The principle can be alternatively stated as:
The motion of the system from instant t_1 to instant t_2 is such that the line integral

$$J = \int_{t_1}^{t_2} L dt \tag{5.3.1}$$

where L = kinetic energy - potential energy = $T - V$, is an extremum for the path of the motion.

For a system with n degrees of freedom, the Lagrangian L therefore can be expressed in the general form as

$$L \equiv L(q_1, q_2, ..., q_n; \dot{q}_1, \dot{q}_2, ..., \dot{q}_n; t)$$

In terms of the calculus of variation, we can state Hamilton's principle as

$$\delta J = \delta \int_{t_1}^{t_2} L dt = 0 \tag{5.3.2}$$

with variations zero at $t = t_1$ and $t = t_2$.

Equation (5.3.2) can, therefore, be written as

$$\delta J = \delta \int_{t_1}^{t_2} L(q_1, q_2, ..., q_n; \dot{q}_1, \dot{q}_2, ..., \dot{q}_n; t) dt = 0$$

$$\Rightarrow \qquad \delta J = \delta \int_{t_1}^{t_2} L(q_i, \dot{q}_i; t) dt = 0 \tag{5.3.3}$$

Note !

It must be remembered that the Hamilton's principle in this form is consistent with dynamical problems under holonomic constraints, *i.e.*, constraints that are dependent on coordinates only, not dependent on velocity or other quantites except time. Further, holonomic constraints must be expressed as equations, connected by equality sign; and not related through inequality. However, the principle can be extended to systems with non-holonomic constraints. Under holonomic constraints, also one can separately view the principle applicable to conservative systems, where the Lagrangian is explicitly independent of time $\left(\frac{\partial L}{\partial t} = 0\right)$; and to non-conservative system where the associated Lagrangian explicitly depends on time, *i.e.*, $\frac{\partial L}{\partial t} \neq 0$.

Comparing equation (5.3.3) with equation (5.1.21) we obtain Euler-Lagrange equation (5.1.16), *viz.*,

$$\frac{\partial f}{\partial y_i} - \frac{d}{dx}\frac{\partial f}{\partial y_i'} = 0, \qquad i = 1, 2, ..., n$$

become

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = 0, \qquad i = 1, 2, ..., n$$

$$\Rightarrow \qquad \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0, \qquad i = 1, 2, ..., n \tag{5.3.4}$$

Equations (5.3.4) are the Lagrange's equations of motion of a system of particles and the quantity $L = T - V$ is called Lagrange's function or the Lagrangian of the system. In terms of the Lagrangian L, Hamilton's principle can be stated as:

Of all the possible paths, along which a dynamical system may move from one point to another in the configuration space within a given interval of time, the actual path followed is that for which the time integral of the Lagrangian function for the system is an extremum.

5.3.1 Hamilton's varitational principle for holonomic and conservative systems

For a conservative system, the applied forces can be expressed in terms of a scalar function which is the potential energy of the system. The kinetic energy $T \equiv T(q_i(t), \dot{q}_i(t))$ for such a conservative system is a function of the generalised coordinates and generalised velocities where the time t appears implicitly, through generalised coordinates and generalised velocities. T is not a function of time explicitly. Similarly the potential energy $V \equiv V(q_i(t))$ is a function of generalised coordinates with implicit, and not explicit time dependence. The Lagrangian $L = T - V$ for a conservative system is then reduced to the functional form as

$$L \equiv L(q_i(t), \dot{q}_i(t)) = T(q_i(t), \dot{q}_i(t)) - V(q_i(t)) \equiv T(q_i, \dot{q}_i) - V(q_i)$$

The Hamilton's principle for a conservative system can therefore be written as

$$\delta \int_{t_A}^{t_B} [T(q_i, \dot{q}_i) - V(q_i)]\, dt = 0$$

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

or
$$\int_{t_A}^{t_B} \sum_j \left(\frac{\partial T}{\partial q_j} - \frac{\partial V}{\partial q_j} \right) \delta q_j dt + \int_{t_A}^{t_B} \sum_j \frac{\partial T}{\partial \dot{q}_j}\delta \dot{q}_j dt = 0$$

As the delta variation and the total variations are independent of each other, we can interchange the operators *i.e.*,

$$\delta \dot{q}_j = \delta \frac{d}{dt} q_j = \frac{d}{dt}\delta q_j$$

. Hence the above equation reduces to

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

Integrating by parts, the second term, we have

$$\int_{t_A}^{t_B} \sum_j \left(\frac{\partial T}{\partial q_j} - \frac{\partial V}{\partial q_j}\right) \delta q_j dt + \sum_j \frac{\partial T}{\partial \dot{q}_j} \delta q_j \Bigg|_{t_A}^{t_B} - \int_{t_A}^{t_B} \sum_j \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) \delta q_j dt = 0$$

Since in variational problems, the variations are considered in such a way that they vanish at the end points, *i.e.*, $\delta q_j \Big|_{t_A}^{t_B} = 0.$ Therefore, the above equation reduces to

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

The delta variations δq_j are actually independent of each other in the sense that variations in one generalised coordinate does not influence other. So we can equate the coefficients of every δq_j to zero. Equating the coefficient corresponding to the variation of the j-th generalised coordinate,

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

or

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

Now, for conservative systems, V is not a function of velocities $\dot{q}_j$, but only of the co-ordinates. Therefore,

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

or

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0, \qquad \text{where} \quad L = T - V.$$

This set of equations is the Lagrange's equation of motion corresponding to a holonomic conservative dynamic system.

5.3.2 Hamilton's varitational principle for holonomic and non-conservative systems

When a dynamical system is acted on by non-conservative forces, *i.e.*, forces that are not derivable from a scalar function, the Hamilton's principle is written in the form

$$\delta I = \delta \int_{t_A}^{t_B} (T + \dot{W}) dt = 0 \tag{5.3.5}$$

with fixed end points, where $\delta W = \delta\left(\sum_i \vec{F}_i . \vec{r}_i\right) = \sum_i \vec{F}_i . \delta \vec{r}_i$ represents the amount of work done by the force on the system for a virtual displacement from the actual path to the varied paths. The possible varied paths can be parametrised in the generalised coordinates through a quantity

α. That is, the varied paths can be represented by the generalised co-ordinates $q_j(t, \alpha)$. To proceed further, we write the transformation equations can be written as

$$\vec{r}_i = \vec{r}_i[q_j(t, \alpha), t]$$

from which we find,

$$\delta\vec{r}_i = \sum_j \frac{\partial\vec{r}_i}{\partial q_j}\delta q_j$$

The expression for the components of the generalised force is then

$$\begin{aligned}\delta W &= \sum_i \vec{F}_i.\delta\vec{r}_i \\ &= \sum_{i,j} \vec{F}_i.\frac{\partial\vec{r}_i}{\partial q_j}\delta q_j \\ &= \sum_j Q_j\delta q_j, \\ \text{where} \quad Q_j &= \sum_i \vec{F}_i.\frac{\partial\vec{r}_i}{\partial q_j}\end{aligned}$$

Substituting these values to (5.3.5), we have

$$\delta\int_{t_A}^{t_B} T dt + \int_{t_A}^{t_B} \sum_j Q_j\delta q_j dt = 0 \tag{5.3.6}$$

Further, the kinetic energy term is a function of the generalised coordinates q_j and generalised velocities $\dot{q}_j$, we can write the first term of the left hand side of equation (5.3.6) as

$$\begin{aligned}\delta\int_{t_A}^{t_B} T(q_j, \dot{q}_j)dt &= \int_{t_A}^{t_B} \sum_j \left(\frac{\partial T}{\partial q_j}\delta q_j + \frac{\partial T}{\partial \dot{q}_j}\delta\dot{q}_j\right) dt \\ &= \int_{t_A}^{t_B} \sum_j \left(\frac{\partial T}{\partial q_j}\delta q_j\right) dt + \sum_j \frac{\partial T}{\partial \dot{q}_j}\delta q_j\Bigg|_{t_A}^{t_B} - \int_{t_A}^{t_B} \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right)\delta q_j dt\end{aligned}$$

The middle term reduces to zero in view of the fact that the variations at the fixed end points are considered zero. Therefore,

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

Hence (5.3.6) can be written as

$$\int_{t_A}^{t_B} \sum_j \left[\frac{\partial T}{\partial q_j} - \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) + Q_j\right]\delta q_j dt = 0 \tag{5.3.7}$$

As we know, for a system with holonomic constraints, δq_j are independent of each other. Hence, for the above integration to vanish, if and only if the coefficients of each δq_j separately vanish, *i.e.*,

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

which are the Lagrange's equations of motion for holonomic and non-conservative system.

5.3.3 Hamilton's principle for non-holonomic systems

For certain types of non-holonomic system, it is normally possible to extend Hamilton's principle. First we take note of the fact that for systems with non-holonomic constraints, the generalised co-ordinates are not independent of each other, and it is not possible to reduce the number of independent coordinates by means of constraint equations of the type $f(q_1, q_2, ..., q_n, t)$. Another point to be kept in mind that the virtual displacements that are considered in variational problems must be consistent with the constraints that operate on the system. With non-holonomic systems, this may not be always possible to construct variational paths consistent with the constraints.

It is however possible to use variational technique when the equations of constraint can be put in the form of linear relation connecting the differentials of the q's, *i.e.*,

$$\sum_k a_{ik} dq_k + a_{it} dt = 0, \tag{5.3.8}$$

We assuming that there are m such relations so that $i = 1, 2, 3, ..., m$, and the coefficients a_{ik} , a_{it} in each such relations, may be functions of the q's and time.

The virtual dispalcemnets referred in Hamilton's variational principle are actually taken at constant times, *i.e.*, $\delta t = 0$ and hence from the constraints as laid in (5.3.8) we can write its variational form as

$$\sum_k a_{ik} \delta q_k = 0 \tag{5.3.9}$$

which is the set of m equations. These equations can now be used to reduce the number of virtual displacements to only that of the independent displacements, or the independent generalised coordinates. To implement this, we use the method of Lagrange's undetermined multipliers.

