other modes. Indeed eq. (6) apparently predicts infinite amplitude when the driving frequency agrees exactly with one of the ω_i s – The familiar phenomenon of Resonance.

In many physical systems these forces when present, are proportional to the particle velocities and can therefore be derived from a dissipation function f. Let us first consider the effects of frictional forces on the free modes of vibration.

From its definition, of must be a homogeneous quadratic function of the velocities

$$\mathfrak{I} = \frac{1}{2}\mathfrak{I}_{ij}\dot{\eta}_i\dot{\eta}_j \qquad \ldots(8)$$

The coefficients $\mathfrak{I}_{ij}$ are clearly symmetric $\mathfrak{I}_{ij} = \mathfrak{I}_{ji}$ and in general will be functions of the coordinates. In eq. (8) we shall take $\mathfrak{I}_{ij}$ as denoting these constant factors. It will be remembered that $2\mathfrak{I}$ is the rate of energy dissipation due to the frictional forces.

The complete set of Lagrange equations of motion now became

$$T_{ij}\ddot{\eta}_j + \mathfrak{I}_{ij}\dot{\eta}_j + V_{ij}\eta_j = 0 \qquad \ldots(9)$$

If the frictional force is proportional both to the particle's velocity and its mass, then $\mathfrak{J}$ will be diagonal whenever T is. When such simultaneous diagonalization is feasible, then the equations of motion are decoupled in the normal coordinates with the form

$$\ddot{\zeta}_i - \mathfrak{J}_i \dot{\zeta}_i + \omega_i^2 \zeta_i = 0 \qquad \text{(no summation)}$$
$$\qquad \qquad \dots(10)$$

Here the $\mathfrak{J}_i$ are the non-negative coefficients in the diagonalized form of $\mathfrak{J}$ when expressed in terms of ζ_i. Being a set of linear differential equations with constant coefficients, eq. (10), eq. (10) is solved by functions of the form

$$\zeta_i = C_i e^{-i\omega_i t}$$

where ω_i' satisfies the quadratic equation

$$\omega_i'^2 + i\omega_i' \mathfrak{J}_i - \omega_i^2 = 0 \qquad \text{(no summation)} \dots(11)$$

Eq. (10) has the two solutions

$$\omega_i' = \pm \sqrt{\omega_i^2 - \frac{\mathfrak{J}_i^2}{4}} - i\frac{\mathfrak{J}_i}{2} \qquad \dots(12)$$

The motion is therefore not a pure oscillation, for ω' is complex. It clear from eq. (12) that the imaginary part of ω_i' results in a factor $\exp -\frac{\mathfrak{J}_i t}{2}$ because of the non-negative nature of the $\mathfrak{J}_i$'s this is always an exponentially decreasing function of time. As the particle vibrate, they do work against the frictional forces, and the energy of the system must decrease with time. The real part of eq. (12) corresponds to the oscillatory factor in the motion, and it will noted that the presence of friction also effects the frequency of the vibration. If the dissipation is small, the squared term in $\mathfrak{J}_i$ is neglected and the frequency of oscillation reduces to the friction free value.

The complete motion is then simply an exponential damping of the free modes of vibration.

$$\zeta_i = C_i e^{-\mathfrak{J}_i t/2} e^{-i\omega_i t} \qquad \dots(13)$$

If the dissipation function cannot be diagonalized along with T and V, the solution is much more difficult to obtain.

Suppose seek a solution to eq. (8) of the form

$$\eta_j = Ca_j e^{-i\omega t} = Ca_j e^{-KT} e^{-2\pi i r t} \qquad \dots(14)$$

with this solution eq. (9) become a set of simultaneous linear equations.

$$V_{ij}a_j - i\omega \mathfrak{J}_{ij}a_j - \omega^2 T_{ij} a_j = 0 \qquad \dots(15)$$

It is convenient to write 'ω' as iv so that

$$v = -i\omega = -K - 2\pi i v \qquad \dots(16)$$

and thus '-k' is the real part of δ. In terms of square matrices of V, T and $\mathfrak{J}$ the set of equation (15) become a column matrix equation involving v:

$$Va + vFa + v^2 Ta = 0 \qquad \dots(17)$$

The set of homogeneous eq. (15) or (17) can be solved for the a_j only for certain values of ω (or) δ.

Convert the matrix equation (17) into a scalar equation for δ by multiplying from the left with a':

$$a' Va + v a' Fu + v^2 a' Ta = 0 \dots(18)$$

Eq. (17) is a quadratic equation for v with coefficients that are matrix products of the same general type as those encountered in equation.

$$\left(\lambda_k - \lambda_k^*\right) a_k^+ T a_k = 0$$

y virtue of the symmetry of V, F and T the matrix products are all real, and can be seen by expanding 'a' as $c + i\beta$. Hence if δ is a solution of the quadratic equation its complex conjugate δ' also be a solution. Now, the sum of the two roots of a quadratic equation is negative of the coefficient of the linear term divided by the coefficient of the square term.

$$v + v^k = -2k = -\frac{a + Fa}{a + Fa} \qquad ...(19)$$

Hence 'k' can be expressed in terms of the real and imaginary parts of a_j as

$$K = \frac{1}{2} \frac{\Im_{ij}(\alpha_i \alpha_j + \beta_i \beta_j)}{T_{ij}(\alpha_i \alpha_j + \beta_j \beta_i)} \qquad ...(20)$$

The dissipation function must always be positive, and 't' is positive definite, hence 'k' cannot be negative. The oscillations of the system may decrease exponentially with time, but they can never increase with time. The frequencies of oscillation, given by the real part of ω, be affected by the dissipative forces, but the change will be small if the damping is not very large during a period of oscillation.

Consider forced sinusoidal oscillations in the presence of dissipative forces. Representing the variation of the driving force with time by

$$F_j = F_{0j} e^{-i\omega t}$$

Where F_{0j} may be complex, the equations of motion are

$$V_{ij}\eta_j + \Im_{ij}\dot{\eta}_j + T_{ij}\ddot{\eta}_j = F_i e^{-i\omega t} \qquad ...(21)$$

Let the solution of these eq. (21) is of the form $\eta_j = A_j e^{-i\omega t}$

Then the following set of inhomogeneous linear equations for the amplitudes A_j are obtained.

$$(V_{ij} - i\omega \Im_{ij} - \omega^2 T_{ij})A_j - F_i = 0 \qquad ...(22)$$

The solution to these equations may easily be obtained from Cramer's rule.

$$A_j = D_j(\omega) / D(\omega) \qquad ...(23)$$

Where $D(\omega)$ is the determinant of the coefficients of A_j in eq. (22) and $D_j(\omega)$ is the modification in $D(\omega)$ resulting when the jth column is replaced by $F_{01} ....F_{0n}$

For a system of 'n' degrees of freedom it is therefore possible to represent $D(\omega)$ as

$$D(\omega) = G(\omega - \omega_1)(\omega - \omega_2).....(\omega - \omega_n)(\omega + \omega_1^*)(\omega + \omega_2^*).....(\omega + \omega_n^*)$$

Where 'G' is some constant.

Using product notation and denoting 'ω' by $2\pi \delta$, this representation can be written as

$$D(\omega) = G\prod_{i=1}^{n} (2\pi(v - v_i) + ik_i)(2\pi(v + v_i) + ik_i) \qquad ...(24)$$

When rationalize eq. (24) to separate A_i into its real and imaginary parts, the denominator will be

$$D^*(\omega)D(\omega) = GG^*\prod_{i=1}^{n}(4\pi^2(v - v_i)^2 + k_i^2)(4\pi^2(v + v_i)^2 + k_i^2) \qquad ...(25)$$

The amplitudes of the forced oscillation thus exhibit typical resonance behaviour in the neighborhood of the frequencies of free oscillations $+v_i$. As a result of the presence of the damping constants k_i, the resonance denominators no longer vanish at the free mode frequencies, and the amplitudes remain finite. The driving frequency at which the amplitude peaks is no longer exactly at the free frequencies because of frequency dependence of terms in 'A_j' other than the particular resonance denominator. If the damping is small enough to preserve a recognizable resonant peak, the shift in the resonance frequencies is usually small.

Summary : A vibrating system can be represented as a superposition of simple harmonic oscillators. Lagrange's equation of motion are suitable to represent for small oscillations of a system . However the problem can be best represented if represented in terms of normal coordinates. So, normal coordinates and normal frequencies are defined and free vibrations of a linear tri atomic molecule are explained

Key words: Principle axes transformation – Normal coordinates and normal frequencies – Free vibrations of a linear tri atomic molecule – Forced vibrations - Dissipative forces

Self-Assessment Questions
1.Obtain the Lagrange's equation of motion for small oscillations of a system.
2.Explain the oscillations of a linear tri atomic molecule.
3.Explain the effect of dissipative force.

Reference Books :
1.Mechanics: Simon
2.Classical mechanics: H. Goldstein.
3. Classical mechanics: Gupta, Kumar and Sharma
4.Analytical Dynamics: Whittaker.

Statistical Mechanics Part
Unit – III
Lesson 1

CLASSICAL STATISTICAL MECHANICS

Objectives:

1. To introduce the concept of statistical mechanics
2. To explain concept of phase space, μ space, Γ space, and an ensemble
3. To explain the three types of ensembles, like micro canonical, canonical and grand canonical ensembles
4. To explain the ensemble average, density of distribution in the phase space
5. To discuss the equipartition theorem.

Structure:

1.1 INTRODUCTION:

In statistical mechanics we study the physical systems consisting of very large number of particles ($N \sim 10^{23}$). The simplest physical system of interest is a perfect gas in thermal equilibrium. From the macroscopic point of view it appears to be a continuum. A complete set of thermodynamic variables, characterizing its equilibrium state, is the energy E, volume V and the number of molecules N. The N, although referred to molecules for convenience, is a macroscopic variable as it is directly related to the mass of the gas.

From the microscopic point of view the gas (matter) consists of discrete particles, like atoms or molecules. In classical mechanics, the microscopic description will specify at a given time, the positions and velocities (or momenta) of all the particles in the gas. It is impossible to measure them instantaneously as $N \sim 10^{23}$. With lapse of time, the description of the behavior of the gas will require solving an enormous number of equations of motion involving collisions. To get out of this impasse, we can try to relate the macroscopic description based on a few variables with the microscopic description based on a large number of variables, by using the method of (i) kinetic theory, or (ii) statistical mechanics.

The drastic reduction in the number of variables occurs because a measurement of the macroscopic property, like pressure P, gives an average of the values over a finite time interval ($\sim 1s$). During this period the molecules undergo a very large number of collisions as the time interval between two successive collisions is of the order of 10^{-10}s. The mathematical process of averaging over a coordinate obviously eliminates it, resulting in simplicity. For example, consider the kinetic theory calculation of pressure of a gas in a cubical box of side length L.

Let C_{xi} be the velocity component along the edge parallel to the x-axis for the ith molecule of mass m. The rate of momentum transfer to a wall normal to the $x-$axis is given by (number of collisions per s) x (Momentum imparted per collision) $= (C_{xi} / 2L)$ x $(2mC_{xi})$. Therefore, the pressure P exerted by the molecules on that wall is

$$P = \frac{\text{total force on the wall}}{\text{area of the wall}} = \frac{m(C_{x1}^2 + ---- + C_{xN}^2)}{L^3} = \frac{1}{L^3}mN\overline{C_x^2},$$

$$\text{Where } \overline{C_x^2} = \frac{1}{N}\sum_{i=1}^{N}C_{xi}^2$$

Where $\overline{C_x^2}$ is the average value of C_x^2 for all the molecules. Thus $N(\sim 10^{23})$ velocity coordinates are reduced to a single suitably averaged coordinate.

In kinetic theory certain basic assumptions are made regarding the nature of molecules and their mutual interactions. Statistical mechanics does not concern itself with such details as it deals mainly with the energy aspects of the molecules. It makes assumptions of a more general nature, uses the theory of probability and it is mathematically simpler. We shall discuss the method of statistical mechanics as first formulated by Josiah Willard Gibbs.

1.2 PHASE SPACE:

First consider a very simple case. A bead of mass m moves freely and arbitrarily a string stretched along the $x-$axis. It has one degree of freedom. The position of the bead at time 't' is $x(t)$ and its velocity $V_x =$

x (or momentum $p_x = m\dot{x}$) at that instant. The state of the bead at any instant can be represented by a point P in a hypothetical two-dimensional space, called the phase space, whose coordinates are x and p_x. As the bead moves on the string, the value of x changes. Under accelerating force, p_x also changes. As a result the point P traces a trajectory in the phase space with the passage of time (Fig 1.1).
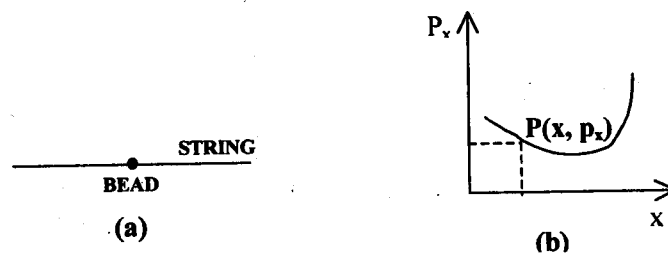
STRING

BEAD

(a)

P_x

$P(x, p_x)$

x

(b)

Fig 1.1(a) A bead sliding on a string (b) Phase space and phase line for the bead.

A molecule of an ideal gas can be represented as a structure-less particle. Such a molecule has three translational degrees of freedom. Its phase space has six dimensions whose Cartesian coordinates are x_1, x_2, x_3, p_1, p_2, p_3. It is called the μ - space, where μ stands for molecule. The instantaneous translational state of the molecule is given by the representative point in this hypothetical space. For a system of N molecules (gas) the instantaneous state (Fig.1.2a) is represented by a set of N points in the μ - space, one for each molecule (Fig.1.2b). It is a symbolic picture of the space because, it is not possible to display a six-dimensional space. The total number of translational degree of freedom is 3xN = 3N. Following Ehrenfest, we can construct a phase space, for all the molecules, which has 6N dimensions. It is called the Γ space, where Γ stands for gas. It is spanned by 3N coordinate axes and 3N momentum axes. The 6N coordinates

$(x_{11}, x_{21}, x_{31}, \ldots\ldots, x_{1N}, x_{2N}, x_{3N}, p_{11}, p_{21}, p_{31}, \ldots\ldots\ldots, p_{1N}, p_{2N}, p_{3N})$

represents the positions and momenta of all the molecules (state of the system) at a given time. In the Γ space, the instantaneous state of the whole system (gas of N molecules) is given by a single representative point (or phase point), Fig 1.2c. The notation [x], [p] stands for the 3N coordinate axes and 3N momentum axes.

z

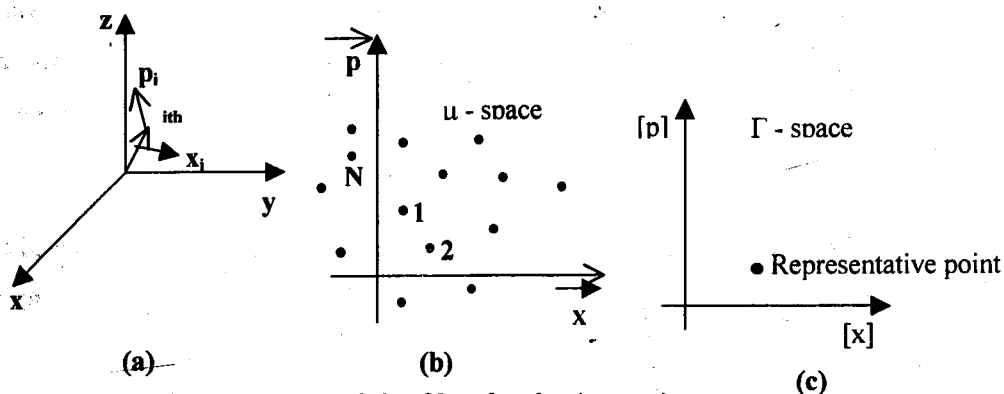p_i

ith

x_i

y

x

(a)

p

μ - space

N

•1

•2

x

(b)

[p]

Γ - space

• Representative point

[x]

(c)

Fig 1.2(a) A gas containing N molecules (system).
(b) μ-space for the system, (c) Γ-space and the representative point for the entire system.