Let us consider m number of unknown constants λ_i, $i = 1, 2, 3, ..., m$, which may be, in general, functions of time. Multiplying (5.3.9) by these constants, and integrating over time t_A to t_B, we get

$$\int_{t_A}^{t_B} \sum_{k,i} \lambda_i a_{ik} \delta q_k dt = 0. \tag{5.3.10}$$

As per the prescription of the method of Lagrange's undetermined multipliers, we now combine this equation with the Hamilton's principle for holonomic system (conservative system in particular)

$$\delta \int_{t_A}^{t_B} L dt = \int_{t_A}^{t_B} \left[\sum_{k=1}^{n} \left(\frac{\partial L}{\partial q_k} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} \right) \delta q_k \right] dt = 0$$

with the (5.3.10) so as to get

$$\int_{t_A}^{t_B} \left[\sum_{k=1}^{n} \left(\frac{\partial L}{\partial q_k} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} + \sum_i \lambda_i a_{ik} \right) \delta q_k \right] dt = 0 \tag{5.3.11}$$

The fact to be noted here is that all the n number of δq_k's are not independent of each other and so we cannot equate the integrand to zero for each k from 1 to n. In fact, the m equations in (5.3.9) connect the δq_k's. Once we choose first $(n - m)$ number of the δq_k's independently, the rest m number of δq_k's will be given from (5.3.9). Now since we are free to choose m number of λ_i's, we make our choices of each of the λ_i in such a way that

$$\frac{\partial L}{\partial q_k} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} + \sum_i \lambda_i a_{ik} = 0, \qquad k = n - m + 1, n - m + 2, \ldots\ldots, n. \tag{5.3.12}$$

holds true. These are the equations of motion for the last m of the generalised coordinates q_k. Now with the choices of λ_i's through (5.3.12), we can write (5.3.11) as

$$\int_{t_A}^{t_B} \left[\sum_{k=1}^{n-m} \left(\frac{\partial L}{\partial q_k} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} + \sum_i \lambda_i a_{ik} \right) \delta q_k \right] dt = 0 \tag{5.3.13}$$

Note that the δq_k's involved in (5.3.13) are independent of each other and hence the integrand is zero separately for each k from 1 to $n - m$, *i.e.*,

$$\frac{\partial L}{\partial q_k} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} + \sum_i \lambda_i a_{ik} = 0, \qquad k = 1, 2, 3, \ldots\ldots, n - m. \tag{5.3.14}$$

Combining (5.3.12) and (5.3.14), we get the complete set of Lagrange's equations for nonholonomic systems as

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} - \frac{\partial L}{\partial q_k} = \sum_i \lambda_i a_{ik}, \qquad k = 1, 2, 3, \ldots\ldots, n. \tag{5.3.15}$$

Note !

1. Deducing the equation of motion for nonholonomic systems using the Hamilton's variational principle involves $n + m$ quantites to be determined (n generalised coordinates and m undetermined multipliers λ_i), with only n available equations. The extra m equations for complete solutions, come from the constraint equations (5.3.8) expressed in differential form as

$$\sum_k a_{ik}\dot{q}_k + a_{it} = 0, \qquad \text{with} \quad i = 1, 2, ..., m.$$

Thus the complete solution of the equations does not only provide the generalised coordinates but also the values of the undetermined multipliers.

2. The physical significance of the unknown multipliers λ_i's can be seen by considering a holonomic system, additionally being acted on by constraint forces. These constraint forces can always be thought of as some equivalent external forces Q'_k, (supposedly not conservative) that keep the motion of the system unchanged *i.e.*, the governing equations remain the same. The forces Q'_k are then given from

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} - \frac{\partial L}{\partial q_k} = Q'_k$$

and must be identical with (5.3.15). Hence we can identify $\sum \lambda_i a_{ik}$ with Q'_k, the generalised forces of constraint.

5.3.4 Summary

In this unit of study we have discussed the variational calculus which forms the essential ingradients to deduce the Lagrange's dynamical equations from an integral principle. The integral principle basically tells that the governing equations follows from the minimization of a given integral when varied between two points in configuration space with some constraints. Starting with the definition, the discussion progressed towards the concept of the δ-variation and the deduction of the Euler-Lagrange equation. A few general applications have then been discussed so that the readers can grasp with the basic concepts concerning the calculus of variation. The connection of the Euler-Lagrange equation with the Lagrange's dynamical equation was then established so that once the quantities to be minimized are found out as integral, the Euler equations can be directly written down. The minimization principle were then extended to discuss problems by altering for varieties of constraints including the non-holonomic constraints and the relevant equations of motions conforming to the constraints are deduced.

Self study questions:

1. Describe the method and philosophy of the calculus of variation and deduce the relevant equations.

2. In some unknown space, the distance between two infinitesimally separated points in a given plane is described by $ds^2 = dx^2 - dy^2$. Find the equation of the shortest distance drawn between two general points, say A and B in this space.

3. State the Hamilton's variational principle for the case of a holonomic and conservative system and apply this to deduce the relevant Hamilton's canonical equations.

4. Illustrate how one can modify the Hamilton's principle in the case of a nonholonomic system.

UNIT 6

Rigid Body Dynamics-I

Preparatory inputs to this unit

1. Newtonian dynamics.
2. Lagrange's equations.
3. Basics of vecotor Algebra.
4. Basics of ordinary differential equations.

6.1 Introduction

Most of the objects we deal in our daily life are not the point masses, the hypothetical constructs, or the particles occupying zero volume, but essentially the extended bodies. The extended bodies or masses consist of a large collection of particles occupying a definite non-zero volume in 3-dimensional space. In the previous chapters we could see the extensive use of the concepts of the point masses in deducing various results in Newtonian dynamics. The dynamics of extended objects on the other hand needs special and careful attention in the deduction of the results concerning the system. This is because of two reasons- (1) a large conglomeration of mass points which apparently requires the dynamics to be developed for each points; (2) the relative locations of any pair of such mass points may keep changing, either in magnitude or in direction, during the course of motion, as exemplified by the case of compressible materials or in rotating bodies. If the distance between any two mass points does not alter with time, the body is said to be incompressible. If this condition prevails strictly, the body is not at all compressible and such bodies may be called rigid bodies.

Euler began work on the general motion of a rigid body and found necessary and sufficient conditions for a body to possess permanent rotation, without actually looking for solutions. He also argued that a body cannot rotate freely unless the term called the *products of inertia* vanishes entirely. His experience of earlier researches in hydraulics during the 1740s, helped Euler to adopt a fundamentally different approach to mechanics in general and rigid bodies in particular. Euler's approach consists in how the Newton's equations $\vec{F} = m\vec{a}$ defined in rectangular coordinates can be used to write down the governing differential equations for the general motion of a rigid body (in particular, three-dimensional rigid bodies). He assumed internal forces within the body can be ignored for forming the correspondng torque since such forces cannot change the shape of the body. Thus, Euler eventually arrived at the Euler's equations of rigid dynamics, expressed in terms of the angular velocity vector and the intertia tensor. Euler's equations of motion for a rigid body, in fact consists of three non-linear, coupled differential equations and its complete general solution is yet to be known. Only in some special cases the solution could so far be found-the torque free motion of rigid bodies and the motion of symmetric rigid bodies.

6.1.1 Rigid Body

A rigid body is defined as a system of mass points subject to the holonomic constraints of the distance between any pair of mass points remaining a constant throughout the motion.

In fact, rigid body is an abstract idea: there is no such material body which is perfectly rigid because the constituent particles of a body are never at rest. But compared to the magnitude of displacement of the centre of mass of the body, the individual motions of the constituent particles are very small so that the distance between any two particles may be reasonably considered to be constant and the body can be assumed to be *sufficiently* a perfect body. We will also consider such a body to be a perfect body during our discussion below.

A rigid body can have two type of motion - a translational motion and rotational motion.

The motion of a rigid body can be completely described if the position and the orientation of the body is given. If the body is fixed at one point, it can rotate about any axis passing through that point. For one additional point of the body is fixed, *i.e.*, for a rigid body with two points fixed, the body can rotate about the axis passing through these two fixed points. If we fix one more point, not lying on the straight line passing through the earlier two fixed points, the now cannot execute rotational motion and the coordinate of the third point will help to locate the rigid body in space.

6.1.2 Degree of Freedom of a Rigid Body

The number of degree of freedom of a free rigid body is the minimum number of independent coordinates required to describe all possible configurations of the rigid body. In the discussion above, we see that the distance between any two constituent particles of a rigid body remain unchanged throughout the motion of the body *i.e.,* the motion of a rigid body is restricted by the requirement that the distance between *any two* of its particles remain same for all time. Mathematically this is expressed as equations of the form

$$r_{ij} = c_{ij} \tag{6.1.1}$$

where r_{ij} is the distance between the i-th and j-th particles of the rigid body and the c's are constants.

Now consider a rigid body consisting of N particles. Ideally there should be $3N$ degrees of freedom, had all the N particles been free to move in 3-dimensional space. But as the body is rigid, the restriction (6.1.1) is in operation and hence the number of degrees of freedom will be greatly reduce. To find the degrees of freedom for a rigid body we proceed as follows:

The idea is to place a given rigid body into a 3-dimensional coordinate system. To do so, we first choose any three particles of the rigid body which are non-collinearly located. Now we place these three paricles in the coordinate system one by one.

Choice of location of the first particle: We can move the first particle in three independent ways to fix in a given location in a coordinate system. As for example, in the Cartesian coordiantes, we can cause it to move along the x-axis, y-axis or z-axis by any amount at our choice. So the degrees of freedom in placing the first particle is 3.

Choice of location for the second particle: Once the first particle is fixed, the second particle has to be fixed so as to maintain a given fixed distance by virtue of (6.1.1), *i.e.*, the second particle is placed such that the constant distance between the first and the second particle r_{12} equals the constant c_{12}. This means the second particle can move to take a place anywhere on the surface of a sphere of radius c_{12} with the first particle at the centre. The degree of freedom of the second particle having freedom to move on the surface is therefore; 2.

Choice of location for the third particle: The third particle which needs to maintain a constant distance between the first as well as the second particle, *i.e.* $r_{13} = c_{13}$ and $r_{23} = c_{23}$ can have the freedom to move anywhere around a circular path the axis joining the first and the second particle and thus its degree of freedom, moving around a circle is 1.

Now considering the location of the fourth particle maintaining the conditions $r_{14} = c_{14}$, $r_{24} = c_{24}$, $r_{34} = c_{34}$, we find that there is no freedom left for the fourth particle to move, *i.e.*, the location of the particle is fixed by the locations of the first three particles and thus its degree of freedom will be zero. Similar is the case for other particles *viz.*, fifth, sixth, seventh and so on, building up the rigid body conforming to the constraints (6.1.1) and in all these cases the degree of freedom will be zero.

Thus the total degree of freedom for a complete specification of the configureation of a rigid body is the sum of the degrees of freedom above, which turns out to be *six*. This means, a total of six independent generalised coordinates is sufficient to describe the configureation of a rigid body.

That a rigid body has six degrees of freedom, can also be understood from the following:
As the degree of freedom decides the number of independent generalised coordinates required for specifying the configuration, we can start counting this requirement by constructing some orthogonal axes.

First, we will require three independent coordinates to specify a point in the rigid body which allows the latter to undergo a translational motion relative to some fixed or inertial frame of reference. This means three degrees of freedom will be used up to describe the translational motion in the rigid body in a 3-dimensional orhtogonal coordinate system, which is an inertial frame of reference in the external space. This frame of reference is known as the *space frame of reference* in relation to the rigid body in question.

Next, the rigid body may also execute rotational motion about any axis passing through the origin of the body system of coordinates, which can be shown to be a vector sum of three independent rotations about the three mutually perpendicular axes which form the *body frame of reference*, a non-inertial reference frame for specifying rotation components. Correspondingly three more independent coordinates, and hence three degrees of freedom are required to describe the rotation in the rigid body.

So, we see that in a rigid body, the available six degrees of freedom are shared for the description of the translational and rotational motion in the rigid body taking up three degrees of freedom each.