In general, if f independent position coordinates and f moment coordinates are required to fully specify the state of a system, then the system is said to possess f degrees of freedom.

Any set of f generalized coordinates q_1, q_2, ...q_f (cartisian, polar or some other convenient set) can be used to uniquely determine the configuration of the system. The corresponding generalized momenta are p_1, p_2, ...p_f. The Γ space is then a conceptual Euclidean space having 2f rectangular axes [q], [p]. The microscopic state of the whole system is specified by a representative point in this space. With the lapse of time, some or all of the 2f coordinates take on different values (Fig 1.3). As a result, the representative point traces a phase line (or phase trajectory) in the accessible phase space (Fig 1.4). Each point on the phase line represents one such possible microscopic state. A point in the phase space is accessible if it corresponds to the physical specification of the system under observation. For example, the states of the crystalline form of sodium are inaccessible at very high temperature. The system is likely to pass through all the accessible states. In this sense, the 2f coordinates take on all possible values. We can say that they are randomized. The phase line tends to fill the accessible phase space. The measurement of macroscopic variables (like P,V,T etc) involves taking time averages over an appropriate portion of the phase line of the system.

EVN
(a)

EVN
(b)

EVN
(c)
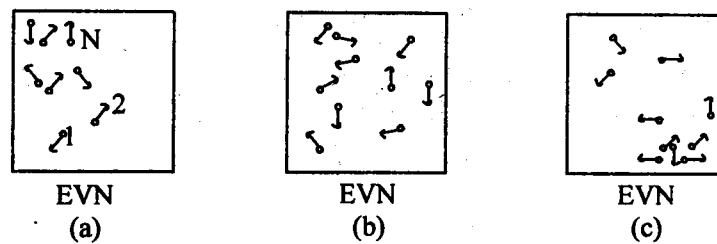
Fig 1.3 A few possible state of the system (gas containing N molecules). For convenience only few molecules are shown.
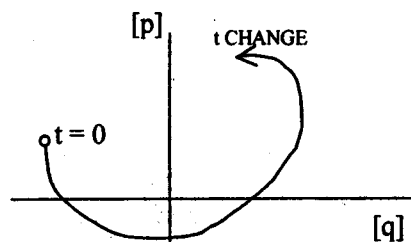

Fig 1.4 Phase space and a portion of the phase line.

So for we have not introduced any concept of statistical mechanics. The problem of solving an enormous number of equations of motion and of calculating the time averages of interest is still with us. Around 1900 Gibbs suggested that a way out is to introduce the idea of an ensemble of system.

1.3 ENSEMBLE:

Each phase point on the phase line of single system develops out of the previous point in time, according to the laws of mechanics. Gibbs replaced the time dependent picture by static picture in which the entire phase line exist at one time (Fig 1.5a). Then each phase point represents a separate system with the same macroscopic properties (E, V, N) as the system of interest but a different microscopic state. In other words, we 'imagine' a large number M(M→α) of system, similar in structure to the system of interest but suitable randomized in the accessible, unobservable, microscopic states. Instead of taking the time average, we take an average over this artificially constructed group existing simultaneously at one time. Such a group of replicas or collection of similar non-interacting, independent, imaginary systems is called an ensemble by Gibbs (Fig 1.5b). We have assumed that the time average of some property of a system in equilibrium is same as the instantaneous ensemble average. This is known as the ergodic hypothesis.

All the members of an ensemble, which are identical in features like N, V, E, are referred to as elements. These elements, though identical in structure (same macroscopic state), are randomized in the sense that they differ from one another in the coordinates and momenta of the individual molecules, that is, the elements differ in their unobservable microscopic states. The various elements, being imaginary, do not interact with each other. Each element behaves independently, in accordance with the laws of mechanics (classical or quantum).
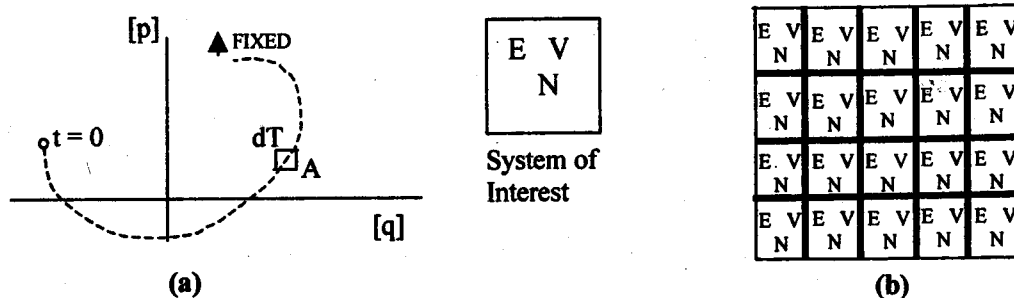
**Fig 1.5(a) & (b) a: The ensemble (a small portion) at one time
b: Schematic lattice representation of an ensemble**

A clear difference exists between the actual system of interest and an element of the ensemble. The system is physical object about which we intend to make predictions. The elements of ensemble are mental copies of it to enable us to use the probability theory.

Thus an ensemble of systems consists of randomized 'mental' pictures of the system of interest that exist simultaneously. It is to be viewed as an intellectual exercise to imitate and represent at one time the state of the actual system as developed in the course of time. It is easier to compute the statistical behavior of such a suitable chosen ensemble than to study the behavior of any particular complex system. Results so obtained enable us to predict the probable behavior of system of interest.

An ensemble average is the average at a fixed time over all the elements in an ensemble. It is difficult to prove the exact equivalence of the ensemble average and the time average over a single system. However, one can hope that the former would closely approximate the latter, if the following essential conditions are satisfied.

1. The system of interest is a macroscopic system consisting of a large number of molecules (N→∞) so that we can randomize in a true sense the microscopic variables.

2. The number of imagined elements that form the ensemble at one time is large (M→∞) so that they can truly represent the range of states available to the actual system over a really long period of time (t→∞). In statistical mechanics we shall use the terms system and ensemble in this above sense only.

In Fig 1.5a, each phase point corresponds to an element in the ensemble (Fig 1.5b). In an appropriate ensemble the phase points would be distributed continuously.

An ensemble is also defined as a collection of a very large number of assemblies which are essentially independent (i.e., in calculating the possible eigen-state of an ensemble we do not have to worry about any interaction between the assembly of interest and any of the other assemblies) of one another but which have been made macroscopically as identical as possible. By being macroscopically identical, we mean that each assembly is characterized by the same values of set of macroscopic parameters, which uniquely determine the equilibrium state of the assembly.

1.3.1 Micro-canonical ensemble:

The micro-canonical ensemble is a collection of essentially independent assemblies having the same energy E, volume V and number N of systems, where for simplicity we assume that we have only one type of system. The individual assemblies are separated by rigid, impermeable, and well-insulated walls (Fig 1.6). We cannot actually specify the macroscopic energy of an assembly exactly.

E,V,N	E,V,N	E,V,N
E,V,N	E,V,N	E,V,N
E,V,N	E,V,N	E,V,N
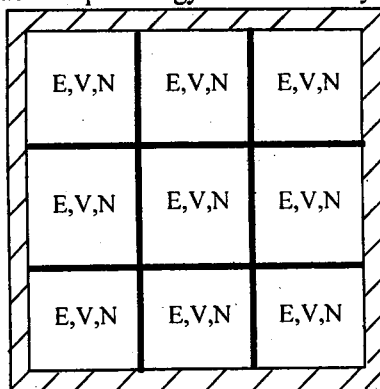
Consider a closed system for which the total energy H(q.p) = E remains constant, according to the equation

$$E(q_1, \ldots q_f, p_1, \ldots p_f) = \text{constant.}$$

The locus of all the phase points having equal energies in phase space is called an energy surface or ergodic surface. We can now imagine a family of such energy surface constructed in phase space and let us consider two neighboring surfaces with energies E and E + δE. Each surface divides the phase space in to two parts, one of higher and the other of lower energy, so that, they will never intersect each other. As they include some phase volume in between them, they will contain certain number of phase points. The number of phase points between them will be constant. A very useful ensemble can be obtained by taking the density as equal to zero for all values of the energy except in a selected narrow range E & E+δE of Gibbs, such an ensemble, specified by

$$\rho = \text{constant (in the range E to E+δE)}$$
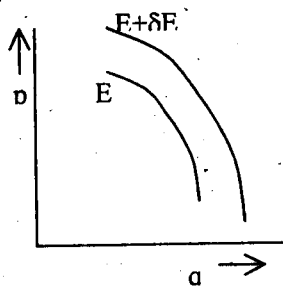
$$\rho = 0 \quad \text{(outside this range)}$$

Fig 1.7 Energy shell in phase space.

may be called a micro-canonical ensemble.

We observe the following properties: In case of micro-canonical ensemble

(i)   As ρ is a function of energy, this ensemble is in statistical equilibrium

(ii)   The average properties predicated by such ensemble will not vary in the time being in statistical equilibrium.

(iii)   As ρ is constant with in the energy shell, the distribution of phase points is uniform (by Liouville's theorem).

An ensemble of this kind can be regarded as obtained from an originally uniform ensemble by discarding all systems having phase points with positions that do not fall with in the limits in the phase space that correspond the energy range E to E+δE.

$$\overline{x} = \frac{1}{N} \int_{-\infty}^{+\infty} x N(x) dx \, , \, N = \int_{-\infty}^{+\infty} N(x) dx \quad \text{--------- (1.2)}$$

In general, if $R(x)$ is any arbitrary property of the points.

$$\overline{R} = \frac{\int_{-\infty}^{+\infty} R(x) N(x) dx}{\int_{-\infty}^{+\infty} N(x) dx} \quad \text{--------- (1.3)}$$

Generalization to higher dimensions is straight forward.

1.5 Density of Distribution in the phase space:

The use of ensembles in statistical mechanics is guided by the following factors:

1.The aim is to know only the number of system or elements that would be found in different states, that is, in different regions of the Γ space, at any time, All the elements being similar in structure, we need not distinguish between them.

2. The number of elements in the ensemble is so large $(m \rightarrow \alpha)$ that there is a continuous change in their number in passing from one region of the phase space to another.

We can therefore, describe the condition of an ensemble by a density D with which the phase points are distributed in the Γ space. It is called the distribution function (or density of distribution or probability density).

In an ensemble of systems of f degrees of freedom, D is a function of 2f position and momentum coordinates q_1, q_2, ...q_f, p_1, p_2, ...p_f which correspond to the 2f axes in the phase space. It can also depend on time t explicitly. The reason is that although we are free to fix the distribution at any given time t_0, we have as yet no assurance for the distribution to remain same. If it remains same, the particular distribution would be one of equilibrium. We shall discuss this later on. Thus, in general,

$$D = D(q_1, q_2, \ldots\ldots q_f, p_1, p_2, \ldots\ldots p_f, t) = D(q, p, t) \quad \text{--------- (1.4)}$$

Consider a small region A of the Γ space such that the position coordinates lie between q_1 and q_1+dq_1, ...,q_f and q_f+dq_f, and the momenta lie between p_1 and p_1+dp_1, ... p_f and p_f+dp_f (Fig 1.5a). The hyper volume of this region is

$$d\Gamma = dq_1 \ldots\ldots dq_f dp_1 \ldots\ldots dp_f = dpdq \quad \text{-------- (1.5)}$$

By the definition of density, the number of system or elements dM lying in the specified infinitesimal region situated at the phase point q_1, ... p_f at the instant t is

$$dM = D(q, p, t) d\Gamma \quad \text{--------- (1.6)}$$

If M is the total number of elements in phase space, then at every instant t,

$$M = \int D d\Gamma \text{ --------- (1.7)}$$

Where the integration is over the whole phase space.

As suggested by eqn (1.3), the ensemble average of a quantity R(q,p) is defined by

$$\overline{R} = \frac{\int R(q, p) D(q, p, t) d\Gamma}{\int D(q, p, t) d\Gamma} = \frac{1}{M} \int R(q, p) D(q, p, t) dt \text{ --------- (1.8)}$$

If a system is selected at random from the ensemble, the probability of selecting one whose phase point lies in the small region at the point $q_1, \ldots\ldots p_f$ is simply ρdT, where

$$\rho = \frac{D}{\int D dT} = \frac{D}{M} \int \rho dT = 1 \text{ --------- (1.9)}$$

In terms of $\rho(q, p, t)$, called the normalized density of distribution,

$$\overline{R} = \frac{\int R(q, p) \rho(q, p, t) dp\, dq}{\int \rho(q, p, t) dp\, dq} = \int R \rho dT \text{ --------- (1.10)}$$

The ensemble average equation (1.10) gives the average value of the physical quantity R for the actual system of interest.

The macroscopic average properties (like E, V, N) of a system in thermodynamic equilibrium do not change with time. Therefore, our ensemble representing it must be such that the ensemble averages are time independent. This is a reasonable requirement. It follows that to construct a suitable ensemble we should study the behavior of ρ(or D) with time.

1.6 EQUIPARTITION THEOREM:

(Law of Equipartition of energy)

Statement: Equipartition theorem states that if the energy of the system E(p, q) is a homogeneous[1] quadratic[2] function of any set of l of the generalized coordinates and momenta then the average energy of the system is < E > or $\overline{E} = lkT/2$, provided that for each momentum in the set

---

1. A function f $(x_1, x_2, \ldots.x_m)$ is homogeneous of degree 'm' in the variables x_1, x_2 if

   $f(\lambda x_1, \lambda x_2, x_3 \ldots x_m) = \lambda^n f(x_1, x_2, \ldots.x_m)$

2. A function E_i is quadratic in p_i if it is of the form $E_i(p_i) = bp_i^2$, where b is constant.

$$\left[p_i \exp\left(\frac{-E}{KT}\right) \right]_{p_i^{'}}^{p_i^{''}} = 0,$$

and for each coordinate in the set

$$\left[q_i \exp\left(\frac{-E}{KT}\right) \right]_{q_i^{'}}^{q_i^{''}} = 0,$$

Where $p_i^{'}$, $p_i^{''}$, $q_i^{'}$ and $q_i^{''}$ are the lower and upper limits respectively of p_i and q_i .

In other words , it states that the mean value of each independent quadratic term in the energy is equal to $\frac{1}{2}$ kT. As a law of equipartition of energy it states that, when the total energy of the system is expressible as an a additive quadratic function of momenta or position coordinates, the energy associated with each degree of freedom of the system is $\frac{1}{2}$ kT per particle.