Thought Capsule 6.1 *Does it require that the origin of the body system of coordinates to describe the rotational motion of the rigid body need to be within the physical boundary of rigid body? Explain.*

6.2 Euler's theorem

Euler's theorem is one of the basic theorems used for describing the motion in a rigid body. The theorem states that

Theorem 6.2.1 *Any general displacement of a rigid body, one point of which is fixed is a rotation about some axis passing through the fixed point.*

As per the statement of the theorem, one point of the rigid body has been fixed, so the body cannot execute translational motion. Hence according to the theorem it is aloways possible to find out a single rotation about some axis when the body rotates from its original to final orientation.

To elaborate, we take the body-set of coordinates fixed within the rigid body so that the origin of the system coincides with the fixed point. During the rotational motion in a body we see that the position vector of any particle does not change in its magnitude. Now if we can find a straight line such that every particle of the body maintains a constant distance from the straight line during the course of its rotation, the Euler's theorem is established and the straight line will be the axis of rotation.

Let us consider the positions A and B to be occupied by two particles in the rigid body, which occupy positions A' and B' respectively after an arbitrary rotational displacement. Let O be the fixed point (Figure 6.1). The rigid body with the initial configuration OAB has now been changed to $OA'B'$. We draw two perpendicular planes $P1$ and $P2$ as shown in the figure 6.1. Let the planes divide the angles of the triangles OAA' and OBB' at O and intersect each other long the straight line OC. Now we can see that every point on the plane $P1$ is equidistant from points A and A' and B and B' are equidistant from the plane $P2$. The straight line OC formed by the intersection of the

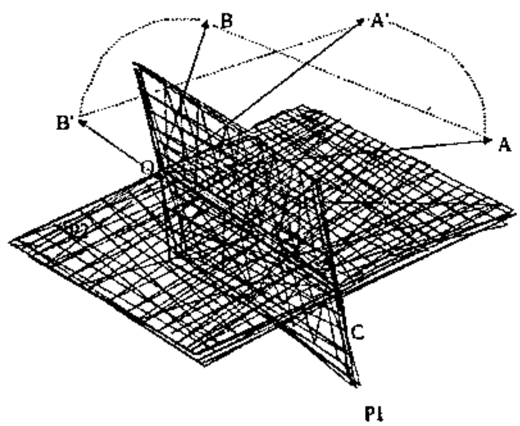


Figure 6.1: Euler Theorem

planes $P1$ and $P2$ is the line which maintains equal distance from the two positions before and after the rotation. The straight line OC is therefore the axis of the rotation, both of A to A' and B to B'. As the particles at A and B are the part of the rigid body governed by the distance constraint (6.1.1), we can say that the line OC remain unchanged and the displacement is equivalent to a rotation about OC. This proves the theorem.

Note!

There also exists a general version of the Eulĕr theorem and is known as the *Chasles' Theorem.* The Chasles' theorem states that *the most general dispalcement of the rigid body is a translation plus a rotation about some axis.* The essence of the Chasles' theorem is that it is possible to separate the discussion of the rigid body motions into two parts - the translation and the rotation, with the proper sharing of the available six degrees of freedom *viz.*, three describing the translational motion and the rest three the rotational motion. Further, it is convenient to choose the centre of mass of the rigid body as the said fixed point. It then turns out that the total angular momentum or the total kinetic energy of the body equals the sum of the angular momentum/kinetic energy of the centre of mass and the angular momenta/kinetic energy of the constituent particles about the centre of mass.

6.3 Rate of change of a vector: Rotating co-ordinate system

Though it appears slightly digressing from the main course of the rigid body dynamics, it still forms a preparatory part in the analysis of rigid body dynamics, particularly when the rigid body executes rotatory motion. Let us consider an orthogonal cartesian coordinate system $O'(x', y', z')$ with its origin O' fixed in space and another similar coordinate system $O(x, y, z)$ whose origin O is coincident with O' and rotating with an angular velocity ω about some instantaneous axis passing through the common origin (figure 6.2).

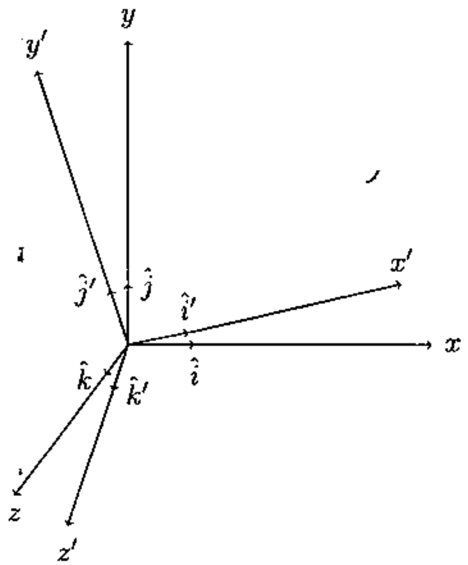


Figure 6.2: Rotating coordinate system

It is obvious that the unit vectors $(\hat{i}', \hat{j}', \hat{k}')$ corresponding to the fixed primed coordinate system are accordingly fixed in space while those in the unprimed coordinate system, $(\hat{i}, \hat{j}, \hat{k})$ are constantly changing their directions because of the rotation of the system itself.

The position vector of a particle at P, say, can be written as

$$\left.\begin{aligned} \vec{r} &= \hat{\imath}'x' + \hat{\jmath}'y' + \hat{k}'z' \\ \text{Further,} \qquad \vec{r} &= \hat{\imath}x + \hat{\jmath}y + \hat{k}z \end{aligned}\right\} \tag{6.3.1}$$

To transform the equations from unprimed system to primed system, first we need to take the dot product of $\vec{r}$ with the unit vectors $\hat{\imath}', \hat{\jmath}'$ and $\hat{k}'$. The result is

$$\left.\begin{aligned} x' &= (\vec{r}.\hat{\imath}') = \hat{\imath}.\hat{\imath}'x + \hat{\jmath}.\hat{\imath}'y + \hat{k}.\hat{\imath}'z \\ y' &= (\vec{r}.\hat{\jmath}') = \hat{\imath}.\hat{\jmath}'x + \hat{\jmath}.\hat{\jmath}'y + \hat{k}.\hat{\jmath}'z \\ z' &= (\vec{r}.\hat{k}') = \hat{\imath}.\hat{k}'x + \hat{\jmath}.\hat{k}'y + \hat{k}.\hat{k}'z \end{aligned}\right\} \tag{6.3.2}$$

The dot products of the right hand side of (6.3.2) are the direction cosines of the angles between the corresponding axes. In the similar fashion we can obtain the components in the unprimed coordinates by an inverse transformation, *i.e.*, by taking dot products of $\vec{r}$ with the unit vectors $\hat{\imath}$, $\hat{\jmath}$ and $\hat{k}$:

$$\left.\begin{aligned} x &= (\vec{r}.\,\hat{\imath}) = \hat{\imath}'.\,\hat{\imath}\,x' + \hat{\jmath}'.\,\hat{\imath}\,y' + \hat{k}'.\,\hat{\imath}\,z' \\ y &= (\vec{r}.\,\hat{\jmath}) = \hat{\imath}'.\,\hat{\jmath}\,x' + \hat{\jmath}'.\,\hat{\jmath}\,y' + \hat{k}'.\,\hat{\jmath}\,z' \\ z &= (\vec{r}.\,\hat{k}) = \hat{\imath}'.\,\hat{k}\,x' + \hat{\jmath}'.\,\hat{k}\,y' + \hat{k}'.\,\hat{k}\,z' \end{aligned}\right\} \tag{6.3.3}$$

The above transformations (6.3.2), (6.3.3) are not only true for the position vectors, but can also be extended for any vector function, and also not necessary that the vector function has to pass through the origin. Thus for a vector function $\vec{V}(t)$ we can write

$$\vec{V} = \hat{\imath}V_x + \hat{\jmath}V_y + \hat{k}V_z = \hat{\imath}'V_x' + \hat{\jmath}'V_y' + \hat{k}'V_z'$$

The time derivatives of the vector function does not behave similarly, as we can see from the following.

In the primed or the fixed system, the time rate of change of $\vec{V}$ can be expressed as

$$\left(\frac{d\vec{V}}{dt}\right)_{\text{fix}} = \dot{\vec{V}}_{\text{fix}} = \hat{\imath}'\dot{V}_x' + \hat{\jmath}'\dot{V}_y' + \hat{k}'\dot{V}_z'$$

Here as the unit vectors as seen by an observer in the primed system are constant vectors, their time derivatives vanish. However, as the unprimed frame of reference is rotating, the unit vectors will also rotate with the frame and therefore their time derivatives will also contribute to $\dfrac{d\vec{V}}{dt}$. Therefore, as seen from the fixed prime frame, the expression of $\dfrac{d\vec{V}}{dt}$ can be expressed in terms of the quantities of the rotating frame as

$$\left(\frac{d\vec{V}}{dt}\right)_{\text{fix}} = \hat{\imath}\,\dot{V}_x + \hat{\jmath}\,\dot{V}_y + \hat{k}\,\dot{V}_z + \frac{d\hat{\imath}}{dt}\,V_x + \frac{d\hat{\jmath}}{dt}\,V_y + \frac{d\hat{k}}{dt}\,V_z \tag{6.3.4}$$

Here the last three terms appear because of the rotating nature of the frame and does not constitute the actual or the inherent rate of change of $\vec{V}$. So subtracting them, from the expression will actually represent the actual rate of change of the vector as seen from the rotating frame, *i.e.*,

$$\left(\frac{d\vec{V}}{dt}\right)_{\text{rot}} = \hat{\imath}\,\dot{V}_x + \hat{\jmath}\,\dot{V}_y + \hat{k}\,\dot{V}_z \tag{6.3.5}$$

Now, the linear velocity of a particle having a poistion vector $\vec{r}$ and rotating with angular velocity $\vec{\omega}$ about the axis passing through the same origin is given by

$$\vec{v} = \frac{d\vec{r}}{dt} = \vec{\omega} \times \vec{r}$$

Here $\vec{r}$ represents any position vector rotating in the body and the left-hand side is the time derivative of it. Hence this formula can be applied to rotating unit vectors too.

The rotating unit vectors $(\hat{\imath}, \hat{\jmath}, \hat{k})$ in our case, which are rotating with the same angular velocity $\vec{\omega}$, we have

$$\frac{d\hat{\imath}}{dt} = \vec{\omega} \times \hat{\imath}, \qquad \frac{d\hat{\jmath}}{dt} = \vec{\omega} \times \hat{\jmath}, \qquad \text{and} \qquad \frac{d\hat{k}}{dt} = \vec{\omega} \times \hat{k} \tag{6.3.6}$$

Substituting (6.3.6) to (6.3.4) along with the consideration of (6.3.5) we get

$$\left(\frac{d\vec{V}}{dt}\right)_{\text{fix}} = \left(\frac{d\vec{V}}{dt}\right)_{\text{rot}} + \vec{\omega} \times \vec{V} \tag{6.3.7}$$

(6.3.7) can be considered as an operator equation giving the relations between the time derivatives in the fixed and the rotating frames of references, *i.e.*,

$$\boxed{\left(\frac{d}{dt}\right)_{\text{fix}} = \left(\frac{d}{dt}\right)_{\text{rot}} + \vec{\omega} \times} \tag{6.3.8}$$

which can be operated on any vector. In particular, if we operate on the angular momentum vector $\vec{\omega}$, then we get

$$\begin{aligned}\left(\frac{d\vec{\omega}}{dt}\right)_{\text{fix}} &= \left(\frac{d\vec{\omega}}{dt}\right)_{\text{rot}} + \vec{\omega} \times \vec{\omega} \\ &= \left(\frac{d\vec{\omega}}{dt}\right)_{\text{rot}} \qquad [\because \quad \vec{\omega} \times \vec{\omega} = 0\,] \\ &= \dot{\vec{\omega}}\end{aligned} \tag{6.3.9}$$

(6.3.9) shows that the angular acceleration $\dot{\vec{\omega}}$ is the same in the fixed and the rotating frames.