Proof: We consider a system described classically in terms of f position coordinates $(q_1....q_2 ......q_f)$ and corresponding f momenta $(p_1...... p_2, ...p_f)$. The total energy of the system can then be written as a function of q's and p's, i.e.

$$E = E(q_1....q_f, p_1.....p_f) \text{-------- (1.11)}$$

This total energy can be broken into a sum of two parts.

$$E = E_i(p_i) + E^{'}(q_1, q_2 ...... q_f, p_1, p_2, ...p_f) \text{-------- (1.12)}$$

Where $E_i(p_i)$ is a function of particular momentum p_i alone and the second term $E^{'}(q_1 ..... p_f)$ is a function of all the position coordinates and all momenta $(p_1 .... p_f)$ excluding the particular momentum p_i which is the only variable of the first term of the sum. Such a split of energy is of course feasible as the position and momentum coordinates are independent of each other.

$$E_i (p_i) = bp_i^2 \text{--------- (1.13)}$$

Where b is some constant

It the system is in equilibrium at a particular temperature T, it will be characterized by canonical distribution and the mean value of energy E_i can be expressed as

$$<E_i> = \frac{\sum E_i e^{-\beta E(q_1....p_f)}}{\sum e^{-\beta E(q_1....p_f)}} \text{---------(1.14)}$$

But in classical theory variation of energy is continuous and we can charge the summation into integration,

i.e.,

$$\langle E_i \rangle = \frac{\int\limits_{-\infty}^{\infty}...(2f)...\int\limits_{-\infty}^{\infty} E_i e^{-\beta E(q_1....p_f)}dq_1....dp_f}{\int\limits_{-\infty}^{\infty}...(2f)...\int\limits_{-\infty}^{+\infty} e^{-\beta E(q_1....p_f)}dq_1....dp_f} \quad\text{------- (1.15)}$$

using equation (1.12)

$$\langle E_i \rangle = \frac{\int\limits_{-\infty}^{\infty}...(2f)...\int\limits_{-\infty}^{+\infty}E_i e^{-\beta(E_i+E^1)}dq_1...dp_f}{\int\limits_{-\infty}^{\infty}...(2f)...\int\limits_{-\infty}^{\infty}e^{-\beta(E_i+E^1)}dq_1...dp_f} = \frac{\int\limits_{-\infty}^{\infty}e^{-\beta E_i(p_i)}E_i dp_i \int\limits_{-\infty}^{\infty}...(2f-1)...\int\limits_{-\infty}^{\infty}e^{-\beta E^1}dq_1...dp_f}{\int\limits_{-\infty}^{\infty}e^{-\beta E_i(p_i)}dp_i. \int\limits_{-\infty}^{\infty}...(2f-1)...\int\limits_{-\infty}^{\infty}e^{-\beta E^1}dq_1...dp_f}$$

where the integrals containing exp $(-\beta E')$ extend over all terms q and p except p_i. These integrals are equal and thus cancel leaving

$$\langle E_i \rangle = \frac{\int\limits_{-\infty}^{\infty}e^{-\beta E_i}E_i dp_i}{\int\limits_{-\infty}^{\infty}e^{-\beta E_i}i\, dp_i} = \frac{-\dfrac{\partial}{\partial\beta}\left(\int\limits_{-\infty}^{\infty}e^{-\beta E_i(p_i)}dp_i\right)}{\int\limits_{-\infty}^{\infty}e^{-\beta E_i}dp_i}$$

$$= -\frac{\partial}{\partial\beta}\log_e\left(\int\limits_{-\infty}^{\infty}e^{-\beta E_i(p_i)}dp_i\right)$$

using equation (1.13), we find

$$\langle E_i \rangle = -\frac{\partial}{\partial\beta}\log_e\left(\int\limits_{-\infty}^{\infty}e^{-\beta b p_i^2}dp_i\right) \quad\text{--------- (1.16)}$$

Let us introduce the variable $\qquad y = \beta^{\frac{1}{2}}p_i$

only in the argument of logarithm so that

$$\langle E_i \rangle = -\frac{\partial}{\partial\beta}\log_e\left(\beta^{-\frac{1}{2}}\int\limits_{-\infty}^{\infty}e^{-by^2}dy\right)$$

$$= -\frac{\partial}{\partial\beta}\left[-\tfrac{1}{2}\log_e\beta + \log_e\int\limits_{-\infty}^{\infty}e^{-by^2}dy\right] = -\frac{\partial}{\partial\beta}(-\tfrac{1}{2}\log_e\beta)$$

$$= \frac{1}{2\beta} = \tfrac{1}{2}kT \quad\text{--------- (1.17)} \quad (\because \beta = \frac{1}{kT})$$

Equation (1.17) is the so called 'equipartition theorem' of classical statistical mechanics and obviously states that the mean value of each independent quadratic term in the energy is equal to $\frac{1}{2}$ kT.

We have earlier mentioned that classical theory assumes a continuous variation of energy and thus enabling us to change the summation into integration and hence a subsequent deduction of equipartition theorem. It should therefore, be emphasized that the equipartition theorem is valid only in classical statistical mechanics. However, when the temperature is sufficiently high (and thus the mean energy E of the system is sufficiently high). The spacing ΔE between the energy levels around the mean energy E is small compared to the thermal energy kT, i.e., $\Delta E \ll$ kT and them the variation of energy can be approximated as continuous. In that case quantum mechanical description of discrete energy levels can be ignored and equipartition theorem can be taken as valid. But when $\Delta E \geq$ kT, classical description breaks down.

At present we shall calculate the mean energy for one- dimensional harmonic oscillator.
Hamiltonian for a one- dimensional harmonic oscillator is

$$H = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 q^2 \quad \text{--------} \quad (1.18)$$

Which involves two quadratic terms.

The mean energy is then

$$<E_i> = \frac{\frac{1}{2m}\int_{-\infty}^{\infty} p^2 e^{-\beta p^2/2m} dp}{\int_{-\infty}^{\infty} e^{-\beta p^2/2m} dp} + \frac{\frac{m\omega^2}{2}\int_{-\infty}^{\infty} q^2 e^{-\beta m\omega^2 q^2/2} dq}{\int_{-\infty}^{\infty} e^{-\beta m\omega^2 q^2/2} dq} \quad \text{---------} \quad (1.19)$$

$$= \frac{1}{2\beta} + \frac{1}{2\beta} = \frac{1}{\beta} = kT \quad (\because \beta = \frac{1}{kT})$$

(Based on the definite integrals $\int_{-\infty}^{\infty} x^2 e^{-ax^2} dx = \frac{1}{2}\left(\frac{\pi}{a^3}\right)^{1/2}$ and $\int_{-\infty}^{\infty} e^{-ax^2} dx = \left(\frac{\pi}{a}\right)^{1/2}$

the integrals in equation (1.19) are evaluated)

which is expected since there are two quadratic terms and each would contribute $\frac{1}{2}$ kT in accordance with equation (1.17).

Summary:

Statistical mechanics is a scheme for studying the macroscopic properties of systems in terms of their microscopic properties. It is a mechanics because it deals with mechanical systems and statistical because it makes only imprecise (or statistical) predictions, the general character of statistical mechanics does not depend on the mechanics (either classical or quantum) which applies only to idealized situations in which we obtain the complete information concerning the motion of a mechanical system. Statistical mechanics aims to study the physical properties of mechanical system in a situation when the description is incomplete. The basic mathematical tool in this approach is the theory of probability. Only statistical or probabilistic consideration can yield a theory, which predicts macroscopic properties from an atomic model of the system.

Further statistical mechanics predicts the average value of thermodynamic (or macroscopic) properties, which differs from the exact value. However the deviation from the exact value decreases with increasing the size of the system (i.e. with the increase in the number of particles in the system). So the statistical predictions are nearly correct for a system consisting of a very large number of particles. Thus the statistical mechanics is an extremely useful tool for solving many body problems.

A molecule of an ideal gas can be represented as a structure less particle. Such a particle has three translational degrees of freedom. As a result, its phase space has six dimensions whose Cartesian coordinates are $x_1, x_2, x_3, p_1, p_2, p_3$ and its phase space is called μ - space where μ-stands for molecule. We can also construct the phase space for the system (gas) consisting of N molecules and it is called the Γ - space, where Γ - stands for gas. It is spanned by 3N coordinates and 3N momentum axes. The 6N coordinates in the Γ- space represents, the positions and momenta of all the molecules (state of the system) at a given time. Therefore the instantaneous state of the whole system (gas of N molecules) is given by a single representative point (or phase point). Each point on the phase line represents one such possible microscopic state.

Each phase point on the phase line of a single system develops out of the previous point in time, according to the laws of mechanics. Gibbs replaced the time dependent picture by static picture in which the entire phase line exist at one time. Let us imagine a large no. of M (M→ α) systems, similar in structure to the system of interest but suitably randomized in the accessible, unobservable microscopic states. Instead of taking the time average, we take an average over this artificially constructed group existing simultaneously at one time. Therefore, such a group of replicas of collections of similar, non-interacting, independent, imaginary systems is called an ensemble.

There are three types of ensembles. (1) Micro canonical (2) Canonical and (3) Grand canonical ensembles. In the micro canonical ensemble neither exchange of energy nor exchange of particles occur

among the assemblies (or systems). But the systems can exchange energy among themselves but not the particles in the canonical ensemble. In the Grand canonical ensemble, the exchange of energy and particles will occur between the systems.

According to law of equipartition of energy, the mean value of each independent quadratic term to the energy is equal to ½ kT per molecule. Equipartition theorem is valid only in classical statistical mechanics because the classical theory assumes continuous variation of energy or the energy levels are closely spaced.

Key terminology:

1.Phase space: Suppose a bead of mass 'm' moves freely and arbitrarily on a string stretched along the x –

axis. It has one degree of freedom. The position of the bead at a time 't' is x(t) and its velocity is $v_x = \dot{x}$ or

(momentum $p_x = m\dot{x}$) at that instant. Therefore, the state of the bead at any instant can be represented by a point 'P' in the hypothetical two dimensional space called the phase space. whose coordinates are x and p_x.

2.Ensemble: Collection of similar , non interacting, imaginary, independent systems is called an ensemble. The members of an ensemble are identical in macroscopic properties. They differ only in their microscopic states. The members of an ensemble are called elements.

Locus: The path joining all phase points in the phase space is called locus.

Self – assessment questions:

1. Define phase space. What do you mean by μ space and Γ space?

2. Define an ensemble. Distinguish three types of ensembles, namely, micro canonical, canonical and grand canonical ensembles.

3. What do you mean by an ensemble average?

4. State and prove equipartition theorem.

5. Calculate the mean energy for one dimensional linear harmonic oscillator.

Reference:

1. Statistical Mechanics by Gupta & Kumar, K Nath & Co.

2. Statistical Mechanics by Kerson Huang, Wiley Eastern Ltd.

3. Statistical Mechanics by M. Eisner and B.K. Agarwal, Wiley Eastern Ltd.,

4. Statistical Mechanics: Theory and Applications by S.K. Sinha ,Tata McGraw Hill Co, Ltd.,

MICROCANONICAL ENSEMBLE
(Isolated system)

Objectives:

1. To discuss the classical ideal gas on the basis of micro-canonical ensemble.

2. To explain the concept of Gibb's paradox and how it was resolved.

Structure:

2.1 Introduction

2.2 Perfect Gas (or) classical ideal Gas-Micro canonical Ensemble.

2.3 Gibbs Paradox.

2.1 Introduction

An Ensemble in which systems have the same energy, same number of particles and same volume is called a micro-canonical ensemble. In this ensemble, density ρ for a closed isolated thermodynamically system (a system whose energy remains constant is called an isolated system) is a function of energy, and we take

$$\rho(E) \begin{cases} = \text{constant} & \text{between the energy shells E and E+}\delta\text{E of phase space.} \\ = 0 & \text{outside the region of phase space.} \end{cases}$$

2.2 Perfect Gas (or) classical ideal Gas – Micro canonical Ensemble:

We consider a micro canonical ensemble of perfect gas. Let there be n point particles of mass m in a volume V with total energy in the phase space is given by δE at E. The corresponding volume in phase space is given by

$$\Delta\Gamma = \int dq_1 \ldots dq_{3n} \int dp_1 \ldots dp_{3n} \quad \text{------- (2.1)}$$

where q's refer to position and p's to momentum co-ordinates. Since in a perfect gas, interaction between the particles is not present, energy of a perfect gas is independent of the positions of the particles and we can write

$$\int dq_1 \ldots dq_{3n} = \int dx_1\, dy_1\, dz_1 \int dx_2\, dy_2\, dz_2 \ldots \int dx_n\, dy_n\, dz_n = V^n.$$

Therefore, the corresponding volume $\Delta\Gamma$ in the phase space and shown in equations (2.1) can be written as

$$\Delta\Gamma = V^n \int dp_1...dp_{3n} \quad \text{------} \quad (2.2)$$

where the momentum space integral is to be evaluated subject to the following constraint imposed by the ensemble:

$$E - \delta E \leq E_r \leq E,$$

since energy $E_r = \sum_{i=1}^{3n} p_i^2 / 2m$ (non – relativistic case)

The above constraint becomes

$$E - \delta E \leq \sum_{i=1}^{3n} p_i^2 / 2m \leq E \quad \text{------} \quad (2.3)$$

The integral in equation (2.2) is just the volume contained between the 3n dimensional hyper sphere of radius $(2mE)^{1/2}$ and 3n dimensional hyper sphere of radius $[2m(E - \Delta E)]^{1/2}$. The volume of n dimensional hyper-sphere of radius R is

$$V_n(R) = C_n R^n, \quad \text{------} \quad (2.3(a))$$

$$\text{where} \quad C_n = \frac{\pi^{n/2}}{\Gamma\left(\frac{n+2}{2}\right)} = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)} = \frac{\pi^{n/2}}{\frac{n}{2}!}$$

$$\because \Gamma(n+1) = n!$$

based on the above equation 2.3(a), the volume of three dimensional sphere can be written as.

$$V_3(R) = C_3 R^3$$

$$\text{Where} \quad C_3 = \frac{\pi^{3/2}}{\Gamma\left(\frac{3}{2}+1\right)} = \frac{\pi^{3/2}}{\Gamma\left(\frac{5}{2}\right)} = \frac{\pi^{3/2}}{\frac{3}{2}\Gamma\frac{3}{2}} = \frac{\pi^{3/2}}{\frac{3}{2}\cdot\frac{1}{2}\Gamma\frac{1}{2}} = \frac{\pi^{3/2}}{\frac{3}{4}\sqrt{\pi}} = \frac{4}{3}\pi$$

Therefore, the volume of the 3 dimensional hyper sphere becomes

$$V_3(R) = \frac{4}{3}\pi R^3$$

In the similar manner the volume for 3n dimensional hyper sphere of radius $(2mE)^{1/2}$ will be

$$V_{3n} = \frac{\pi^{3n/2}}{\frac{3n}{2}!} (2mE)^{3n/2}.$$

Therefore the volume contained between hyper sphere of radius $(2mE)^{1/2}$ and of radius $[2m(E-\Delta E)]^{1/2}$ will be

$$\int dp_1...dp_{3n} = \frac{\pi^{3n/2}}{\frac{3n}{2}!} [(2mE)^{3n/2} - [2m(E-\Delta E)]^{3n/2}]$$

$$= \frac{\pi^{3n/2}}{\frac{3n}{2}!} (2mE)^{3n/2} \left[1 - \left(1 - \frac{\Delta E}{E} \right)^{3n/2} \right]$$

By the definition of exponential function, $\left(1 - \frac{\Delta E}{E} \right)^{3n/2} \cong e^{-\frac{3n}{2}\frac{\Delta E}{E}}$

Since

$$(1-x)^n = 1 - nx + \frac{n(n-1)}{2!}x^2 - \frac{n(n-1)(n-2)}{3!}x^3 + ---$$

$$= 1 - nx + \frac{n^2}{2!}x^2 - \frac{n^3}{3!}x^3 + ------ \qquad \text{when n is sufficiently large}$$

$$= e^{-nx}$$

$$\therefore \int dp_1...dp_{3n} = \frac{\pi^{3n/2}}{\frac{3n}{2}!}(2mE)^{3n/2} [1 - e^{-3n/2 . \Delta E/E}]$$

For a macroscopic system $3n \approx 10^{23}$ and for

$$\frac{3n}{2}\Delta E >> E.$$

The exponential term in the bracket vanishes. Hence,

$$\int dp_1...dp_{3n} = \frac{\pi^{3n/2}}{\frac{3n}{2}!}(2mE)^{3n/2}$$

Substituting this in equation (2.2), we get the volume in the 6n dimensional phase space.

$$\Delta\Gamma = V^n \frac{\pi^{3n/2}}{\frac{3n}{2}!}(2mE)^{3n/2}$$

In classical statistical mechanics, the entropy σ of a system in statistical equilibrium is defined as

$$\sigma = \log_e \Delta\Gamma = \log_e \left[V^n \frac{\pi^{3n/2}}{\frac{3n}{2}!} (2mE)^{3n/2} \right]$$

$$= n \log_e \left[V \pi^{3/2} (2mE)^{3/2} \left(\frac{2}{3n} \right)^{3/2} \right] + \frac{3n}{2}$$

$$= n \log_e \left[V \left(\frac{4\pi m}{3} \right)^{3/2} \left(\frac{E}{n} \right)^{3/2} \right] + \frac{3n}{2} \quad \text{---------} \quad (2.4)$$

[By making use of the Stirling's approximation to evaluate the factorial

$\log_e n! \simeq \log_e n - n$, with e = 2.7182 as the base of natural logarithm.]

As the entropy should not depend upon the units of hyper volume $\Delta\Gamma$, to make it dimensionless we divide

it by h^{3n}. $\left(\frac{\Delta\Gamma}{h^{3n}} \right)$ represents the number of states accessible to the ensemble. Introducing h^{3n} (where h is

plank's constant) as the limit of volume in the phase space.

$$\sigma = \log \left(\frac{\Delta\Gamma}{h^{3n}} \right)$$

$$\sigma = n \log \left[\frac{V \left(\frac{4\pi m}{3} \right)^{3/2} \left(\frac{E}{n} \right)^{3/2}}{h^3} \right] + \frac{3n}{2} \quad \text{---------} \quad (2.5)$$

This does not satisfy the additive property of the entropy as the volume V appears in the argument of the

logarithm. It is therefore not possible to divide the system into two parts and to write the total entropy as the

sum ($\sigma_1 + \sigma_2$). This result will, however, bear, the desired additive property if we introduce n! term, so

that.

$$\sigma = \log \left[\frac{\Delta\Gamma}{n! h^{3n}} \right]$$

$$= n \log_e \left[\frac{V \left(\frac{4\pi m}{3} \right)^{3/2} \left(\frac{V}{n} \right) \left(\frac{E}{n} \right)^{3/2}}{h^3} \right] + \frac{5n}{2}$$

$$\sigma = n \log_e \left[\left(\frac{4\pi m}{3h^2} \right)^{3/2} \left(\frac{V}{n} \right) \left(\frac{E}{n} \right)^{3/2} \right] + \frac{5n}{2} \text{--------- (2.6)}$$

We note that now in the argument of logarithm $\left(\dfrac{V}{n} \right)$ and $\left(\dfrac{E}{n} \right)$ occur which are volume per particle and the energy per particle respectively and hence σ is now additive.

We, now, establish the relations between the thermo dynamical and statistical quantities.

(a) Internal Energy:

We can write for statistical temperature. τ

$$\frac{1}{\tau} = \left(\frac{\partial \sigma}{\partial E} \right)_{V,n}$$

$$= \frac{\partial}{\partial E} \left[n \log \left(\frac{4\pi m}{3h^2} \right)^{3/2} + n \log V - n \log n + \frac{3n}{2} \log E - \frac{3n}{2} \log n \right] + \frac{\partial}{\partial E} \cdot \frac{5n}{2}$$

$$= \frac{\partial}{\partial E} \left(\frac{3n}{2} \log E \right).$$

(Because V and n are constants and the derivative of the rest of the terms w.r.t. E will be zero.)

$$\frac{1}{\tau} = \frac{3n}{2} \cdot \frac{1}{E} \qquad\qquad (\because \frac{\partial}{\partial E} (\log E) = \frac{1}{E})$$

So that $\quad E = \dfrac{3n}{2} \cdot \tau \quad = \dfrac{3}{2} nkT \quad (\because \tau = kT)$ ——— (2.7)

Which is the familiar result for the internal energy of a perfect gas.

(b) Relation between τ and T (thermodynamic temperature):

From equation 2.7 it is obvious that

$$\tau = kT \text{ ———— (2.8)}$$

This is the relation between statistical temperature τ and thermodynamic temperature T.

(c) Relation between τ and P (thermodynamic pressure):

$$\frac{p}{\tau} = \left[\frac{\partial \sigma}{\partial V} \right]_{n,E}$$

$$= \frac{\partial}{\partial V}\left[n \log_e\left(\frac{4\pi m}{3h^2}\right)^{3/2} + \frac{3n}{2}\log E - \frac{3n}{2}\log n + n \log V - n \log n \right] + \frac{\partial}{\partial V}\cdot\left(\frac{5n}{2}\right)$$

$$= \frac{\partial}{\partial V}\ (n \log V) \qquad\qquad (\because \text{The derivatives of all other terms should be zero}).$$

$$\frac{p}{\tau} = n\frac{\partial}{\partial V}\log_e V = n\ \frac{1}{V}$$

So that, $pV = n\tau$

$$PV = nkT \qquad (\because \tau = kT) \qquad \text{---------} (2.9)$$

(d) Entropy S (thermodynamic) of perfect gas:

We know that the relation between S and σ is

$$S = k\sigma$$

$$= nk \log\left[\left(\frac{4\pi m}{3h^2}\right)^{3/2}\left(\frac{V}{n}\right)\left(\frac{E}{n}\right)^{3/2}\right] + \frac{5n}{2}k$$

Putting $\quad E = \frac{3}{2}nkT$, we get

$$S = nk \log\left[\left(\frac{2\pi mkT}{h^2}\right)^{3/2}\left(\frac{V}{n}\right)\right] + \frac{5}{2}nk \qquad \text{--------} (2.10)$$

Which is famous Sackur – Tetrode formula for the entropy of a perfect gas.

The thermal de-Broglie wavelength associated with a molecule may be defined as

$\lambda = h$ / average thermal momentum of a molecule.

We note that the quantity $(2\pi mkT)^{1/2}$ has the character of an average thermal momentum of a molecule, then

$$\lambda = \frac{h}{(2\pi mkT)^{1/2}}$$

putting this in equation (2.10), we get

$$S = nk \log\left[\left(\frac{V}{n}\right)\cdot\frac{1}{\lambda^3}\right] + \frac{5}{2}nk \qquad \text{-------} (2.11)$$

Where the argument of logarithm has a ratio of volume per particle $\left(\dfrac{V}{n}\right)$ to the volume λ^3 associated with

de Broglie wavelength

(e) Chemical Potential, μ, of a Perfect gas:

We know that

$$-\frac{\mu}{\tau} = \left(\frac{\partial \sigma}{\partial n}\right)_{E,V}$$

$$= \frac{\partial}{\partial n}\left[n \log\left(\frac{V}{n}\right) - n \log \lambda^3\right] + \frac{\partial}{\partial n}\cdot\left(\frac{5}{2}n\right)$$

$$= \frac{\partial}{\partial n}[n \log V - n \log n - n \log \lambda^3] + \frac{5}{2}$$

$$= \log V - 1 - \log n - \log \lambda^3 + \frac{5}{2}$$

$$= \log\left(\frac{V}{n\lambda^3}\right) + \frac{3}{2}$$

$$\frac{\mu}{\tau} = -\log\left(\frac{V}{n\lambda^3}\right) - \frac{3}{2}$$

$$= \log\left(\frac{V}{n\lambda^3}\right)^{-1} - \frac{3}{2}$$

$$\frac{\mu}{\tau} = \log\left[\frac{n\lambda^3}{V}\right] - \frac{3}{2}$$

Putting $\dfrac{p}{\tau} = \dfrac{n}{V}$ from relation (c)

$$\frac{\mu}{\tau} = \log\left[\frac{p\lambda^3}{\tau}\right] - \frac{3}{2}$$

$$\mu = \tau\left(\log p\lambda^3 - \log\tau - \frac{3}{2}\right)$$

$$= \tau\left(\log p + \log\lambda^3 - \log\tau - \frac{3}{2}\right)$$

$$= \tau \log p + \tau \log \frac{\lambda^3}{\tau} - \frac{3\tau}{2}$$

$$= \tau \log p + f(\tau), \quad \text{------} \quad (2.12)$$

Where $f(\tau)$ is a function of temperature alone

$$\therefore f(\tau) = \tau \log \frac{\lambda^3}{\tau} - \frac{3\tau}{2}.$$

2.3 Gibbs Paradox:

Previously we defined the entropy of a system in statistical equilibrium as

$$\sigma = \ln \Delta\Gamma$$

after evaluating the $\Delta\Gamma$, the expression for the entropy of a system is written as

$$\sigma = n \log \left[V \left(\frac{4\pi m}{3} \right)^{3/2} \left(\frac{E}{n} \right)^{3/2} \right] + \frac{3}{2} n$$

The above equation runs into two main difficulties

(1) The entropy σ should not depend upon the units of hyper volume $\Delta\Gamma$ (since $\Delta\Gamma$ has the dimensions of distance x momentum)

(2) Entropy σ is not additive because the volume V (not V/n)

occurs in the argument of logarithm. This prevents us from dividing the system into two parts and writing $\sigma = \sigma_1 + \sigma_2$

Of the above two difficulties, the first difficulty is easily removed by dividing $\Delta\Gamma$ by h^{3n}. Therefore

$\sim \dfrac{\Delta\Gamma}{h^{3n}}$ represents the number of states accessible to the ensemble.

The second difficulty is not easily removed. In fact, it leads to the famous Gibb's paradox (apparently self contradictory term having the truth). After dividing $\Delta\Gamma$ by h^{3n}, the expression for entropy becomes

$$\sigma = n \log \left[V \left(\frac{E}{n} \right) \left(\frac{4\pi m}{3h^2} \right)^{3/2} \right] + \frac{3n}{2}$$

$$= n \log \left[V \left(\frac{E}{n} \right)^{3/2} \right] + n \log \left(\frac{4\pi m}{3h^2} \right)^{3/2} + \frac{3n}{2}$$

$$= n \log \left[V \left(\frac{E}{n} \right)^{3/2} \right] + \frac{3n}{2} \left[\log \left(\frac{4\pi m}{3h^2} \right) + 1 \right]$$

$$= n \log \left[V \left(\frac{E}{n} \right)^{3/2} \right] + n \sigma_0 \quad \text{------} \quad (2.13)$$

$$\text{Where } \sigma_0 = \frac{3}{2} \left[\log \left(\frac{4\pi m}{3h^2} \right) + 1 \right] \quad \text{-------} \quad (2.14)$$

(A) Mixing of Two Different Ideal Gases:

The mixing of two different gases is an irreversible process. It is therefore attended by an increase of the entropy. Consider two different ideal gases (n_1, v_1, T) and (n_2, v_2, T), Fig 2.1 They are allowed to mix by removing the partition reversibly.

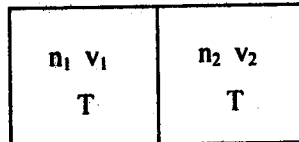| $n_1 \; v_1$	$n_2 \; v_2$
T	T

Fig 2.1 Mixing of two gases

It can be regarded as the expansion of each of the gases to the volume $v = v_1 + v_2$. The temperature, and therefore $\frac{E}{n}$, remains unchanged for each gas.

From equation 2.13, the change in entropy is given by

$$\Delta \sigma = \sigma_{12} - (\sigma_1 + \sigma_2)$$

$$= (n_1 \ln v + n_2 \ln v) - (n_1 \ln v_1 + n_2 \ln v_2)$$

$$= n_1 \ln \left(\frac{v}{v_1} \right) + n_2 \ln \left(\frac{v}{v_2} \right) > 0 \quad \text{-------} \quad (2.15)$$

This gives the entropy of mixing for two different ideal gases is in agreement with experiments. For the case $n_1 = n_2 = n$ and $v_1 = v_2 = \frac{1}{2} v$, we get $\Delta \sigma = 2n \ln 2$.

(B) Mixing of one Ideal Gas with the same Ideal gas:

Suppose the two gases are the same. Then the removal of the partition should not affect the distribution of systems over the accessible states. The final entropy ought to be the same with, or without, the partition,

$$\Delta\sigma = \sigma_{12} - (\sigma_1 + \sigma_2) = 0 \quad \text{---------- (2.16)}$$

This result is in agreement with the thermodynamics of a reversible process and also with experiments, but contradicts (2.13, 2.15). The derivation of (2.13, 2.15) does not depend on the identity of molecules and so would give the same increase in entropy (2.15) even in this case. In particular, for the case $n_1 = n_2 = n$ and $v_1 = v_2 = \dfrac{v}{2}$, we get an unobserved and therefore unaccountable increase of entropy by $2n \ln 2$ when a partition is simply removed from a box containing the same gas throughout. This is the Gibbs paradox.

This Gibbs paradox implies that the entropy of a given gas depends on the history of the gas. For example if we imagine the present state of the gas to be achieved just by slowly removing one by one a large number of their partitions, then the final entropy can have any value one desires. This is certainly not tenable.

Resolution of the Paradox:

In the case(A), the removal of partition leads to the diffusion of the molecules through out the whole volume v (twice the volume if $v_1 = v_2 = \dfrac{v}{2}$). There is a random mixing of the different molecules and so an increase of disorder. This is an irreversible process and the increase of entropy (2.15) makes sense.

We can imagine the mixing to be a process in which the position of some of the molecules of one gas are interchanged with those of the other gas. Each such exchange creates a new state. Therefore, the number of accessible states increases or equivalently the entropy increases.

On the other hand, in the case(B), any such interchange of molecules is always an interchange between identical molecules. Therefore, no new state is created when the partition is removed. It follows that in this case the application of (2.13, 2.15) overestimates the number of accessible states because classically we have taken all the molecules, even of the same gas, as distinguishable.

The way out of the paradox is to regard all the identical molecules in the case (B) to be indistinguishable. If there are n molecules, then n! possible permutations among themselves do not lead to physically distinct situations. There is just one way of arranging them. Therefore, our

estimation of number of accessible states, or equivalently of $\Delta\Gamma$, is too large by a factor of n! we should replace $\Delta\Gamma$ by $\Delta\Gamma/n!$ (Boltzmann counting), so that (2.13) becomes.

$$\sigma = \ln\left(\frac{\Delta\Gamma}{n!h^{3n}}\right) = \ln\left(\frac{\Delta\Gamma}{h^{3n}}\right) - \ln n!$$

$$= \ln\left(\frac{\Delta\Gamma}{h^{3n}}\right) - (n \ln n - n)$$

$$= n \ln\left[\left(\frac{V}{n}\right)\left(\frac{E}{n}\right)^{3/2}\right] + n\sigma_0 \quad \text{---------} \quad (2.17)$$

$$\sigma_0 = \frac{3}{2}\ln\frac{4\pi m}{3h^2} + \frac{5}{2}$$

use of (2.17) gives the correct result (2.16) for the case B and reproduces the result (2.15) for the case (A). Thus the Gibbs paradox is resolved because of the appearance of the extra terms $- n \ln n$.