The second order derivative of $\vec{V}$ can also be found out similarly. But before it we would simplify the notations as

$$\left(\frac{d}{dt}\right)_{\text{fix}} = \frac{d'}{dt}, \qquad \text{and} \qquad \left(\frac{d}{dt}\right)_{\text{rot}} = \frac{d}{dt} \tag{6.3.10}$$

With these, we then write

$$\frac{d'^2\vec{V}}{dt^2} = \frac{d'}{dt}\left(\frac{d'V}{dt}\right)$$
$$= \frac{d'}{dt}\left[\frac{d\vec{V}}{dt} + \vec{\omega}\times\vec{V}\right]$$
$$= \left[\frac{d}{dt} + \vec{\omega}\times\right]\left[\frac{d\vec{V}}{dt} + \vec{\omega}\times\vec{V}\right]$$
$$= \frac{d^2\vec{V}}{dt^2} + \frac{d\vec{\omega}}{dt}\times\vec{V} + \vec{\omega}\times\frac{d\vec{V}}{dt} + \vec{\omega}\times\left[\frac{d\vec{V}}{dt} + \vec{\omega}\times\vec{V}\right]$$

or,
$$\frac{d'^2\vec{V}}{dt^2} = \frac{d^2\vec{V}}{dt^2} + 2\vec{\omega}\times\frac{d\vec{V}}{dt} + \vec{\omega}\times\vec{\omega}\times\vec{V} + \frac{d\vec{\omega}}{dt}\times\vec{V} \tag{6.3.11}$$

The relations can be used to obtain expressions for velocity and acceleration of the particle situated at the point P.

We would now generalise the case to include the translational motion of the oirgin O of the rotating coordinate system with respect to that of the fixed frame, O'. We recall here that if $\vec{R}$ is position vector at any instant of time t of the origin O of the rotating coordinate with respect to the point O', the origin of the fixed frame, then the position vector $\vec{r}$ of the point P with respect to O' and the position vector $\vec{r}'$ with respect to O are related by

$$\vec{r}\,' = \vec{R} + \vec{r}$$

and its differentiation with respect to time, as applied to the situation above is

$$\left(\frac{d\vec{r}\,'}{dt}\right)_{\text{fix}} = \left(\frac{d\vec{R}}{dt}\right)_{\text{fix}} + \left(\frac{d\vec{r}}{dt}\right)_{\text{fix}}$$

Consideration of the results of (6.3.8), (6.3.9) and (6.3.11) enables us to write

$$\left(\frac{d\vec{r}\,'}{dt}\right)_{\text{fix}} = \left(\frac{d\vec{R}}{dt}\right)_{\text{fix}} + \left(\frac{d\vec{r}}{dt}\right)_{\text{rot}} + \vec{\omega}\times\vec{r} \tag{6.3.12}$$

and

$$\left(\frac{d^2\vec{r}\,'}{dt^2}\right)_{\text{fix}} = \left(\frac{d^2\vec{R}}{dt^2}\right)_{\text{fix}} + \left(\frac{d^2\vec{r}}{dt^2}\right)_{\text{rot}} + 2\,\vec{\omega}\times\left(\frac{d\vec{r}}{dt}\right)_{\text{rot}} + \vec{\omega}\times(\vec{\omega}\times\vec{r}) + \frac{d\vec{\omega}}{dt}\times\vec{r} \tag{6.3.13}$$

Let us put these formulae in a compact form by using the suffix f for *fix* and r for *rot* so that the equations (6.3.12) and (6.3.13) appear as

$$\dot{\vec{r}}\,'_f = \dot{\vec{R}}_f + \dot{\vec{r}}_r + \vec{\omega}\times\vec{r} \tag{6.3.14}$$

and
$$\ddot{\vec{r}}\,'_f = \ddot{\vec{R}}_f + \ddot{\vec{r}}_r + 2\vec{\omega}\times\dot{\vec{r}}_r + \vec{\omega}\times(\vec{\omega}\times\vec{r}) + \dot{\vec{\omega}}\times\vec{r} \tag{6.3.15}$$

where	$\dot{\vec{r}}'_f$	$\equiv$	velocity of the moving particle relative to fixed axes,
	$\ddot{\vec{r}}'_f$	$\equiv$	acceleration in the particle relative to fixed axes,
	$\dot{\vec{R}}'_f$	$\equiv$	linear velocity of the origin of the rotating axes,
	$\ddot{\vec{R}}'_f$	$\equiv$	linear acceleration of the origin of the rotting axes,
	$\dot{\vec{r}}_r$	$\equiv$	velocity of the particle relative to the rotating frame of reference,
	$\ddot{\vec{r}}_r$	$\equiv$	accelertion in the particle as observed from the rotating frame of reference
	ω	$\equiv$	angular velocity of the rotating axes,
	$\vec{\omega} \times \vec{r}$	$\equiv$	velocity due to rotation of the axes,
	$2\,\vec{\omega} \times \dot{\vec{r}}$	$\equiv$	the *coriolis acceleration*
	$\vec{\omega} \times \vec{\omega}$	$\equiv$	angular acceleration of the particle due to the acceleration of the rotating axes,
and	$\dot{\vec{\omega}} \times \vec{r}$	$\equiv$	angular acceleration of the article because of the acceleration of the rotating axes.

The Non-Intertial force

To discuss the coriolis force, we first recall that the Newton's laws of motion, in particular the second law, is valid only in the inertial frame of reference, *i.e.*, if a particle of mass m is acted on by an external force $\vec{F}$ resulting in an acceleration $\vec{a} = \frac{d^2\vec{r}}{dt^2}$, all the quantities are measured by an observer in a given inertial frame of reference, designated as *fix*, then the corresponding Newton's second law of motion can be written as

$$\vec{F} = m\vec{a} = \left(\frac{d^2\vec{r}}{dt^2}\right)_{\text{fix}} \tag{6.3.16}$$

The suffix *fix* here reminds us that the associated differentiations must be carried out with respect to the fixed reference frame. Now suppose we want to rewrite the equation in a rotating frame of reference rotating say with a constant angular velocity, such that it preserves the form of the equation. Under this circumstance, there is no angular acceleration of the rotating from such that $\dot{\vec{\omega}} = 0$. If further the origins of the fixed and the rotating frames coincide, then $\vec{R} = 0$ so that $\vec{r}' = \vec{r}$. Therefore equation (6.3.15) reduces to

$$\begin{aligned} m\left(\frac{d^2\vec{r}}{dt^2}\right)_{\text{rot}} &= m\left(\frac{d^2\vec{r}}{dt^2}\right)_{\text{fix}} - 2m\vec{\omega} \times \left(\frac{d\vec{r}}{dt}\right)_{\text{rot}} - m\vec{\omega} \times (\vec{\omega} \times \vec{r}) \\ &= \vec{F}_{\text{eff}} \end{aligned} \tag{6.3.17}$$

Thus, we can consider the equation (6.3.17) to represent a situation where a particle is acted on by a couple of forces in the rotating frame, *viz.*,

1. the real force $\vec{F} = m\left(\frac{d^2\vec{r}}{dt^2}\right)_{\text{fix}}$,
2. the centrifugal force $-m\vec{\omega} \times (\vec{\omega} \times \vec{r})$ arising because of the rotation of the coordinate axes, and
3. the coriolis force $-2m\vec{\omega} \times \left(\frac{d\vec{r}}{dt}\right)_{\text{rot}}$ arising as a result of the motion of the particle in the rotating frame.

Whereas the same was seen from the fixed reference frame as the particle being acted only by the real force. The last two forces are seen to arise out of the selection of the reference frame, *e.g.*, the rotating frame here which is actually a *non-intertial frame*. The coriolis force and the centrifugal forces are thus the non-inertial forces and are to be added to the force term $F = m\left(\frac{d^2\vec{r}}{dt^2}\right)_{\text{rot}}$ so as to give $\vec{F}_{\text{eff}}$ and hence to resemble the equation of motion to that in the fixed frame, *i.e.*, the equation (6.3.16). Hence the thumb rule to relate the forces in the inertial and non-inertial frames is

$$\boxed{\vec{F} + \text{non-inertial forces} = \vec{F}_{\text{eff}}}$$

6.4 Angular momentum and kinetic energy of a rigid body

Consider a rigid body composed of n particles having masses $m_i \quad (i = 1, 2, \ldots, n)$ and rotating with instantaneous angular velocity $\vec{\omega}$. Let one of the points in the body be fixed. Hence, the body cannot execute translational motion. The only possible motion of such a body is the rotational motion. We shall now find out the expressions for the angular momentum and the kinetic energy associated with the rotation of the body.

The linear velocity $\vec{v}_i$ of the i-th particle of mass m_i and position vector $\vec{r}_i$ with respect to the fixed point is given by

$$\vec{v}_i = \dot{\vec{r}}_i = \vec{\omega} \times \vec{r}_i.$$

The linear momentum of the i-th particle is therefore

$$\vec{p}_i = m_i \vec{v}_i$$

and the corresponding angular momentum about an axis passing through the origin,

$$\vec{l}_i = \vec{r}_i \times m_i \vec{v}_i$$

The total angular momentum $\vec{L}$ of the rigid body is the sum of angular momenta $\vec{l}_i$ of the individual particle and is given by

$$\begin{aligned}
\vec{L} &= \sum_{i=1}^{n} \vec{l}_i = \sum_i \vec{r}_i \times m_i \vec{v}_i \\
&= \sum_i m_i \vec{r}_i \times (\vec{\omega} \times \vec{r}_i) = \sum_i m_i r_i^2 \vec{\omega} - \sum_i m_i (\vec{r}_i . \vec{\omega}) \vec{r}_i, \\
\text{where,} \quad \vec{r}_i . \vec{r}_i &= |\vec{r}_i|^2 = r_i^2
\end{aligned}$$

The summation is carried out over all the particles of rigid body. If we place a cartesian coordinate system in the rigid body with its origin at the fixed point, so that

$$\vec{r}_i \equiv (x_i, y_i, z_i); \qquad \vec{v}_i = \dot{\vec{r}}_i \equiv (\dot{x}_i, \dot{y}_i, \dot{z}_i) \qquad \vec{\omega} \equiv (\omega_x, \omega_y, \omega_z)$$

then expression of the x-component of the angular momentum of the rigid body can be written as

$$L_x = \sum_i m_i[r_i^2 - (x_i^2\omega_x + x_i y_i \omega_y + x_i z_i \omega_z)] \tag{6.4.1}$$

$$= \sum_i m_i(r_i^2 - x_i^2)\omega_x - \sum_i m_i x_i y_i \omega_y - \sum_i m_i x_i z_i \omega_z. \tag{6.4.2}$$