It makes the entropy properly additive, as now $\frac{V}{n}$, rather than V, appears in the argument of the

logarithm.

As indistinguishability of identical particles is assumed in quantum mechanics, the Gibbs paradox will not occur if $\Delta\Gamma(E)$ is calculated in quantum mechanics.

Summary:

The ensemble in which systems have same energy, no of particles and volume is called micro canonical ensemble. In this ensemble the density distribution function ρ for a closed isolated thermodynamically system is a function of energy and it remains constant in a particular energy shell in the phase space and it is zero outside the region of the shell.

The entropy of a system in statistical equilibrium should not depend upon the units of hyper volume $\Delta\Gamma$ in order to make it as a dimensional less quantity, we must divide $\Delta\Gamma$ by h^{3n}. Therefore $\frac{\Delta\Gamma}{h^{3n}}$ gives the no. of accessible states to ensemble. Further, it is not possible to under stand classically why we must divide $\frac{\Delta\Gamma}{h^{3n}}$ by n! in order to prove that entropy is an additive quantity.

The reason is inherently quantum mechanical, we know that quantum mechanically, atoms or molecules are inherently indistinguishable in the sense that a state of the gas or system is described by n - particle wave function, which is either symmetric or antisymmetric with respect to

interchange of any two particles. A permutation of the particles can at most change the wave function by sign and it does not produces a new state of the system. Therefore from this fact it seems reasonable that the Γ space volume element dq dp (dΓ) corresponds to not one but only

dΓ / n! states of the system. Hence we should divide $\left(\dfrac{\Delta\Gamma}{h^{3n}}\right)$ by n!. This rule of counting is known

as "correct Boltzmann counting". Therefore, in classical statistical mechanics we do not have consistent way in which we can regard the particles as indistinguishable. In all classical considerations except counting of the states we must continue to regard the particles in a gas as distinguishable. Therefore Gibbs paradox should be resolved quantum mechanically in which we can regard the particles as indistinguishable.

Key Terminology:

Microcanonical ensemble: Microcanonical ensemble is a collection of essentially independent assemblies having same energy, same number of particles and same volume, in which the density function is constant in the particular energy shell in the phase space. and out side this region, it is zero.

Isolated system: The system whose energy remains constant is called an isolated system.

Ideal gas: A gas in which particles move freely and independently with out having any interaction between them.

Self – assessment questions:

1. Derive an expression for entropy of classical ideal gas on the basis of microcanonical ensemble.
2. Establish the relation between the statistical and thermodynamical parameters.
3. What do you mean by Gibb's paradox? Explain how it can be resolved.

Reference:

1. Statistical Mechanics by Gupta & Kumar, K Nath & Co.
2. Statistical Mechanics by Kerson Huang, Wiley Eastern Ltd.
3. Statistical Mechanics by M.Eisner and B.K.Agarwal, Wiley Eastern Ltd.,
4. Statistical Mechanics :Theory and Applications by S.K. Sinha, Tata McGraw Hill Co, Ltd.,

Statistical Mechanics Part

UNIT III

Lesson 3

CANONICAL ENSEMBLE

Objectives:

1. To introduce the concept of distribution function and partition functions.

2. To explain the energy fluctuations in the canonical ensemble

3. To establish the equivalence between the canonical and micro canonical ensembles.

Structure:

3.1 Canonical Ensemble

3.2 Energy Fluctuations in the Canonical Ensemble

3.1 Canonical Ensemble:

In the first chapter, we studied the microcanonical ensemble in which the macrostate of the systems is defined by the fixed number of particles N, fixed volume V, and fixed energy E. However for most physical systems the concept of fixed energy does not appear satisfactory as the total energy E of a system is hardly ever measured. It is convenient to keep a fixed temperature T of the system in place of a fixed energy E. For this purpose the system is placed in contact with an appropriate heat reservoir (or thermostat). Now, an infinitely large number of mental copies of a given system in thermal contact with a heat reservoir form an ensemble in which the macrostate of the system is defined by the fixed parameters T, V, N. Such an ensemble is called a canonical ensemble. On the other hand, the microcanonical ensemble describes the systems, which are perfectly isolated.

Therefore in a canonical ensemble, the energy E of a system is necessarily variable. However the system in contact with the heat reservoir forms a composite system, whose energy is fixed. Such a composite system is described by the microcanonical ensemble.

We wish to consider the question, "what ensemble is appropriate for the description of a system not in isolation, but in thermal equilibrium with a large system?" To answer it we must find the probability that the system has energy E, because this probability is proportional to the density in Γ − space for the ensemble we want.

Consider an isolated composite system made up of two sub systems whose Hamiltonians are respectively H_1 (p_1, q_1) and H_2 (p_2, q_2), with number of particles N_1 and N_2 respectively. We assume that $N_2 \gg N_1$ but, that both N_1 and N_2 are macroscopically large. We are interested in system 1 only. Consider a micro-canonical ensemble of the composite system with total energy between E and E + 2δ. The energies E_1 and E_2 of the sub systems accordingly can have any values satisfying.

$$E < (E_1 + E_2) < E + 2\delta \text{ --------- (3.1)}$$

Although this includes a range of values of E_1, E_2, the analysis of micro canonical ensemble shows that only one set of values, namely $\bar{E}_1, \bar{E}_2$, is important. We assume that $\bar{E}_2 \gg \bar{E}_1$. Let Γ_2 (E_2) be the volume occupies by system 2 in its own Γ - space. The probability of finding system 1 in a state within dp_1 dq_1 of (p_1, q_1), regardless of the state of system 2, is proportional to dp_1 dq_1 Γ_2 (E_2), where $E_2 = E - E_1$. Therefore up to a proportionality constant the density in the Γ - space for system 1 is

$$\rho(\ p_1, q_1) \propto \Gamma_2(\ E - E_1) \text{ ---------- (3.2)}$$

since only the values near $E_1 = \bar{E}_1$ are expected to be important, and $\bar{E}_1 \ll E$, we may perform the expansion.

$$k \log \Gamma_2(\ E - E_1) = S_2(\ E - E_1) = S_2(E_2) \text{ ---------- (3.2(a))}$$

Since $\bar{E}_1 \ll E$, $\bar{E}_2$ approaching E, expand $S_2(E_2)$ in the above equation in Taylor's form above $E_2 = E$, Therefore

$$k \log \Gamma_2 (E - E_1) = S_2(E_2) = S_2(E) + (E_2 - E) \left(\frac{\partial S_2(E_2)}{\partial E_2} \right)_{E_2 = E} + \ldots\ldots\ldots$$

$$k \log \Gamma_2(\ E - E_1) = S_2(E) - E_1 \frac{\partial S_2(E)}{\partial E} + \ldots\ldots$$

$$k \log \Gamma_2(\ E - E_1) = S_2(\ E_2) - \frac{E_1}{T} + \ldots\ldots \left(\because \frac{\partial S_2(E)}{\partial E} = \frac{1}{T} \right)$$

$$k \log \Gamma_2(\ E - E_1) \approx S_2 (E) - \frac{E_1}{T} \text{ --------- (3.3)}$$

Where T is the temperature of the larger subsystem. Hence

$$\Gamma_2 (E - E_1) \approx \exp\left[\frac{1}{k}S_2(E)\right] \exp\left(-\frac{E_1}{kT}\right) \quad \text{---------} \quad (3.4)$$

The first factor is independent of E_1 and is thus a constant as far as the small sub system is concerned. Owing to (3.2) and the fact that $E_1 = H_1$ (p_1, q_1), we may take the ensemble density for the small sub system to be

$$\rho(p, q) = e^{-H(p,q)/kT} \quad \text{----------} \quad (3.5)$$

Where the subscript 1 labeling the subsystem has been omitted, since we may now forget about the larger subsystem, apart from the information that its temperature is T. The larger subsystem in fact behaves like a heat reservoir in thermodynamics. The ensemble defined by (3.5), appropriate for a system whose temperature is determined through contact with a heat reservoir, is called the canonical ensemble'.

The volume in Γ- space occupied by the canonical ensemble is called the partition function:

$$Q_N(V, T) \equiv \int \frac{d^{3N}p\, d^{3N}q}{N!\, h^{3N}} e^{-\beta H(p,\, q)} \quad \text{----------} \quad (3.6)$$

Where $\beta = 1/kT$, and where we have introduced a constant h, which has dimension of momentum x distance, in order to make Q_N dimensionless. The factor $1/N!$ appears, in accordance with the rule of "correct Boltzmann counting." These constants are of no importance for the equation of state.

Strictly speaking we should not integrate over the entire Γ- space in (3.6), because (3.2) requires that $\rho($ p_1, q_1) vanishes if $E_1 > E$. The justification for ignoring such a restriction is that in the integral (3.6) only one value of the energy H (p, q) contributes to the integral and that this value will lie in the range where the approximation (3.4) is valid. We prove this contention in sec 3.2.

The thermodynamics of the system is to be obtained from the formula

$$Q_N(V, T) = e^{-\beta A(V,\, T)} \quad \text{----------} \quad (3.7)$$

Where A (V, T) is the Helmholtz free energy. To justify this identification we show that

(a) A is an extensive quantity,

(b) A is related to the internal energy $U \equiv <H>$ and the entropy

$$S \equiv -\left(\frac{\partial A}{\partial T}\right)_V \quad \text{by the thermodynamic relation}$$

$$A = U - TS$$

That A is an extensive quantity follows from (3.6), because if the system is made up of two subsystems whose mutual interaction can be neglected, then Q_N is a product of two factors. To prove the relation (b), we first convert (b) into the following differential equation for A:

$$<H> = A - T \left(\frac{\partial A}{\partial T} \right)_V \qquad (3.8)$$

To prove (3.8), note the identity from equations 3.6 and 3.7

$$\frac{1}{N! h^{3N}} \int dp \, dq \, e^{\beta[A(V, T) - H(p, q)]} = 1 \qquad (3.9)$$

Differentiating with respect to β on both sides, we obtain

$$\frac{1}{N! h^{3N}} \int dp \, dq \, e^{\beta[A(V, T) - H(p, q)]} \left[A(V, T) - H(p, q) + \beta \left(\frac{\partial A}{\partial \beta} \right)_V \right] = 0 \qquad (3.9(a))$$

This is the same as

$$A(V, T) - U(V, T) - T \left(\frac{\partial A}{\partial T} \right)_V = 0 \quad \text{since} \quad \beta \left(\frac{\partial A}{\partial \beta} \right)_V = -T \left(\frac{\partial A}{\partial T} \right)_V , \quad <H> \equiv U$$

All other thermodynamic functions may be found from A (V, T) by the Maxwell relations in thermodynamics:

$$P = - \left(\frac{\partial A}{\partial V} \right)_T$$

$$S = - \left(\frac{\partial A}{\partial T} \right)_V$$

$$G = A + PV$$

$$U = <H> = A + TS$$

Therefore all calculations in the canonical ensembles begin (and nearly end) with the calculation of the partition function (3.6).

3.2 ENERGY FLUCTUATIONS IN THE CANONICAL ENSEMBLE:

We now show that the canonical ensemble is mathematically equivalent to the micro canonical ensemble in the sense that although the canonical ensemble contains systems of all energies the overwhelming majority of them have the same energy. To do this we calculate the mean square fluctuation of energy in the canonical ensemble. The average energy is

$$U = <H> = \frac{\int dpdq\, H\, e^{-\beta H}}{\int dpdq\, e^{-\beta H}} \text{ ------- (3.10)}$$

Hence multiplying both sides by $e^{\beta\, A(V,\, T)}$

$$\int dp\, dq[U - H\,(p, q)]\, e^{\beta[A(V,\, T) - H(p,\, q)]} = 0 \text{ -------- (3.11)}$$

Differentiating both sides with respect to β, we obtain

$$\int dp\, dq\, \frac{\partial U}{\partial \beta}\, e^{\beta[A(V,\, T) - H(p,\, q)]} +$$

$$\int dp\, dq(U - H)\, e^{\beta[A(V,\, T) - H(p,\, q)]}\, [A(V, T) - H(p, q) + \beta\left(\frac{\partial A}{\partial \beta}\right)_V] = 0$$

we know that $\beta\left(\dfrac{\partial A}{\partial \beta}\right)_V = - T\left(\dfrac{\partial A}{\partial T}\right)_V$

$$\int dp\, dq\, \frac{\partial U}{\partial \beta}\, e^{\beta[A(V,\, T) - H(p,\, q)]} +$$

$$\int dp\, dq(U - H)\, e^{\beta[A(V,\, T) - H(p,\, q)]} \left[A(V,T) - H(p,q) - T\left(\frac{\partial A}{\partial T}\right)_V \right] = 0$$

Divide the above equation throughout by $\int dp\, dq\, e^{\beta[A(V,\, T) - H(p,q)]}$

$$\frac{\partial U}{\partial \beta} + \frac{\int dpdq\, (U - H)^2\, e^{\beta(A-H)}}{\int dpdq\, e^{\beta(A-H)}} \text{ --------- (3.12)} \qquad \text{since} \left[A - H - T\left(\frac{\partial A}{\partial T}\right)_V \right] = (U - H)$$

Equation 3.12 can also be written as

$$\frac{\partial U}{\partial \beta} + <(U - H)^2> = 0 \text{ -------- (3.13)}$$

$$<(U - H)^2> = - \frac{\partial U}{\partial \beta} \text{ -------- (3.13(a))}$$

$$\left(\frac{\partial U}{\partial \beta}\right)_V = \left(\frac{\partial U}{\partial T}\right)_V \left(\frac{\partial T}{\partial \beta}\right)$$

we know that $\beta = \dfrac{1}{kT}$

$$\Rightarrow T = \frac{1}{k\beta}$$

$$\Rightarrow \frac{\partial T}{\partial \beta} = -\frac{1}{k\beta^2}$$

$$\Rightarrow \frac{\partial T}{\partial \beta} = -k\left(\frac{1}{k\beta}\right)^2$$

$$\Rightarrow \frac{\partial T}{\partial \beta} = -kT^2$$

$$\therefore \left(\frac{\partial U}{\partial \beta}\right)_V = -kT^2\left(\frac{\partial U}{\partial T}\right)_V = -kT^2 C_V \qquad \left[\text{since } C_V = \left(\frac{\partial U}{\partial T}\right)_V\right]$$

$$\therefore <(U - H)^2> = kT^2 C_V$$

$$<(U^2 - 2UH + H^2)> = kT^2 C_V$$

$$<(<H>^2 - 2<H>H + H^2)> = kT^2 C_V \quad (\because U = <H>)$$

$$<H>^2 - 2<H>^2 + <H^2> = kT^2 C_V$$

Therefore the mean square fluctuation of energy is

$$<H^2> - <H>^2 = kT^2 C_V \text{ ---------- (3.14)}$$

For a macroscopic system $<H> \propto N$ and $C_V \propto N$. Hence (3.14) is a normal fluctuation. As N $\rightarrow \infty$, almost all systems in the ensemble have the energy $<H>$, which is the internal energy. Therefore the canonical ensemble is equivalent to the micro canonical ensemble.

It is instructive to calculate the fluctuations in another way. We begin by calculating the partition function in the following manner:

$$Q_N(v, T) \equiv \frac{\int dp dq \, e^{-\beta H}}{N! h^{3N}}$$

But $\dfrac{\int dp dq \, e^{-\beta H}}{N! h^{3N}} = \dfrac{\Delta\Gamma}{N! h^{3N}} = \sum_i \omega(E_i) = $ No. of states accessible to the ensemble

E_i is the energy eigen value of the ith state that lies between E and E+δE.