Similarly we can find the expressions for L_y and L_z as

$$L_y = \sum_i m_i x_i y_i \omega_x + \sum_i m_i(r_i^2 - y_i^2)\omega_y - \sum_i m_i y_i z_i \omega_z. \tag{6.4.3}$$

and

$$L_z = -\sum_i m_i x_i z_i \omega_x - \sum_i m_i y_i z_i \omega_y + \sum_i m_i(r_i^2 - z_i^2)\omega_z. \tag{6.4.4}$$

Let us now introduce some subscripted quantities I_{ab} for the coefficients of ω_x, ω_y and ω_z of the equation (6.4.2), where the suffix a and b may denote any of the three cartesian components, *viz.*, x, y or z. We write

$$I_{xx} = \sum_i m_i(r_i^2 - x_i^2) = \sum_i m_i(y_i^2 + z_i^2) \tag{6.4.5a}$$

$$I_{yy} = \sum_i m_i(x_i^2 + z_i^2) \tag{6.4.5b}$$

$$I_{zz} = \sum_i m_i(x_i^2 + y_i^2) \tag{6.4.5c}$$

$$I_{xy} = -\sum_i m_i x_i y_i = I_{yx} \tag{6.4.5d}$$

$$I_{yz} = -\sum_i m_i y_i z_i = I_{zy} \tag{6.4.5e}$$

$$I_{zx} = -\sum_i m_i x_i z_i = I_{xz} \tag{6.4.5f}$$

These quantities in equations (6.4.5a) through (6.4.5f) actually define *moments of inertia* about various axes and the *products of inertia*. The quantities I_{xx}, I_{yy} and I_{zz} (with matching subscripts) are called the moments of inertia about x, y and z axes, respectively, and the terms with the non-matching subscripts such as I_{xy}, I_{xz} and I_{yz} are the the products of inertia.

Using these quantities, we can express the components of the angular momentum vector in a tidy manner, so that

$$\left.\begin{aligned} L_x &= I_{xx}\omega_x + I_{xy}\omega_y + I_{xz}\omega_z \\ L_y &= I_{yx}\omega_x + I_{yy}\omega_y + I_{yz}\omega_z \\ L_z &= I_{zx}\omega_x + I_{zy}\omega_y + I_{zz}\omega_z \end{aligned}\right\} \tag{6.4.6}$$

Next we proceed to deduce the expression for the rotational kinetic energy of a rigid body by recalling and using the relations as above. The total energy T of a rigid body is the sum of the individual rotational kinetic energies of each particles. For the i-th particle with mass m_i moving with a velocity $\vec{v}_i$, we have,

$$
\begin{aligned}
T &= \sum_i \frac{1}{2} m_i v_i^2, \qquad \text{with} \qquad |\vec{v}_i| = v_i \\
2T &= \sum_i m_i |v_i|^2 = \sum_i m_i \dot{\vec{r}}_i . \dot{\vec{r}}_i \\
&= \sum_i m_i \dot{\vec{r}}_i . (\vec{\omega} \times \vec{r}_i) \\
&= \sum_i m_i \vec{\omega} . (\vec{r}_i \times \dot{\vec{r}}_i) \\
&= \vec{\omega} . \sum_i (\vec{r}_i \times m_i \dot{\vec{r}}_i) \\
&= \vec{\omega} . \sum_i \vec{l}_i \\
T &= \frac{1}{2} \vec{\omega} . \vec{L}
\end{aligned} \tag{6.4.7}
$$

Thus the total rotational kinetic energy of a rigid body is given by halving the result on the dot product of the angular momentum and angular velocity vectors.

We can alter some notations so that the equations (6.4.6) and (6.4.7) look more compact. Let us number the x, y and z axes with 1, 2, and 3 respectively. Then $L_x \to L_1$, $\omega_x \to \omega_1$ etc., so that the components of the angular momentum vector are

$$
\boxed{L_i = \sum_i I_{ij} \omega_j,} \qquad i = 1, 2, 3. \tag{6.4.8}
$$

and the total rotational kinetic energy is

$$
\begin{aligned}
T &= \frac{1}{2} \vec{\omega} . \vec{L} \\
&= \frac{1}{2} \sum_i \vec{\omega}_i . \vec{l}_i \\
\text{or,} \qquad T &= \frac{1}{2} \sum_{ij} I_{ij} \vec{\omega}_i . \vec{\omega}_j
\end{aligned} \tag{6.4.9}
$$

So, the total kinetic energy in terms of the moments of inertia and products of inertia is given by

$$
\boxed{T = \frac{1}{2} \sum_{ij} I_{ij} \vec{\omega}_i . \vec{\omega}_j}
$$

Thus from the boxed equations above we see tht once the mass distribution (the mass at each location (x, y, z)) are known, the quantities I_{ij} can be calculated using the relations (6.4.5a) through (6.4.5f). It then becomes easy to calculate the components of the angular velocity vector $\vec{L} \equiv (L_x,\ L_y,\ L_z)$ with the knowledge of the angular velocity components $(\omega_1,\ \omega_2,\ \omega_3)$ with respect to a given axis. The total rotational kinetic energy of the rigid body is then found out either by using the angular momentum components and angular velocity components (6.4.7), or directly using (6.4.9).

6.5 Euler Angle and Translation Matrix

We have already found that there are six degree of freedom in a rigid body. We have also learnt that the degrees of freedom in a rigid body can be interpreted in different ways. Out of these six, three degrees of freedom correspond to three independent co-ordinates which serve to locate a point on the instantaneous axis of rotation in the rigid body undergoing a translational motion in relation to some fixed inertial frame of reference. Two more degrees of freedom and hence two co-ordinates are required to locate the axis of rotation passing through some fixed point already located. Lastly, the orientation of the body can be specified in terms of an angle. Thus, in summary we find that three co-ordinates are needed to represent the position of a fixed point and the remaining three coordinates are designated by Euler angles, as described below.

The Euler angles (ϕ, θ, ψ) relate two orthogonal coordinates systems having a common origin. The transformation from one coordinate system to the another is possible through a series of successive two-dimensional rotations. More specifically, the Euler angles are the three successive angles of rotation. The sequence starts by rotating the inital system of axes, by an angle ϕ counterclockwise about the z axis, followed by a second rotation about the new x axis by an amount θ. The third rotation in the sequence is the rotation about the latest z axis of the coordinates by an angle ϕ.

Essentially, we have some initial co-ordinates (x_0, y_0, z_0) of an inertial coordinate system S_0 in three dimensions which is rotated to a final co-ordinate system S_3 with the co-ordinate (x_3, y_3, z_3), after a succession of three rotations. Usually,S_3 is identified as the body frame of reference as we said earlier and the S_0 as the fixed frame. Let the position vector $\mathbf{X}^0$ denote the set of initial coordinates (x_0, y_0, z_0) and the position vecotr $\mathbf{X}^3$ to denote the final set of coordinates (x_3, y_3, z_3). The details of each of the three rotations, designated as the *First rotation, Second rotation* and the *Third rotation* are given below.

First Rotation

The S_0 system is rotated about z_0-axis by an angle ϕ, $(0 \leq \phi \leq 2\pi)$ in anticlockwise direction so that the plane contained within x_0-y_0 axes takes new position x_1-y_1. This plane, including the same z_0 axis as the z_1 axis forms the new orthogonal coordinate system. Let this new co-ordinate system is S_1. The transformation equations are

$$\begin{aligned} x_1 &= x_0 \cos\phi + y_0 \sin\phi + 0.z_0 \\ y_1 &= -x_0 \sin\phi + y_0 \cos\phi + 0.z_0 \\ z_1 &= 0.x_0 + 0.y_0 + 1.z_0 \end{aligned}$$

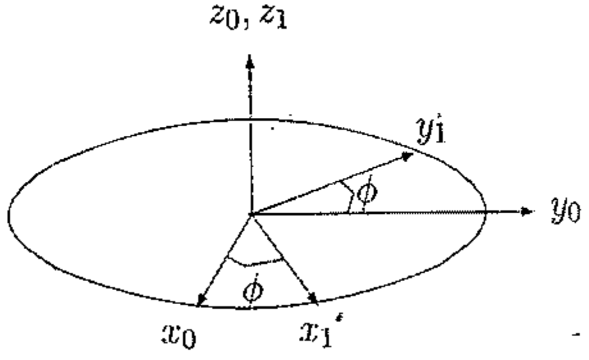


Figure 6.3: First rotation.

These relations can be written in terms of matrix, *i.e.*,

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} = R_z(\phi) \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$$

where the first rotational matrix is

$$R_z(\phi) = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

If we take the transpose of $R_z(\phi)$ *i.e.*, find $R_z^T(\phi)$ then

$$R_z^T(\phi) == \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

which, on postmultiplication with $R_z(\phi)$, results in

$$\begin{aligned} R_z^T(\phi)R_z(\phi) &= \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \cos^2\phi + \sin^2\phi & \sin\phi\cos\phi - \sin\phi\cos\phi & 0 \\ -\sin\phi\cos\phi + \sin\phi\cos\phi & \cos^2\phi + \sin^2\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_{3\times 3} \end{aligned}$$

Similarly, on premultiplication by $R_z(\phi)$ yields

$$R_z(\phi)R_z^T(\phi) = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_{3\times 3}$$

In both the cases we see that the first rotational matrix is orthogonal, *i.e.*, $R_z^T(\phi) = R_z^{-1}(\phi)$.

Second Rotation

The second rotation is rotation of the S_1 system about x_1-axis through an angle θ, $(0 \leq \theta \leq \pi)$ in anticlockwise direction to generate the new co-ordinate system S_2.

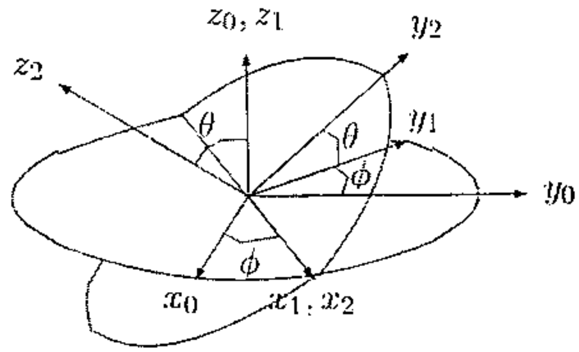


Figure 6.4: Second rotation.

Here the transformation is represented in the matrix form by

$$\begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = R_x(\theta) \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}$$

where the second rotational matrix is

$$R_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}$$

One can prove easily that the second rotational matrix is also an orthogonal matrix.

Third Rotation

The third rotation involves the Euler angle ψ, $0 \leq \psi \leq 2\pi$. The S_2 system is rotated about the z_2-axis through an angle ψ in the anticlockwise direction, generating the co-ordinate system S_3. Thus we see that the z coordinate in S_3 is identical with that of S_2,, *i.e.*, z_2-axis remain unaffected on the third rotation, which is obvious because the third rotation is about the z_2-axis.

The third transformation can be written as

$$\begin{pmatrix} x_3 \\ y_3 \\ z_3 \end{pmatrix} = R_z(\psi) \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}$$

where the third rotational matrix is

$$R_z(\psi) = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

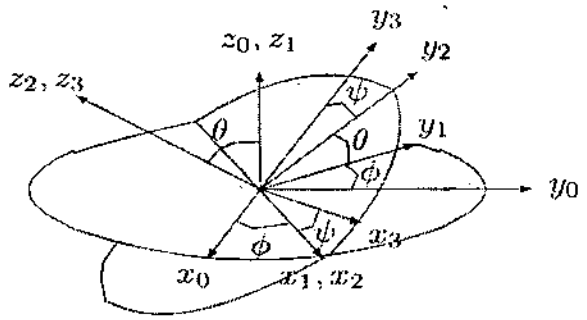


Figure 6.5: Third rotation

It is easy to see that the matrix $R_z(\psi)$ is also an orthogonal matrix.

As the third rotation completes the required transformation, we are now in a position to find the single matrix of trnansformation which transforms the reference frame S_1 to the final S_3 frame. This single matrix is known as the *Eulerian Rotational Matrix*.

Eulerian Rotational Matrix

Let us define a vector X^0 with components given by (x_0, y_0, z_0) in a fixed reference frame S_0. Let S_0 be rotated as per the prescription of the Eulerian angles to a new reference frame S_3. The components of the vector X^0 are then obviously transformed to a new set of values (x_3, y_3, z_3) and define a new vector X^3 in the frame S_3. The two vectors are related by

$$X^3 = R(\phi, \theta, \psi) X^0$$

where,

$$\begin{aligned} R(\phi,\theta,\psi) &= R_z(\psi)R_x(\theta)R_z(\phi) \\ &= \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \cos\psi\cos\phi - \cos\theta\sin\phi\sin\psi & \cos\psi\sin\phi + \cos\theta\cos\phi\sin\psi & \sin\psi\sin\theta \\ -\sin\psi\cos\phi - \cos\theta\sin\phi\cos\psi & -\sin\phi\sin\psi + \cos\theta\cos\phi\cos\psi & \cos\psi\sin\theta \\ \sin\theta\sin\phi & -\sin\theta\cos\phi & \cos\theta \end{pmatrix} \end{aligned}$$

and,

$$\mathbf{X}^3 = \begin{pmatrix} x_3 \\ y_3 \\ z_3 \end{pmatrix}, \qquad \mathbf{X}^0 = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$$

The matrix $R(\phi, \theta, \psi)$ which is obtained by the matrix multiplications of the the matrices corresponding to each rotation, ϕ, θ and ψ, taken in reverse order, is known as the *Eulerian Rotation*

Matrix.

We have already seen that the matrices R_z, R_x etc. are each orthogonal. Their matrix product given by $R = R(\phi, \theta, \psi)$ will also be orthogonal. This fact follows from the transpose and inverse properties of orthogonal matrics; this can be easily shown that the product of two orthogonal matrices results in another orthogonal matrix. From the discussion, we can easily find that the inverse of the Eulerian rotation matrix $R(\phi, \theta, \psi)$ equals its transpose, *i.e.*, $\mathbf{X}^0 = R^{-1}\mathbf{X}^3 = R^T\mathbf{X}^3$.

6.6 Summary

In this unit we have discussed at length on the basic definition of a rigid body. Starting with an introduction we have dealt upon the degrees of freedom a rigid body possesses, along with a the Euler Theorem will tells us about the kind of motion a rigid body may execute. We have seen that a rigid body may have translational motion as well as a rotation, particularly about an axis passing through a fixed axis. Further, we have seen that out of the six degrees of freedom available, three degrees of freedom describe the translational motion and the rest three are used for describing the rotational motion. An elaboration of the rate of change of a vector in a rotating coordinate system has been made which demonstrates that such a non-inertial frame are beset with pseudo- forces like the centrifugal and coriolis force which make their appearances while transforming the results of the inertial frame to a non-inertial frame. The angular momentum and the kinetic energy possessed by a rotating rigid body can be well described with the introduction of the inertia tensor. Finally, a new set of angles as generalised coordinates -the·Eulerian angles have been introduced to describe the rotatory motion of a rigid body.

Self study questions:

1. Justify that a rigid body has six degrees of freedom.
2. State and explain the theorems associated with the motion of a rigid body. What kind of motions are expected in a rigid body system?
3. Define non-inertial forces and explain their role in the dynamics of a rigid body.
4. Explore the properties of the moment of inertia tensor.
5. Show that the moment of inertia of a system consisting of N particles can be expressed as

$$I = \sum_{i=1}^{N} m_i(\vec{r}_i \times \hat{n}) \cdot (\vec{r}_i \times \hat{n}).$$

   where m_i is the mass of the i-th particle and the $\hat{n}$ is the unit vector along the direction of the axis about which the moment of inertia is sought.

6. What are Eulerian angles. Illustrate the rotations in a rigid body in terms of the Eulerian angle. Are the rotations of the rigid body by the sequence of the Eulerian angle orthogonal? Explain.

UNIT 7

Rigid Body Dynamics-II

Preparatory inputs to this unit

1. Newtonian dynamics
2. Lagrange's equations:.
3. Basics of vecotor Algebra
4. Basics of Ordinary differential equations.

7.1 The Inertia Tensor

We have already seen in the previous unit of study that in order to streamline the understanding of the angular momentum vector, the comcept of the moment of inertia was introduced. It was also seen that the moment of inertia has 9 components written in matrix notation and the matrix is symmetric. The condition of symmetricity renders the matrix to posses six independent componets. This moment of inertia matrix, often called a moment of inertia tensor or simply the inertia tensor is actually a mathematical quantity which bears the signature of the nature of the mass distribution in the rigid body. The moment of inertia tensor is denoted by $\overrightarrow{I}$, which in the matrix notation takes the following form, *i.e.*,

$$\overrightarrow{I} = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{xz} & I_{xy} & I_{zx} \end{pmatrix}$$

where,

$$I_{xx} = \sum_i m_i(y_i^2 + z_i^2)$$

$$I_{yy} = \sum_i m_i(z_i^2 + x_i^2)$$

$$I_{zz} = \sum_i m_i(x_i^2 + y_i^2)$$

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

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

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

As the moment of inertia tensor is obviously symmetric, we have

$$I_{ij} = I_{ji}$$

with six independent components. In a rigid body, the matter is continuously distributed and hence the density of matter can be expressed as a continuous function $\rho = \rho(r)$, with functional dependence on the distance from a chosen point, the fixed point. This fact make it easy to rewrite the moments and the products of inertia in a generalized form, as

$$I_{xx} = \int \rho(r)(r^2 - x^2)d\tau$$

and

$$I_{xy} = -\int \rho(r)xyd\tau$$

and so on, so that the most general form of the angular momentum vector can be written for continuous distribution of matter as,

$$\vec{L} = \int \rho(\vec{r}) \left(r^2\vec{\omega} - (\vec{r}.\vec{\omega})\vec{r}\right) d\tau.$$

Here τ denotes the volume the of the rigid body over which the integration is done.

If the three components of the moment of inertia about the principal axes are equal *i.e.*, if $I_1 = I_2 = I_3$, the body is called a *spherical top.* On the other hand, only two of the components of the moment of inertia about the principal axes being equal *i.e.*, $I_1 = I_2 \neq I_3$, is the symmetric top. Finally, if all the components of the moment of inertia about the principal axes are distinct, the body is known to be an *asymmetric top.* A body for which $I_1 = I_2$ and $I_3 = 0$ is called a *rotor.*

Numerical Examples

Example 7.1.1 *Find the Moments and Products of Inertia of a uniform rectangular parallelopiped with respect to its edges.*

Solution: Let us consider a rectangular parallelopiped with edges of length a, b and c. We set up a rectangular cartesian coordinate system at that edge so that the side a lies along x-axis, b along y-axis and c along the z-axis. Further let ρ be the mass density of the parallelopiped which is considered uniform, *i.e.*, the density does not depend on any coordinates. If M be the total mass of the parallelopiped, then the density ρ is related to it by

$$M = \rho abc.$$

Under such circumstances, the moment of intertia about z-axis is calculated as the following :

$$\begin{aligned}
I_{zz} &= \int_0^c \int_0^b \int_0^a \rho(x^2 + y^2) dx dy dz \\
&= \rho \int_0^c dz \int_0^b \int_0^a (x^2 + y^2) dx dy \\
&= \rho c \int_0^b \left(\frac{a^3}{3} + y^2 a \right) dy \\
&= \rho c \left(\frac{a^3 b}{3} + \frac{b^3 a}{3} \right) \\
&= \frac{1}{3} \rho abc(a^2 + b^2) = \frac{1}{3} M(a^2 + b^2) \qquad \text{as} \qquad M = \rho abc.
\end{aligned}$$

Similarly we can calculate out the other moments of inertia, *viz.*,

$$I_{xx} = \frac{1}{3} M(b^2 + c^2), \qquad \text{and} \qquad I_{yy} = \frac{1}{3} M(a^2 + c^2)$$

...roceed to calculate the products of inertia, which are given by

$$I_{xy} = I_{yx} = -\int_0^c \int_0^b \int_0^a \rho xy dx dy dz = -\rho c \frac{a^2}{2}\frac{b^2}{2}$$
$$= -\frac{1}{4}Mab$$

$$I_{xz} = I_{zx} = -\int_0^c \int_0^b \int_0^a \rho xz dx dy dz = -\rho b \frac{a^2}{2}\frac{c^2}{2}$$
$$= -\frac{1}{4}Mac$$

$$I_{yz} = I_{zy} = -\int_0^c \int_0^b \int_0^a \rho yz dx dy dz = -\rho a \frac{b^2}{2}\frac{c^2}{2}$$
$$= -\frac{1}{4}Mbc$$

Example 7.1.2 *Find the Moment of inertia of uniform hemisphere about*

(i) the axis of symmetry, and

(ii) a given axis lying in the base plane and perpendicular to the symmetry axis.