All these are equally probable and these are characterized by canonical distribution function $e^{-\beta E}$, and all these states are closely spaced

Therefore

$$Q_N(v, T) \equiv \frac{\int dp dq \, e^{-\beta H}}{N! h^{3N}} = \int_0^\infty dE \, \omega(E) e^{-\beta E}$$

we know that $S(E) = k \log \omega(E)$

$$\omega(E) = e^{S(E)/k}$$

$$\therefore \omega(E) = e^{\beta TS(E)} \quad (\because \beta = \frac{1}{kT})$$

$$\therefore \frac{1}{N! h^{3N}} \int dp \, dq \, e^{-\beta H(p, q)} = \int_0^\infty dE \, e^{\beta[TS(E) - E]} \quad \text{-------- (3.15)}$$

Where 'S' is the entropy defined in the micro canonical ensemble. Since both S and U are proportional to N, the exponent in the last integrand is enormous. We except that as $N \to \infty$ the integral receives contribution only from the neighborhood of the maximum of the integrand. The maximum of the integrand occurs at $E = \bar{E}$, where $\bar{E}$ satisfies the conditions

$$T\left(\frac{\partial S}{\partial E}\right)_{E=\bar{E}} = 1 \quad \text{---------- (3.16)}$$

$$\left(\frac{\partial^2 S}{\partial E^2}\right)_{E=\bar{E}} < 0 \quad \text{---------- (3.17)}$$

from the first condition we have

$$\left(\frac{\partial S}{\partial E}\right)_{E=\bar{E}} = \frac{1}{T} \quad \text{or} \quad \frac{\partial S}{\partial \bar{E}} = \frac{1}{T} \quad \text{---------- (3.17(a))}$$

previously we defined $\dfrac{\partial S}{\partial U} = \dfrac{1}{T}$ ---------- (3.17(b))

by comparing the equations 3.17(a) and (b) we can say that $\bar{E} = U$

The first condition implies that $\bar{E} = U$, the internal energy. Next we note that

$$\left(\frac{\partial^2 S}{\partial E^2}\right)_{E=\bar{E}} = \left(\frac{\partial}{\partial E} \frac{1}{T}\right)_{E=\bar{E}} = -\frac{1}{T^2}\left(\frac{\partial T}{\partial E}\right)_{E=\bar{E}} = -\frac{1}{T^2 C_V} \quad \text{---------- (3.18)}$$

Thus the condition (3.17) is satisfied if $C_V > 0$, which is true for physical systems.

In order to evaluate the integral of equation 3.15, let us expand the power of exponent in (3.15) about $E = \bar{E}$ in Taylor's form.

$$[TS(E) - E]_{E=\bar{E}} = [TS(\bar{E}) - \bar{E}] + (E - \bar{E})\left[\frac{\partial}{\partial E}(TS(E) - E)\right]_{E=\bar{E}}$$

$$+ \frac{1}{2}(E - \bar{E})^2 \left\{ \frac{\partial^2}{\partial E^2}(TS(E) - E) \right\}_{E=\bar{E}} + \cdots$$

$$= [TS(\bar{E}) - \bar{E}] + [E - \bar{E}]\left[T\frac{\partial S(E)}{\partial \bar{E}} - 1 \right] + \frac{1}{2}(E - \bar{E})^2 T\left\{ \frac{\partial^2 S(E)}{\partial E^2} \right\}_{E=\bar{E}} + \cdots$$

$$= [TS(\bar{E}) - \bar{E}] + \frac{1}{2}(E - \bar{E})^2 T\left\{ \frac{\partial^2 S(E)}{\partial E^2} \right\}_{E=\bar{E}} + \cdots$$

By neglecting higher order terms,

$$[TS(E) - E]_{E=\bar{E}} \approx [TS(\bar{E}) - \bar{E}] + \frac{1}{2}(E - \bar{E})^2 T\left(-\frac{1}{T^2 C_V} \right)$$

$$\approx [TS(\bar{E}) - \bar{E}] - \frac{1}{2}(E - \bar{E})^2 \frac{1}{T C_V}$$

$$\approx [TS(U) - U] - \frac{1}{2T C_V}(E - U)^2 + \cdots \qquad \text{------- (3.19) (since } \bar{E} = U)$$

Hence

$$\frac{1}{N! h^{3N}} \int dp\, dq\, e^{-\beta H(p,\, q)} \approx e^{\beta(TS - U)} \int_0^\infty dE\, e^{-(E-U)^2/2kT C_V} \quad \text{------- (3.20)}$$

Showing that in the canonical ensemble the distribution in energy is a Gaussian distribution centered about the value $E = U$ with a width equal to

$$\Delta E = \sqrt{2kT^2 C_V} \quad \text{------- (3.21)}$$

Since $U \propto N$ and $C_V \propto N$, $\Delta E/U$ is negligibly small as $N \to \infty$ for macroscopic system. As $N \to \infty$ the Gaussian distribution approaches a δ function. Finally, let us perform the integral in (3.20). It is elementary.

$$\int_0^\infty dE\, e^{-(E-U)^2/2kT C_V} = \int_{-U}^\infty dx\, e^{-x^2/2kT C_V} \approx \int_{-\infty}^{+\infty} dx\, e^{-x^2/2kT^2 C_V} = \sqrt{2\pi kT^2 C_V}$$

Therefore

$$\frac{1}{N! h^{3N}} \int dp\, dq\, e^{-\beta H(p,\, q)} \approx e^{\beta(TS - U)} \sqrt{2\pi kT^2 C_V} \quad \text{------- (3.22)}$$

we know that partition function

$$Q_N(V, T) \approx e^{\beta(TS - U)} \sqrt{2\pi kT^2\, C_V}$$

But according to thermodynamics of the system, we have

$$e^{-\beta A(V, T)} \approx e^{\beta(TS - U)} \sqrt{2\pi kT^2\, C_V}$$

$$-\beta A(V, T) \approx \beta(TS - U) + \frac{1}{2} \log\, [2\pi kT^2 C_V]$$

$$A(V, T) \approx (U - TS) - \frac{1}{2\beta}\, \log\, [2\pi kT^2 C_V]$$

$$\therefore A \approx (U - TS) - \frac{1}{2}\, kT \log C_V \quad \text{------- (3.23)}$$

Since $C_V \propto N$ the last term is negligible when $N \rightarrow \infty$. In that limit we have exactly $A = U - TS$. Statement (3.23) shows that the entropy as defined in the canonical and micro-canonical ensemble differs only by terms of the order of log N.

We have shown that almost all systems in the canonical ensemble have the same energy of a system at the given temperature T. The reason for this is easy to see, both mathematically and physically.

In the canonical ensemble we distribute systems in $\Gamma-$ space according to the density function $\rho(p, q) = \exp[-\beta H(p, q)]$, which is represented in fig 3.1. The density of points fall off exponentially as we go away from the origin of $\Gamma-$ space. The distribution in energy is obtained by "counting" the number of points on energy surfaces. As we go away from the origin, the energy increases and the area of the energy surface increases of the peak is due to the rapidity with which the area of the energy surface increases as E increases. For an N – body system this area increases like e^E, where $E \propto N$.

Energy surface
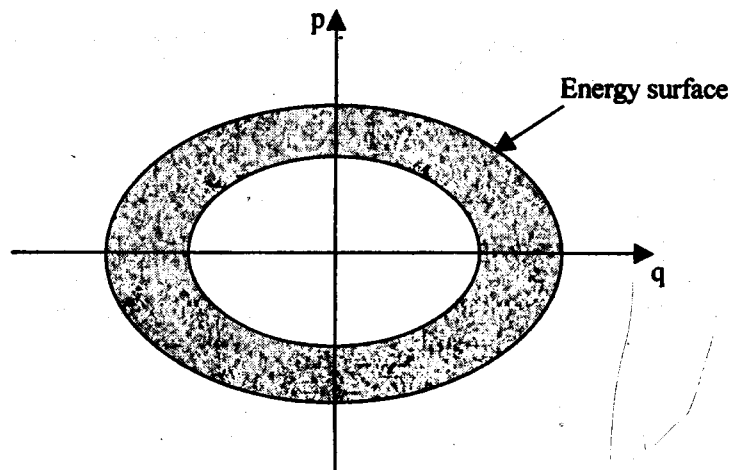
From a physical point of view, a micro-canonical ensemble must be equivalent to a canonical ensemble, otherwise we would seriously doubt the utility of either. A macroscopic substance has the extensive property, i.e, any part of the substance has the same thermodynamic property as the whole substance. If we consider a piece of substance isolated from every thing, it is still true that any part of the substance must be in equilibrium with the rest; and the rest serves as a heat reservoir that defines a temperature for the part on which we focus our attention. Therefore the whole substance must have a well – defined temperatures.

We have seen earlier that in the micro-canonical ensemble it matters little whether we take the entropy to be k times the logarithm of the density of states at the energy E, the number of states with energies between E, E + S, or all the states with energy below E. In all these cases we arrive at the same thermodynamic behavior. Now we see that it matters little whether we specify the energy of the system or the temperature of the system, for specifying one fixes the other, and we find the same thermodynamic behavior in both cases. All these examples illustrate the insensitivity of thermodynamic results to methods of derivation. The reasons behind this insensitivity are, in all cases, the facts that

(a) density of states $\propto e^E$

(b) $E \propto N$

(c) $N \rightarrow \infty$

On these facts depends the validity of statistical mechanics.

Summary:

We know that is the micro canonical the macro state of the systems is defined by the fixed number of particles N, fixed volume V and fixed energy E. However for most physical systems, the concept of fixed energy does not appear satisfactory as the total energy E of a system is hardly ever measured. Therefore, it is convenient to keep a fixed temperature T of the system in place of a fixed energy E. For this purpose, the system is placed is contact with an appropriate heat reservoir (or thermostat). Now, an infinitely large no of mental copies of a given system in thermal contact with a heat reservoir, form an ensemble in which the macrostate of the system is defined by fixed parameters T, V, and N. such an ensemble is called a canonical ensemble. The volume in Γ- space occupied by the canonical distribution

function, $(e^{-\beta H})$ is called partition function $Q_N(v, T)$. Since the partition function is the dimensionless quantity the volume in Γ-space is divided by h^{3N} (since h^{3N} has also volume dimension). Further with regard to the counting of the states $\frac{\Delta\Gamma}{h^{3N}}$ should be divided by N! in accordance with the rule of "correct Boltzmann counting". In a canonical ensemble all thermodynamic functions may be obtained from helmholtz free energy A(v, T) by using the Maxwell's relations in thermodynamics.

For a macroscopic system U or $<H> \alpha$ N, and $C_V \alpha$ N. The mean square fluctuation in energy is normal and as $N \to \infty$, all most all systems in the ensemble have the same energy $<H>$ which is the internal energy. Therefore, the canonical ensemble is equivalent to the microcanonical ensemble. We studied the energy fluctuation by calculating the partition function and here also we have seen that almost all the systems in the canonical ensemble have the same energy called internal energy at a given temperature T.

Key terminology:

1.Thermostat: It is an instrument maintaining a constant temperature by the use of a device that cuts off the supply of heat when the required temperature is exceeded and automatically restores the supply when the temperature falls below that required.

2. Macroscope: Denoting large in contrast to micro – small

3. Mean square fluctuations: In course of time the quantity H varies, fluctuating above its mean value $<H>$. Since the deviation H - $<H>$ is alternately positive and negative the average absolute function $<\Delta H> = <H - <H>>$ is zero and is not suitable. Hence we prefer to use the mean square fluctuation defined by

$$< (\Delta H)^2> \text{ or } <(U - H)^2> = <H^2> - <H>^2.$$

This quantity is called the dispersion of the physical quantity H. It is defined by the difference between the mean square of the quantity and the square of its mean.

Self – assessment questions:

1. Obtain an expression for ensemble density by focusing your attention on a small sub system which is in thermal contract with the larger sub system.

2. Define the partition function $Q_N(V,T)$ in the canonical ensemble and explain its significance.

3. Discuss the energy fluctuations in the canonical ensemble.

4. Establish the equivalence between the canonical and microcanonical ensembles by calculating the partition function.

Reference:

1. Statistical Mechanics by Gupta & Kumar, K Nath & Co.
2. Statistical Mechanics by Kerson Huang, Wiley Eastern Ltd.
3. Statistical Mechanics by M. Eisner and B.K. Agarwal, Wiley Eastern Ltd.,
4. Statistical Mechanics: Theory and Applications by S.K. Sinha, Tata McGraw Hill Co, Ltd.,

Statistical Mechanics Part

UNIT -III

Lesson 4

GRAND CANONICAL ENSEMBLE

Objectives: To

1. explain the grand canonical ensemble in detail
2. discuss the density fluctuations in the grand canonical ensemble
3. establish the equivalence between the grand canonical and canonical ensemble

Structure:

4.1 Grand canonical ensemble

4.2 Density fluctuations in the grand canonical ensemble

4.3 Equivalence of the canonical ensemble and the grand canonical ensemble

4.1 GRAND CANONICAL ENSEMBLE:

Earlier we studied the canonical ensemble in which the number of particles of the system is fixed, but the energy E is variable. However, there are number of physical as well as chemical systems, which are either exchanging particles with other systems or generating an unlimited number of particles of a definite type. For example, electromagnetic radiation in a container may be treated as a gas of photons, the number of particles of which is not definite, since photons are emitted and absorbed continuously by the walls of container. The system in which the number of particles N ad energy E vary is called an open system. We may consider the system immersed in a heat reservoir with which it may exchange both energy and particles. A large number of mental copies of the given system in thermal contact with a heat reservoir form an ensemble, which is called a grand canonical ensemble. In this case, the distribution function depends not only on the energy of the quantum state but also on the number N of the particles in the system.

Therefore in a grand canonical ensemble, both energy E and number N of particles are variable. However the system in contact with the reservoir forms a composite system whose energy and number of particles are fixed. This composite system is described by the microcanonical ensemble.

Although the canonical and the microcanonical ensemble yield equivalent results, it may be a⋯ conceptually the canonical ensemble corresponds more closely to physical situations. In experimen⋯ never deal with a completely isolated systems, nor do we ever directly measure the total energy of a

macroscopic system. We usually deal with systems with a given temperature - a parameter that we can control in the experiments.

By the same attitude, we would consider it unphysical to specify the number of particles of a macroscopic system, for, that is never precisely known. All we can find out from experiments is the average number of particles. This is the motivation for introducing the grand canonical ensemble, in which the systems can have any number of particles, with the average number determined by conditions external to the system. This is analogous to the situation in the canonical ensemble, where, the average energy of a system is determined by the temperature of the heat reservoir with which it is in contact.

Assume that the Γ - space for the grand canonical ensemble is spanned by all the canonical momenta and coordinates of systems with 0, 1, 2, ... number of particles. The density function describing the distribution of representative points in Γ - space is denoted by ρ (p, q, N), which gives the density of points representing systems with N particles with the momenta and coordinates (p, q). To find ρ (p, q, N) we consider the canonical ensemble for a system with N particles, volume V, and temperature T, but we focus out attention as a small sub volume V_1 of the system. The density ρ (p_1, q_1, N_1) is proportional to the probability that in the sub volume V_1 there are N_1 particles with the coordinates (p_1, q_1).

Let $N_2 = N - N_1$ and $V_2 = V - V_1$. We assume that

$N_2 >> N_1$

$V_2 >> V_1$

If there are N_1 particles in V_1, there must be N_2 particles in the remaining V_2. Hence, neglecting molecular interactions across the surface separating V_2 and V_1, we must have

$$\rho\,(p_1, q_1, N_1) \propto e^{-\beta H(p_1,q_1,N_1)} \int_{V_2} dp_2\, dq_2\, e^{-\beta H(p_2,q_2,N_2)} \quad \text{---------- (4.1)}$$

Where the integral in equation (4.1) extends over all p_2, but only over such values of q_2 as to keep the N_2 particles always in the volume V_2. The Hamiltonians $H(p_1, q_1, N_1)$ and $H(p_2, q_2, N_2)$ have the same functional form but refer respectively to N_1 particles and N_2 particles. We arbitrarily choose the proportionality constant in equation (4.1) so as to give

$$\rho\,(p_1, q_1, N_1) = \frac{N!}{N_1! N_2!} \frac{e^{-\beta H(p_1,q_1,N_1)} \int_{V_2} dp_2 dq_2 e^{-\beta H(p_2,q_2,N_2)}}{\int_V dp\, dq\, e^{-\beta H(p,q,N)}} \quad \text{---------- (4.2)}$$

Multiply both numerator and denominator of equation (4.2) by $h^{3N_2} h^{3N}$

$$\rho\,(p_1, q_1, N_1) = \frac{N!}{N_1!N_2!} \; \frac{e^{-\beta H(p_1,q_1,N_1)} \displaystyle\int_{V_2} e^{-\beta H(p_2,q_2,N_2)}\,dp_2\,dq_2\,h^{3N_2}h^{3N}}{h^{3N_2}h^{3N}\displaystyle\int_{V} e^{-\beta H(p,q,N)}\,dp\,dq} \qquad\text{------- (4.2(a))}$$

Hence, equation (4.2(a)) can also be written as

$$\rho\,(p_1, q_1, N_1) = \frac{Q_{N_2}(V_2, T)}{Q_N(V, T)} \; \frac{e^{-\beta H(p_1,q_1,N_1)}}{N_1!\,h^{3N_1}} \qquad\text{------- (4.3)}$$

Where Q_N (V, T) is the partition function defined as $\dfrac{\int dpdq\; e^{-\beta H(p,q)}}{N!h^{3N}}$ in the canonical ensemble (equation 3.6)

Using equation (3.7) we can write

$$\frac{Q_{N_2}(V_2\,T)}{Q_N(V, T)} = e^{-\beta[A(N_2,V_2,T)-A(N,V,T)]} = e^{-\beta[A(N-N_1,V-V_1,T)-A(N,V,T)]} \qquad\text{------- (4.5)}$$

where A(N,V,T) is the Helmholtz free energy. Since $N \gg N_1$ and $V \gg V_1$.

We may use the approximation

$$A(N - N_1, V - V_1, T) - A\,(N, V, T) \approx -N_1\,\mu + V_1\,P \qquad\text{------- (4.6)}$$

Where μ and P are respectively the chemical potential and the pressure of the part of the system external to the small volume V_1:

$$\mu = \left[\frac{\partial A(N_2, V, T)}{\partial N_2}\right]_{N_2 = N} \qquad\text{------- (4.7)}$$

$$P = -\left[\frac{\partial A(N, V_2, T)}{\partial V_2}\right]_{V_2 = V} \qquad\text{------- (4.8)}$$

we introduce the fugacity, defined by

$$z = e^{\beta\mu} \qquad\text{------- (4.9)}$$

Substituting (4.9) and (4.6) into (4.5), and then substituting (4.5) into (4.3) we obtain

$$\rho\,(p, q, N) = \frac{z^N}{N!h^{3N}}\,e^{-\beta PV - \beta H(p,q)} \qquad\text{------- (4.10)}$$

Where the subscript 1 identifying the volume under consideration has been omitted because the system external to the volume can now be forgotten, apart from the information that it has the temperature T, pressure P, and chemical potential μ. We now allow the system external to the volume under consideration to become infinite in size. Then the range of N in (4.10) becomes

$\cap \leq N < \infty$

The thermodynamic functions for the volume under consideration may be found as follows. First of all, the internal energy shall be the ensemble average of H (p, q). Second, the temperature, pressure, and chemical potential shall be respectively equal to T,P, μ. To show that this is a correct recipe, it suffices to remind ourselves that thermodynamics has been derived from the canonical ensemble. It is an elementary thermodynamic exercise to show that if a system is in equilibrium any part of the system must have the some T, ρ, μ as any other part; but this is the desired result.

To obtain a convenient formal recipe for finding all the thermodynamic functions we define the "grand partition" function as follows:

$$Z(z, V, T) \equiv \sum_{N=0}^{\infty} z^N Q_N(V, T) \quad \text{--------- (4.11)}$$

Which in principle can be calculated from a knowledge of the Hamiltonian. Integrating both sides of (4.10) overall (p, q) for a given N, and then summing N from 0 to ∞, we find that

$$\sum_{N=0}^{\infty} \int \rho(p, q, N) dp dq = e^{-\beta pV} \sum_{N=0}^{\infty} z^N \frac{\int dp \, dq \, e^{-\beta H}}{N! h^{3N}}$$

$$\sum_{N=0}^{\infty} \int \rho(p, q, N) dp dq = e^{-\beta pV} \sum_{N=0}^{\infty} z^N Q_N(V, T) = e^{-\beta PV} Z(z, V, T) \quad \text{----------(4.11(a))}$$

$$\beta PV = \log Z(z, V, T) \qquad [\text{since} \sum_{N=0}^{\infty} \int \rho(p, q, N) dp \, dq = 1]$$

$$\frac{PV}{KT} = \log Z(z, V, T) \quad \text{----------- (4.12)}$$

Thus the grand partition function directly gives the pressure as a function of z, V and T. The average number N of particles in the volume V is by definition the ensemble average

$$N = \frac{\sum_{N'=0}^{\infty} N' z^{N'} Q_{N'}(V, T)}{\sum_{N'=0}^{\infty} z^{N'} Q_{N'} V, T)} = z \frac{\partial}{\partial z} \log Z(z, V, T) \quad \text{--------- (4.13)}$$

The equation of state, which is the equation expressing P as a function of N, V, and T, is obtained by eliminating z between (4.12) and (4.13).

All other thermodynamic functions may be obtained from the internal energy:

$$U = -\frac{\partial}{\partial \beta} \log Z(z, V, T) \quad \text{---------- (4.14)}$$

After eliminating z with the help of (4.13), U becomes a function of N, V, and T. We can then use the formulae

$$C_V = \left(\frac{\partial U}{\partial T}\right)_V$$

$$S = \int_0^T dT \frac{C_V}{T}$$

$$A = U - TS$$

Alternatively, all thermodynamic functions may be found from the Helmholtz free energy, which can be shown to be directly obtainable from log Z through the formula

$$A = NkT \log z - kT \log Z(z, V, T) \quad \text{------} \quad (4.15)$$

Again it is necessary to eliminate z with the help of (4.13) in order to obtain A as a function of N, V, and T. The justification for (4.15) is given in the following two sections.

4.2 DENSITY FLUCTUATIONS IN THE GRAND CANONICAL ENSEMBLE:

We begin a study of the equivalence between the grand canonical ensemble and the canonical ensemble. These two are trivially equivalent to each other, if almost all systems in the grand canonical ensemble have the same number of particles. Since all systems have exactly the same volume, this means that the fluctuation of density is small. We first find the condition under which the fluctuation of density is small.

The probability that a system in the grand canonical ensemble has N' particles is given by

$$W(N') = z^{N'} Q_{N'}(V, T) = e^{\beta \mu N' - \beta A(N', V, T)} \quad \text{------} \quad (4.16)$$

Where A(N, V, T) is the Helmholtz free energy calculated from the canonical ensemble for N particles in volume V at temperature T. For the fluctuation of density in the grand canonical ensemble to be small, it is necessary and sufficient that W(N') be essentially zero except in the neighborhood of some points N' = N, where W(N') should have a sharp maximum. That is, we require that there be a value N for which

$$\left[\frac{\partial A(N', V, T)}{\partial N'}\right]_{N'=N} = \mu \quad \text{------} \quad (4.17)$$

$$\gamma \equiv \left[\frac{\partial^2 A(N', V, T)}{\partial N'^2}\right]_{N'=N} > 0 \quad \text{------} \quad (4.18)$$

The first of these conditions requires that the system have the same chemical potential μ as the external system. It is identical with the condition (4.13). To find the meaning of the second condition we first express γ in terms of measurable quantities in the following manner.

Since the Helmholtz free energy is an extensive quantity it may be written in the form.

$$A(N', V, T) = N' \, a(v') \text{ --------- (4.19)}$$

Where $v' = V/N'$, and where the dependence on the temperature is understood. Then

$$\frac{\partial A}{\partial N'} = a(v') - v' \frac{\partial a(v')}{\partial v'}$$

$$\frac{\partial^2 A}{\partial N'^2} = \frac{1}{N'} v'^2 \frac{\partial^2 a(v')}{\partial v'^2} \text{ --------- (4.20)}$$

on the other hand the pressure of the system is

$$P(v') = -\frac{\partial a(v')}{\partial v'}$$

Hence $\dfrac{\partial P(v')}{\partial v'} = -\dfrac{\partial^2 a(v')}{\partial v'^2}$ --------- (4.21)

Comparing (4.21) with (4.20) and (4.18) we obtain

$$\gamma \equiv \left(\frac{\partial^2 A}{\partial N'^2}\right)_{N'=N} = -\frac{v^2}{N} \frac{\partial P(v)}{\partial v} \text{ --------- (4.22)}$$

Where $v = \dfrac{V}{N}$. thus (4.18) is the same as the requirement

$$\frac{\partial P(v)}{\partial v} < 0 \text{ --------- (4.23)}$$

Experimentally the equation of state of a substance is always such that $\dfrac{\partial P(v)}{\partial v} \leq 0$. Hence (4.23) is fulfilled.

We now expand W(N') about N:

$$[W(N')]_{N'=N} = W(N) + \frac{N'-N}{1!}\left[\frac{\partial}{\partial N'}e^{\beta(\mu N'-A)}\right]_{N'=N} + \frac{(N'-N)^2}{2!}\left[\frac{\partial^2}{\partial N'^2}e^{\beta(\mu N'-A)}\right]_{N'=N} + \text{------}$$

$$= W(N) + (N'-N)\left[e^{\beta(\mu N'-A)}\beta\left(\mu - \frac{\partial A}{\partial N'}\right)\right]_{N'=N} +$$

$$\frac{(N'-N)^2}{2!}\frac{\partial}{\partial N'}\left[e^{\beta(\mu N'-A)} \cdot \beta\left(\mu - \frac{\partial A}{\partial N'}\right)\right]_{N'=N} + \text{-----}$$

$$= W(N) + (N'-N)e^{\beta(\mu N'-A)} \cdot \beta(\mu - \mu) + \frac{(N'-N)^2}{2!}\left[e^{\beta(\mu N'-A)}\beta^2\left(\mu - \frac{\partial A}{\partial N'}\right)^2 + e^{\beta(\mu N'-A)}\beta\left(-\frac{\partial^2 A}{\partial N'^2}\right)\right]_{N'=N} + \text{--}$$

$$= W(N) + 0 + \frac{(N'-N)^2}{2!} \left[0 - e^{\beta(\mu N'-A)} \beta \cdot \frac{\partial^2 A}{\partial N'^2} \right]_{N'=N} + \text{----}$$

$$= W(N) - \frac{(N'-N)^2}{2!} \left[e^{\beta(\mu N'-A)} \beta \cdot \frac{\partial^2 A}{\partial N'^2} \right]_{N'=N} + \text{----}$$

$$= W(N) - \frac{(N'-N)^2}{2!} W(N) \beta \gamma + \text{----}$$

$$(\because e^{\beta(\mu N-A)} = W(N) \text{ and } \left(\frac{\partial^2 A}{\partial N'^2} \right)_{N'=N} = \gamma)$$

$$= W(N) \left[1 - \frac{1}{2} \beta \gamma (N'-N)^2 + \text{-----} \right]$$

By the definition of exponential function and by neglecting the higher order terms the above equation can be written as

$$\therefore [W(N')]_{N'=N} \approx W(N) \, e^{-\frac{1}{2}\beta\gamma(N'-N)^2} \quad \text{--------} \quad (4.24)$$

This leads to a Gaussian distribution of N', centered about N, with a width equal to

$$\Delta N = \sqrt{\frac{2}{\beta\gamma}} = \sqrt{\frac{2KTN}{v^2(-\partial P/\partial v)}} \quad \text{--------} \quad (4.25)$$

Thus $\dfrac{\Delta N}{N} \to 0$ as $N \to \infty$, if (4.23) is fulfilled.

We may also directly calculate the mean square fluctuation of N' and find that

$$\frac{\sum\limits_{N'=0}^{\infty}(N'-N)^2 W(N')}{\sum\limits_{N'=0}^{\infty} W(N')} = \frac{KTN}{v^2\left(-\dfrac{\partial P}{\partial v}\right)} \quad \text{--------} \quad (4.26)$$

We have seen that if $\dfrac{\partial P}{\partial v} < 0$ then almost all systems in the grand canonical ensemble have the same

number of particles N. Then the grand canonical ensemble is trivially equivalent to the canonical ensemble for N particles. We must then have

$$Z(z, V, T) \approx z^N Q_N(V, T) \quad \text{--------} \quad (4.27)$$

from equation 4.27, we obtain $-\beta A(N, V, T) = \log Z(z, V, T) - N \log z$

and $\qquad N = z\dfrac{\partial}{\partial z} \log Z(z, V, T)$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ ———— (4.28)

Eliminating z between the two equations of (4.28) yields the Helmholtz free energy, from which all thermodynamic functions can be obtained. In particular we recover the equation of state in the form (4.12) by using the formula $P = -\dfrac{\partial A}{\partial v}$ from the canonical ensemble. There is still the question whether there always exist a z such that the second equation of (4.28) gives any desired value of N. We post pone the question until the next section, where it is answered in the affirmative.

Since we derive the grand canonical ensemble from the canonical ensemble by focusing our attention on a volume within the system, the grand canonical ensemble cannot contain more information than the canonical ensemble. The grand canonical ensemble does, however, make it more convenient for us to consider density fluctuations. These fluctuations give rise to physically observable effects, eg., the fluctuation scattering of light. The formula (4.26) indicates that near the critical point of a gas, where $\dfrac{\partial P}{\partial v} = 0$, the density fluctuations become abnormally large. This is experimentally borne out by the phenomenon of critical opalescence.

4.3 EQUIVALENCE OF THE CANONICAL ENSEMBLE AND THE GRAND CANONICAL ENSEMBLE:

To complete our investigation of the equivalence between the canonical and the grand canonical ensemble it is necessary to consider values of v for which $\dfrac{\partial P}{\partial v} = 0$. It will be shown that in such cases the function W(N) gives in (4.16) will no longer have a sharp maximum: the equation of state as given by the recipe in the grand canonical ensemble nevertheless still agrees with that given by the recipe in the canonical ensemble. In this sense the two ensembles are always equivalent.

Physically the values of 'v' for which $\dfrac{\partial P}{\partial v} = 0$ correspond to the transition region of a first – order phase transition. In this region (4.26) leads us to expect that the fluctuations of density in a given volume of the system will be large. This is also expected physically, for in such a region the system is composed of two or more phases of different densities. Therefore the number of particles in any given volume can have a whole range of values, depending on the amounts of each phase present. At the critical point of a gas – liquid system fluctuation in density are also expected to be large, because throughout the system molecules are

spontaneously forming large clusters and breaking up. It is clear that under these conditions the grand canonical ensemble must continue to yield thermodynamic predictions that are in agreement with those obtained by the canonical ensemble. Otherwise the validity of either as a description of matter would be in doubt, for it is a basic experimental fact that we can obtain the same thermodynamic information whether we look at the whole system or at only a sub volume of the system.