Solution:

(i) As the hemisphere is the half portion of a full sphere, the soltion to the problem will be easier to work out in spherical polar coordinates, $(r,\ \theta,\ \phi)$. The transformation of coordinates are related by

$$x = r\sin\theta\cos\phi$$
$$y = r\sin\theta\sin\phi$$
$$z = r\cos\theta$$

so that the elemental volume $dV = dx.dy.dz$ transform in spherical polar coordinates to $dV = r^2 \sin\theta\, d\theta\, d\phi\, dr$. We calculate the moment of inertia about the z-axis, the axis of symmetry as

$$I_{zz} = \int \rho(x^2+y^2)dV$$
$$= \int_0^{2\pi}\int_0^{\pi/2}\int_0^R \rho r^2 \sin^2\theta . r^2 \sin^2\theta d\theta d\phi dr$$
$$= \int_0^{2\pi}\int_0^{\pi/2}\int_0^R \rho r^4 \sin^3\theta d\theta d\phi dr$$
$$= \rho\frac{R^5}{5}.\frac{4\pi}{3} = \frac{2}{5}MR^2, \qquad \text{where } M = \left(\frac{2\pi R^3}{3}\right) = \text{mass of the hemisphere.}$$

(*ii*) In the similar fashion as laid in *i* above, the moment of inertia about a base axis, say the x-axis as

$$
\begin{aligned}
I_{xx} &= \int \rho(y^2 + z^2)dV \\
&= \int_0^{2\pi} \int_0^{\pi/2} \int_0^R \rho \left(r^2 \sin^2\theta \sin^2\phi + r^2 \cos^2\theta\right) r^2 \sin\theta d\theta d\phi dr \\
&= \int_0^{2\pi} \int_0^{\pi/2} \int_0^R \rho r^4 \left(\sin^3\theta \sin^2\phi + \cos^2\theta \sin\theta\right) d\theta d\phi dr \\
&= \rho \frac{R^2}{5} \int_0^{2\pi} \int_0^{\pi/2} \left[\sin\theta - (1 - \sin^2\phi)\sin^3\theta\right] d\theta d\phi \\
&= \rho \frac{R^5}{5} \left[- \int_0^{2\pi} \cos\theta|_0^{2\pi} d\phi - \int_0^{2\pi} \int_0^{\pi/2} \cos^2\phi \sin^3\theta d\theta d\phi \right] \\
&= \rho \frac{R^5}{5} \left[2\pi - \frac{2}{3} \int_0^{2\pi} \cos 6\phi d\phi \right] \\
&= \rho \frac{R^5}{5} \left[2\pi - \frac{2\pi}{3} \right] \\
&= \frac{2}{5} M R^2, \qquad \because M = \frac{2\pi}{3}\rho R^3 = \text{M is the total mass of the hemisphere.}
\end{aligned}
$$

We can see from symmetry that rotation about both the x-axis or the y-axis are equivalent so that

$$I_{yy} = I_{xx} = \frac{2}{5} M R^2$$

7.2 Euler's equations of motion for a rigid body

7.2.1 Introduction

While discussing the Kinematics of the rigid body in the earlier unit, we observed that one requires a particular reference point in the rigid body so that the motion in the rigid body can be split into parts- one purely translational motion and the other purely rotational motion about the reference point. For a reference point which is made fixed, the motion simply renders to purely rotational; no translational motion is possible. The reason is also obvious. Because of the fixed point, the body cannot execute translational motion.

If the reference point is not predecided, the most convenient point to be chosen as the reference point is the centre of mass of the rigid body. We have also learnt earlier that under the choice of the reference point as the centre of mass, either the total kinetic energy or the total angular momentum can be separated to two neat pieces - one corresponding to translational motion of the centre of mass and the other to the rotation about the centre of mass.

The act of separation facilitates us to extend the method for considering the other aspects of the rigid body problems and the corresponding solutions separately for translational motion of the centre of mass and for rotational motion of the body about the centre of mass. Thus if we consider a rigid body system under holonomic constraints is acted on externally by a torque and if the latter involves a conservative force, we can separate out the Lagrangian into the so-called *Translational Lagrangian* $\mathcal{L}_c(q_c, \dot{q}_c)$ and the *Rotational Lagrangian* $\mathcal{L}_b(q_b, \dot{q}_b)$ involving the generalised coordinates q and generalised velocity $\dot{q}$ so that the total Lagrangian $\mathcal{L}(q, \dot{q})$ of the rigid body system is

$$\mathcal{L}(q, \dot{q}) = \mathcal{L}_c(q_c, \dot{q}_c) + \mathcal{L}_b(q_b, \dot{q}_b)$$

The generalised coordinates to be chosen to analyse the rotational motion are obviously the angular coordinates; the suitable orthogonal set of angles being the three euler angles as discussed earlier.

7.2.2 Deduction of Euler's equations of motion

For deducing the relevant equations govering the rotational motion of rigid bodies about a fixed point or the centre of mass, we have the direct Newtonian approach at hand. The Newton's second law of motion can be suitably adapted to take into account the rotational motion, *i.e.*, the momentum is replaced by the angular momentum $\vec{L}$ and the force part by the external torque $\vec{\tau}$. With respect to a fixed point, not necessarily inside the rigid body, the corresponding Newton's law adapted for rigid body motion is expressed as

$$\left(\frac{d\vec{L}}{dt}\right)_s = \vec{\tau} \tag{7.2.1}$$

where the subscript s referes the tme derivative with respect to a space set of axes, which do not share the rotation of the body. Now we want to transfer the results of the derivatives above, obtained with respect to the space set of axes to the body set of axes fixed in the body. For this, we use the relation already deduced (Equation (6.3.8)), so that

$$\left(\frac{d\vec{L}}{dt}\right)_s = \left(\frac{d\vec{L}}{dt}\right)_b + \vec{\omega} \times \vec{L},$$

Using this to Equation (7.2.1), we can write the equation of motion in terms of body set of axes as

$$\boxed{\frac{d\vec{L}}{dt} + \vec{\omega} \times \vec{L} = \vec{\tau}.} \tag{7.2.2}$$

wherein the subscript b is dropped, because henceforth the entire equation will be meant exclusively for the body set of axes. Equation (7.2.2) is the appropriate form of the Newtonian equation of motion for a rigid body relative to the body set of axes.

We can also write Equation (7.2.2) in the component form, so that the i-th component is expressed as

$$\frac{dL_i}{dt} + \epsilon_{ijk}\, \omega_j L_k = \tau_i \tag{7.2.3}$$

Here ϵ_{ijk} is the 3-indexed Levi-Civita symbol defined as

$$\epsilon_{ijk} = \begin{cases} 1 & \text{for } i,j,k \text{ cyclic} \\ -1 & \text{for } i,j,k \text{ anticyclic} \\ 0 & \text{if any two of } i,j,k \text{ coincide} \end{cases} \tag{7.2.4}$$

We further adopt the Einstein summation convention on indices that *repeated indices in a single term imply summation, unless otherwise specified.*

Let us now arrange the body set of axes for the principal axes of the rigid body relative to the chosen reference point. The angular momentum components in this case can then be expressed simply as $L_k = I_k\omega_k$ (no summation), so that the Equation (7.2.3) takes the form

$$\begin{aligned} &\frac{d}{dt}(I_i\omega_i) + \epsilon_{ijk}\omega_j\omega_k I_k = \tau_i && i = 1,2,3. \\ &I_i\frac{d\omega_i}{dt} + \epsilon_{ijk}\omega_j\omega_k I_k = \tau_i. && (\because \quad I_i \text{ is independent of time.}) \end{aligned} \tag{7.2.5}$$

Expanding (7.2.5) to all values of the index i, we have a set of three equations

$$\left.\begin{aligned} I_1\,\dot{\omega}_1 - \omega_2\,\omega_3\,(I_2 - I_3) = \tau_1 \\ I_2\,\dot{\omega}_2 - \omega_3\,\omega_1\,(I_3 - I_1) = \tau_2 \\ I_3\,\dot{\omega}_3 - \omega_1\,\omega_2\,(I_1 - I_2) = \tau_3 \end{aligned}\right\} \tag{7.2.6}$$

Equations (7.2.6) are called the *Euler's equations of motion* for a rigid body with one point fixed.

We now analyse a special case of the Euler's equations of motion of the rigid body. Let us consider that the rigid body rotates about a fixed axis, say the z-axis. This means, in the angular velocity vector $\vec{\omega} = \omega_1\hat{i} + \omega_2\hat{j} + \omega_3\hat{k}$, the component ω_1 about the x-axis and ω_2 about the y-axis obviously vanish; only the component ω_3 survives to maintain the rotation about the z-axis. Let $\omega_3 = \omega$ *i.e.*,

$$\omega_1 = \omega_2 = 0 \qquad \text{and} \qquad \omega_3 = \omega \text{ (say)}$$

The Euler's equations of motion (7.2.6) then reduces to

$$\tau_1 = \tau_2 = 0$$

and

$$\begin{aligned} \tau_3 &= I_3\dot{\omega}_3 \\ \Longrightarrow \qquad \tau &= I\dot{\omega} \end{aligned}$$

The angular momentum of the rigid body about the z-axis is given by

$$L_3 = I_3\omega_3$$

$$\text{or,} \qquad L = I\omega$$

and the instanteous rotational kinetic energy will be

$$T = \frac{1}{2}I_3\omega_3^2 = \frac{1}{2}I\omega^2$$

7.2.3 Torque-Free Motion

This is a special case in the Euler's equations of motion of a rigid body, if the component of the torque $\vec{\tau}$ along the principal axes of the rotating body is known, a formal solution to the Euler's equations of motion can be found. The motion of a free symmetric top is the simplest type of the motion of a rigid body. The torque acting on such a system is zero. It is noteworthy to mention here that a body is called a free symmetric top if $I_1 = I_2 \neq I_3$.
For such situation, we have

$$\tau_1 = \tau_2 = \tau_3 = 0$$

The Euler's equations of motion can then be written as

$$I_1\dot{\omega}_1 = (I_2 - I_3)\omega_2\omega_3 \qquad (7.2.7)$$

$$I_2\dot{\omega}_2 = (I_3 - I_1)\omega_3\omega_1 \qquad (7.2.8)$$

$$I_3\dot{\omega}_3 = (I_1 - I_2)\omega_1\omega_2 \qquad (7.2.9)$$

Multiplying equations (7.2.7), (7.2.8) and (7.2.9) respectively by ω_1, ω_2 and ω_3 and adding, we get

$$\begin{aligned} I_1\dot{\omega}_1\omega_1 + I_2\dot{\omega}_2\omega_2 + I_3\dot{\omega}_3\omega_3 &= (I_2 - I_3 + I_3 - I_1 + I_1 - I_2)\omega_1\omega_2\omega_3 \\ &= 0 \end{aligned}$$

which gives

$$\frac{d}{dt}\left(\frac{1}{2}I_1\omega_1^2 + \frac{1}{2}I_2\omega_2^2 + \frac{1}{2}I_3\omega_3^2\right) = 0$$

$$\Longrightarrow \qquad \frac{1}{2}I_1\omega_1^2 + \frac{1}{2}I_2\omega_2^2 + \frac{1}{2}I_3\omega_3^2 = \frac{1}{2}\vec{\omega}.\vec{L} \;=\; \text{a constant}$$

which is the principle of conservation of total rotational kinetic energy in absence of external torque.

Further,

$$\begin{aligned} \vec{\tau} &= \frac{d\vec{L}}{dt} = 0 \\ \Longrightarrow \qquad \vec{L} &= I_1\omega_1\hat{i} + I_2\omega_2\hat{j} + I_3\omega_3\hat{k} = \text{constant} \end{aligned}$$

which is the principle of conservation of angular momentum.

In the following we consider some special cases from the above analysis.