The mathematical questions that we try to answer are as follows. Suppose Q_N (V, T) is given, and we wish to calculate

$$Z(z, V, T) = \sum_{N'=0}^{\infty} z^{N'} \, Q_{N'}(V, T) \quad \text{------ (4.29)}$$

For given values of z, V and T.

(a) For a given value of z is the following true for some N

$$Z(z, V, T) \approx Z^N Q_N(V, T) \quad \text{------ (4.30)}$$

(b) Does there always exist a value of z for which N has any given positive value?

The answers are obviously no, if Q_N (V, T) is any function of N, V, T. We are only interested, however, in the answers when Q_N (V, T) is the partition function of a physical system. Thus we must first make some assumptions about Q_N (V, T).

In order to incorporate the salient features of a physical system into our considerations, any yet keep the mathematics simple, we assume that we are dealing with a system.

       (a) Whose molecules interact through an intermolecular potential that contains a hard – sphere repulsion of finite diameter plus a finite potential of finite range, and

       (b) Whose Helmholtz free energy has the form

$$A(N,V) = -\frac{1}{\beta} \log Q_N(V) = -\frac{V}{\beta} f(v) \quad \text{------(4.31)}$$

Where $v = \dfrac{V}{N}$, $\beta = \dfrac{1}{kT}$, and f(v) is finite. The temperature shall be fixed throughout our discussions and will not be displayed unless necessary. The function f(v) is related to the pressure P(v) of the canonical ensemble by

$$f(v) = \frac{1}{v} \int_{v_0}^{v} dv' \, \beta P(v') \quad \text{------ (4.32)}$$

Where the integration is carried out along an isotherm and v_0 is an arbitrary constant corresponding to an arbitrary additive constant in the Helmholtz free energy.

Derivation of (4.32) is as follows:

We know that $A(N,V) = N'a(v')$

$$a(v') = \frac{1}{N'} A(N,V)$$

and $P(v') = -\frac{\partial a(v')}{\partial v'} = -\frac{1}{N'} \frac{\partial A}{\partial v'}$

$$\frac{\partial A}{\partial v'} = -N'P(v')$$

Integrating the above equation we get,

$$A(N,V) = -N' \int_0^v P(v') \, dv' + \text{constant, where } v = \frac{V}{N} \quad \text{------- (4.32(a))}$$

From (4.31), and 4.32(a), we have $-\frac{V}{\beta} f(v) = -N' \int_0^v P(v') \, dv' + \text{constant} \quad \text{-------- (4.32(b))}$

$$f(v) = \frac{N'}{V} \int_0^v \beta P(v') \, dv' - \frac{\beta}{V} \text{constant} \quad \text{-------- (4.32(c))}$$

$$f(v) = \frac{1}{v} \left[\int_0^{v_0} \beta P(v') \, dv' + \int_{v_0}^v \beta P(v') \, dv' \right] - \frac{\beta}{V} \text{constant}$$

$$f(v) = \frac{1}{v} \int_{v_0}^v \beta P(v') \, d(v')$$

In equation 4.31(c), w introduce v_0 in the integral between the range 0 and v such that the integral in the range between 0 and v_0 gives an arbitrary additive constant to nullify the existing constant $-\frac{\beta}{V}$ constant

(c) We further assume that f(v) is such that

$$\frac{\partial P}{\partial v} \leq 0 \quad \text{--------(4.33)}$$

This immediately implies that $\dfrac{\partial^2 f(v)}{\partial \left(\dfrac{1}{v}\right)^2} \leq 0 \quad \text{--------- (4.34)}$

(How that $\dfrac{\partial P}{\partial v} \leq 0$ immediately implies that $\dfrac{\partial^2 f(v)}{\partial \left(\dfrac{1}{v}\right)^2} \leq 0$ is shown below)

$$A(N, V) = N\,a(v) = -\frac{V}{\beta}\,f(v)$$

$$a(v) = -\frac{v}{\beta}f(v)$$

$$P(v) = -\frac{\partial a(v)}{\partial v} = \frac{\partial}{\partial v}\left[\frac{v}{\beta}f(v)\right]$$

$$= \frac{1}{\beta}\left[f(v) + v\frac{\partial f(v)}{\partial v}\right]$$

$$= \frac{1}{\beta}\left[f(v) + v\frac{\partial f(v)}{\partial(1/v)}\cdot\frac{\partial(1/v)}{\partial v}\right]$$

$$= \frac{1}{\beta}\left[f(v) - v\cdot\frac{1}{v^2}\frac{\partial f(v)}{\partial(1/v)}\right]$$

$$P(v) = \frac{1}{\beta}\left[f(v) - \frac{1}{v}\frac{\partial f(v)}{\partial(1/v)}\right]$$

$$\frac{\partial P(v)}{\partial v} = \frac{\partial P}{\partial(1/v)}\frac{\partial(1/v)}{\partial v} \quad;\text{ Here } \frac{\partial}{\partial v}\cdot\left(\frac{1}{v}\right) = -\frac{1}{v^2}$$

$$= -\frac{1}{v^2}\cdot\frac{1}{\beta}\left[\frac{\partial}{\partial(1/v)}\left\{f(v) - \frac{1}{v}\frac{\partial f(v)}{\partial(1/v)}\right\}\right]$$

$$= -\frac{1}{\beta v^2}\left[\frac{\partial f(v)}{\partial(1/v)} - \frac{\partial f(v)}{\partial(1/v)} - \frac{1}{v}\frac{\partial^2 f(v)}{\partial(1/v)^2}\right]$$

$$\frac{\partial P(v)}{\partial v} = \frac{1}{\beta v^3}\frac{\partial^2 f(v)}{\partial(1/v)^2} \le 0$$

With these assumptions the grand partition function may be simplified in the following manner

$$Z(z, V) = \sum_{N=0}^{\infty} \exp\left[V\phi\left(\frac{V}{N}, z\right)\right] \quad\text{------- (4.35)}$$

Derivation of this expression is as follows:

$$Z(z, V, T) = \sum_{N=0}^{\infty} z^N Q_N(V, T)$$

$$= \sum_{N=0}^{\infty} e^{\beta\mu N} e^{Vf(v)} \quad (\because z = e^{\beta\mu}, \text{ from equation 4.31 we have } Q_N(v) = e^{Vf(v)}$$

$$= \sum_{N=0}^{\infty} e^{N \log z} e^{V f(v)} \quad (\because \beta \mu = \log z)$$

$$= \sum_{N=0}^{\infty} e^{[V/v \log z + V f(v)]}$$

$$Z(z, V) = \sum_{N=0}^{\infty} e^{V \cdot \phi(v, z)}$$

Where z is an arbitrary fixed number and

$$\phi(v, z) \equiv f(v) + \frac{1}{v} \log z \quad \text{------- (4.36)}$$

Using (4.32) we obtain

$$\phi(v, z) = \frac{1}{v} \log z + \frac{1}{v} \int_{v_0}^{v} dv' \, \beta P(v') \quad \text{------- (4.37)}$$

By (4.34), we have $\dfrac{\partial^2 \phi}{\partial (1/v)^2} \leq 0$, or

$$\frac{\partial^2 \phi}{\partial v^2} + \frac{2}{v} \frac{\partial \phi}{\partial v} \leq 0 \quad \text{-------(4.38)}$$

We now calculate the grand partition function. For a fixed volume V the partition function $Q_N(V)$ vanishes whenever

$$N > N_0 (V)$$

Where N_0 (V) is the maximum number of particles that can be accommodated in the volume V, such that no two particles are separated by a distance less than the diameter of the hard sphere in the inter particle potential. Therefore Z(z, V) is a polynomial of degree N_0 (V). For large V it is clear that

$$N_0 (V) = aV \quad \text{--------- (4.39)}$$

Where a is a constant. Let the largest value among the terms in this polynomial be exp $[V\phi_0(z)]$, where

$$\phi_0(z) = Max[\phi(V/N, z)] \quad (N = 0, 1, 2, 3, \ldots \ldots) \quad \text{-------(4.40)}$$

Then the following inequality holds:

$$e^{V\phi_0(z)} \leq Z(z, V) \leq N_0(V) \, e^{V\phi_0(z)}$$

Using (4.39) we obtain

$$e^{V\phi_0(z)} \leq Z(z, V) \leq aV \, e^{V\phi_0(z)}$$

or $\phi_0(z) \leq \dfrac{1}{V} \log Z(z, V) \leq \phi_0(z) \dfrac{\log(aV)}{V} \quad \text{------- (4.41)}$

Therefore

$$\underset{V\to\infty}{Lt}\ \frac{1}{V}\ \log Z(z,\ V) = \phi_o(z) \ \text{------} \ (4.42)$$

This shows that (4.27) is true.

Let $\bar{v}$ be a value of v at which $\phi(v, z)$ assumes its largest possible value. Since $\phi(v, z)$ is differentiable, $\bar{v}$ is determined by the conditions

$$\left(\frac{\partial\phi}{\partial v}\right)_{v=\bar{v}} = 0 \ \text{------} \ (4.43)$$

$$\left(\frac{\partial^2\phi}{\partial v^2}\right)_{v=\bar{v}} \le 0 \ \text{------} \ (4.44)$$

By virtue of (4.38) the first condition implies the second. Therefore $\bar{v}$ is determined by (4.43) alone. Differentiate equation (4.37) w.r.t. v

$$\phi(v, z) = \frac{1}{v}\ \log z + \frac{1}{v}\ \int_{v_0}^{v} dv'\beta P(v')$$

$$\frac{\partial\phi}{\partial v} = -\frac{1}{v^2}\ \int_{v_0}^{v} dv'\beta P(v') + \frac{1}{v}\beta P(v) - \frac{1}{v^2}\log z$$
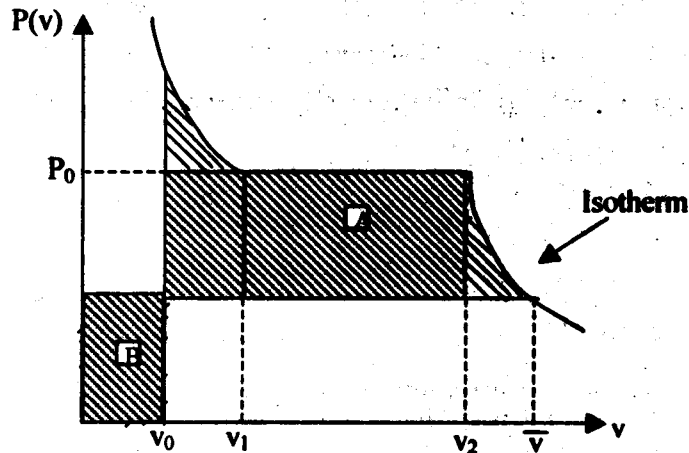
$$\left(\frac{\partial\phi}{\partial v}\right)_{v=\bar{v}} = -\frac{1}{\bar{v}^2}\ \int_{v_0}^{\bar{v}} dv'\beta P(v') + \frac{1}{\bar{v}}\beta P(\bar{v}) - \frac{1}{\bar{v}^2}\log z = 0 \ \text{------} \ (4.44(a))$$

Multiply equation 4.44(a) through out by $-\dfrac{\bar{v}^2}{\beta}$ we get

$$\int_{v_0}^{\bar{v}} dv'\beta P(v') - \bar{v}P(\bar{v}) + \frac{1}{\beta}\log z = 0$$

$$\int_{v_0}^{\bar{v}} dv'P(v') - \bar{v}P(\bar{v}) = -kT\ \log z \ \text{------} \ (4.44(b))$$

$$\text{or}\ \left[\int_{v_0}^{\bar{v}} dv'P(v') - (\bar{v} - v_0)P(\bar{v})\right] - v_0 P(\bar{v}) = -kT\ \log z \ \text{------} \ (4.45)$$

Fig 4.1 Typical Isotherm of a substance in the transition region of a first – order phase transition.

A geometrical representation of this condition is shown in fig 4.1. The volume of $\overline{v}$ is such that the difference between the area of the region A and that of the region B is numerically equal to $-kT \log z$. The result is shown in fig 4.2. It is seen that to every value of $\overline{v}$ greater than the close packing volume, there corresponds a value of z. This answers the question (b) in the affirmative, manner.
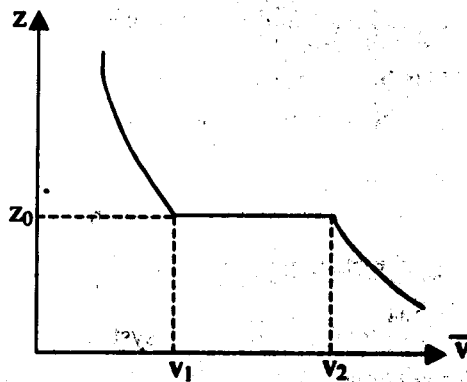
Fig 4.2 z as a function of $\overline{v}$

There is a value of z that corresponds to all the values of $\overline{v}$ lying in the interval $v_1 \leq \overline{v} \leq v_2$. This value denoted by z_0, is given by

$$\log z_0 = \beta v_1 P(v_1) - \int_{v_0}^{v_1} dv' \beta P (v') \qquad \text{------ (4.46)}$$

Summary:

The system in which both the number of particles N and energy E very is called an open system. We may consider the system immersed in a heat reservoir with which it may exchange both energy and particles. A large number of mental copies of the given system in thermal contact with heat reservoir form an ensemble which is called a grand canonical ensemble. Therefore in a grand canonical ensemble the distribution function depends not only on the energy of quantum state but also on the number N of the particles of the system.

The density ρ in the grand canonical ensemble is derived by focusing our attention on a sub volume v_1 of the system with respect to larger volume v_2 which acts as heat reservoir. To obtained a convenient formal recipe for finding all the thermodynamic functions. We have defined the grand partition function.

The density of the function in the grand canonical ensemble shows that if $\frac{\partial P}{\partial v} < 0$ almost all systems in the grand canonical ensemble have the same no. of particles N. As a result the grand canonical ensemble is trivially equivalent to the canonical ensemble. The density fluctuations in the grand canonical ensemble become abnormally large near the critical point of gas liquid system where $\frac{\partial P}{\partial v} = 0$. This is experimentally borne out by the phenomenon of critical opalescence

Key Terminology:

1. **Open system**: The system in which both the number of particles N and energy E vary is called an open system.

2. **Fugacity (z):** It is defined as $e^{\beta \mu}$. The fugacity is anologus to the pressure and it has pressure dimension.

3. **First order phase transition:** In the P –V diagram, it is seen that a certain mass of liquid is converted into gas, the total volume of system expands although P and T remain unchanged where $\frac{\partial P}{\partial v} = 0$ through out the transition. The total volume of the system changes as a relative amount of the substance in the two phases changes because the two phases have different densities. Such a transition is called first order phase transition.

Self – assessment questions:

1. Define a grand canonical ensemble, obtained the various thermodynamic functions in the grand canonical ensemble.

2. Discuss the density fluctuations in the grand canonical ensemble.

3. Prove the equivalence between the grand canonical ensemble and canonical ensemble.

Reference:

1. Statistical Mechanics by Gupta & Kumar, K Nath & Co.

2. Statistical Mechanics by Kerson Huang, Wiley Eastern Ltd.

3. Statistical Mechanics by M. Eisner and B.K. Agarwal, Wiley Eastern Ltd.,

4. Statistical Mechanics: Theory and Applications by S.K. Sinha, Tata McGraw Hill Co, Ltd.,