Case 1: The case of force-free motion of a symmetrical rigid body *i.e.*, a symmetrical top, for which $I_1 = I_2$ and the third principal axis (z-axis) is the axis of symmetry of the body. In this case, Euler's equations become

$$I_1\dot{\omega}_1 = (I_1 - I_3)\omega_2\omega_3 \tag{7.2.10a}$$
$$I_1\dot{\omega}_2 = (I_3 - I_1)\omega_3\omega_1 \tag{7.2.10b}$$
$$\dot{\omega}_3 = 0 \tag{7.2.10c}$$

From the last equation, (7.2.10c), we have

$$\omega_3 = \text{constant}$$

Now, multiplying the equations (7.2.10a) and (7.2.10b) by ω_1 and ω_2 respectively and adding, we get

$$\begin{aligned} I_1\dot{\omega}_1\omega_1 + I_1\dot{\omega}_2\omega_2 &= (I_1 - I_3 + I_3 - I_1)\,\omega_1\omega_2\omega_3 = 0 \\ \Rightarrow \qquad & \dot{\omega}_1\omega_1 + \dot{\omega}_2\omega_2 = 0 \\ \Rightarrow \qquad & \frac{1}{2}\frac{d}{dt}\left(\omega_1^2 + \omega_2^2\right) = 0 \\ i.e., \qquad & \omega_1^2 + \omega_2^2 = \text{constant} = \omega_c^2 \text{ (say,)} \end{aligned}$$

It is an equation of a circle with radius $w_c = \sqrt{\omega_1^2 + \omega_2^2}$. The total angular velocity vector is then given by

$$\vec{\omega} = \omega_1\hat{i} + \omega_2\hat{j} + \omega_3\hat{k},$$

with the magnitude ω, with ω_3=constant, is given as

$$\omega = |\vec{\omega}| = \sqrt{\omega_1^2 + \omega_2^2 + \omega_3^2} = \text{constant}.$$
$$i.e., \qquad \omega = \text{constant}$$

Thus, we can see that the angular velocity vector $\vec{\omega}$ rotates about the body z-axis and describe a cone with the vertex at the origin. This motion of the rigid body is known as the *precession* or the precessional motion; the body is said to precess about the z-axis with a precessional velocity that is decided by the values of I_1, I_3 and ω_3. The cone described by the angular velocity vector $\vec{\omega}$ is called the body cone, and its half angle η is given by

$$\tan\eta = \frac{\sqrt{\omega_1^2 + \omega_2^2}}{\omega_3} = \frac{\omega_c}{\omega_3}.$$

Case 2: If we put $\alpha = \dfrac{I_3 - I_1}{I_1}$ = constant, in equations (7.2.10a) and (7.2.10b), then

$$\dot{\omega}_1 = \left(\frac{I_1 - I_3}{I_1}\omega_3\right)\omega_2 = -\alpha\omega_2$$
$$\dot{\omega}_2 = \left(\frac{I_3 - I_1}{I_1}\omega_3\right)\omega_1 = \alpha\omega_1$$

From which it follows that

$$\begin{aligned} \ddot{\omega}_1 &= -\alpha\dot{\omega}_2 = -\alpha^2\omega_1 \\ \Rightarrow \omega_1 &= a\cos\alpha t + b\sin\alpha t \qquad a,\ b \text{ are constants of integration.} \\ &= c\sin(\alpha t + \beta) \end{aligned}$$

where $a = c\sin\beta$, $b = c\cos\beta$, $c = \sqrt{a^2 + b^2}$ and $\beta = \tan^{-1}\frac{a}{b}$.

Now, at $t = 0$, we take $\omega_1 = 0$, $\Rightarrow$ $\beta = 0$.

$$\therefore \quad \omega_1 = c\sin\alpha t.$$

Taking c =constant= ω_c, we get

$$\omega_1 = \omega_c\sin\alpha t.$$

Thus we see from the equation $\dot{\omega}_1 = -\alpha\omega_2$ in the beginning that

$$\alpha\omega_c\cos\alpha t = -\alpha\omega_2$$

$$\Rightarrow \quad \omega_2 = -\omega_c\cos\alpha t$$

which lead to $\boxed{\omega_1^2 + \omega_2^2 = \omega_c^2}$

Hence, ω_1 and ω_2 satisfy the equation for simple harmonic motion.

Note !

The Euler's equations of motion can also be derived from Lagrange's equations, as the constraints involved are holonomic and the forces involved are conservative. In this case first we have to construct the corresponding Lagrangian in terms of the generalised coordinates. The best suited three generalised coordinates are the Euler angles of rotation and the generalised forces are the associated torques. However, in this case we need to construct Lagrange's equation with respect to one generalised coordinates only, *i.e.*, only one Euler angle. The rest two equations can be constructed from the cyclic permutation of the corresponding indices.

7.2.4 The heavy symmetrical top

We now consider the motion of a symmetrical body in a uniform gravitational field when one point on the symmetry axis is fixed in space. There are numerous situations that we can find physical systems, from the children's playing top to complicated gyroscopic navigational instruments, are based on the analysis of such a heavy symmetrical top.

We consider that the symmetry axis of the top is one of the principal axes, which for convenience is chosen as the z-axis of the coordinate system fixed in the body. The relevant equations of motion for a symmetric top can be developed either by Newtonian approach, *i.e.*, writing the Newton's second law of motion in vector form, or one can adopt the Lagrangian approach, which consists of writing the relevant Lagrangian in terms of some independent, generalised coordinates and their time derivatives, called the generalised velocities and finally using the Lagrange's equations for holonomic systems. Here we follow the latter approach.

The configuration of the top is completely specified by the three Euler angles: θ gives the information of the z-axis from the vertical, ϕ measure the azimuthal angle of the top, while ψ is the rotation angle of the top about its own z-axis. These Euler angles form the generalised coordinates and it will be attempted to write the Lagrangian in terms of these angles and their derivatives. We denote the distance of the center of gravity from the fixed point by l.

So far as the generalised velocities are concerned, the time rate of change of the three Euler angles $(\dot{\phi}, \dot{\theta}, \dot{\psi})$ will serve the purpose. But these components will not be convenient for us to use directly; they must be transformed to usual 'Cartesian type' body set of axes $(x',\ y',\ z')$. This transformation can be effected through the orthogonal Eulerian rotation matrix $R(\phi, \theta, \psi)$. Thus for $\vec{\dot{\phi}} \equiv \vec{\omega}_\phi$ being parallel to the space z-axis, its components along the body set of axes will be found from

$$(\vec{\omega}_\phi)_{x'} = \dot{\phi}\sin\theta\sin\psi, \qquad (\vec{\omega}_\phi)_{y'} = \dot{\phi}\sin\theta\cos\psi, \qquad (\vec{\omega}_\phi)_{z'} = \dot{\phi}\cos\theta.$$

Similarly when $\vec{\dot{\theta}} \equiv \vec{\omega}_\theta$ are transformed, they take the following form

$$(\vec{\omega}_\theta)_{x'} = \dot{\theta}\cos\psi, \qquad (\vec{\omega}_\theta)_{y'} = -\dot{\theta}\sin\psi, \qquad (\vec{\omega}_\theta)_{z'} = 0.$$

and as the vector $\vec{\omega}_\psi$ lies along the z' directin, no transformation of the components of $\vec{\omega}_\psi$ are actually necessary. Thus the angular momentum components in the body set of axes are given by the addition of the above two transformations, *i.e.*,

$$\begin{aligned}
\omega_1 \equiv \omega_{x'} &= \dot{\phi}\sin\theta\sin\psi + \dot{\theta}\cos\psi \\
\omega_2 \equiv \omega_{y'} &= \dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi \\
\omega_3 \equiv \omega_{z'} &= \dot{\phi}\cos\theta + \dot{\psi}
\end{aligned}$$

Since the body is symmetrical, with the angular velocity components ($\omega_1 = \omega_{x'},\ \omega_2 = \omega_{y'},\ \omega_3 = \omega_{z'}$) and the Inertia terms ($I_1, I_2 = I_1, I_3$), the kinetic energy of the top can be written as

$$T = \frac{1}{2}I_1(\omega_1^2 + \omega_2^2) + \frac{1}{2}I_3\omega_3^2$$

or, in terms of Euler's angles,

$$T = \frac{I_1}{2}(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta) + \frac{I_3}{2}(\dot{\psi} + \dot{\phi}\cos\theta)^2 \tag{7.2.11}$$

At this point we refer to a well known elementary theorem that in a constant gravitational field the potential energy is the same as if the body were concentrated at the center of mass. The potential energy of the body, expressed in terms of the Euler angles, is given by,

$$V = Mgl\cos\theta,$$

Here, l is the distance of the centre of gravity from the fixed point. The Lagrangian of the system can now be written as

$$L = T - V = \frac{I_1}{2}(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta) + \frac{I_3}{2}(\dot{\psi} + \dot{\phi}\cos\theta)^2 - Mgl\cos\theta$$

Since the torque of gravity is along the line of nodes, there is no component of the torque along either the vertical or the body z-axis, because both of these axes are perpendicular to the line of nodes and hence the components of the angular momentum along these two axes must be constant in time. This is also obvious from the expression of the Lagrangian where the generalised coordinates ϕ, ψ do not explicitly appear in the Lagrangian. Two immediate first integrals of the motion will then follow:

$$P_\psi = \frac{\partial L}{\partial \dot{\psi}} = I_3\left(\dot{\psi} + \dot{\phi}\cos\theta\right) = I_3\omega_3 = I_1 a \tag{7.2.12}$$

and

$$P_\phi = \frac{\partial L}{\partial \dot{\phi}} = \left(I_1 \sin^2\theta + I_3 \cos^2\theta\right)\dot{\phi} + I_3\dot{\psi}\cos\theta = I_1 b \tag{7.2.13}$$

Here the two constants of the motion are expressed in terms of new constants a and b. Since the system is conservative the total energy E is constant in time, *i.e.*,

$$E = T + V = \frac{I_1}{2}(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta) + \frac{I_3}{2}w_3^2 + Mgl\cos\theta \tag{7.2.14}$$

From the equation (7.2.12) above, we see that $\dot{\psi}$ is related to $\dot{\phi}$ through

$$I_3\dot{\psi} = I_1 a - I_3\dot{\phi}\cos\theta \tag{7.2.15}$$

Substitute this in P_ϕ to eliminate $\dot{\psi}$:

$$I_1\dot{\phi}\sin^2\theta + I_1 a\cos\theta = I_1 b \tag{7.2.16}$$

or, $$\dot{\phi} = \frac{b - a\cos\theta}{\sin^2\theta} \tag{7.2.17}$$

Substituting equation $\dot{\phi}$ back in equation (7.2.15) results in a corresponding expression for $\dot{\psi}$:

$$\dot{\psi} = \frac{I_1 a}{I_3} - \left(\frac{b - a\cos\theta}{\sin^2\theta}\right)\cos\theta \tag{7.2.18}$$

Finally, equations (7.2.17), (7.2.18) are used to eliminate $\dot{\phi}$ and $\dot{\psi}$ from the energy equation (7.2.14), which results in a differential equation involving θ alone. Further, ω_3 is found to equal $\left(\frac{I_1}{I_3}\right)a$, a constant in time. Therefore $E - I_3\frac{\omega_3^2}{2}$ has to be a constant of motion, denoted by E'. The energy equation, then turns out to be

$$E' = E - I_3\frac{\omega_3^2}{2} + \frac{I_1\dot{\theta}^2}{2} + \frac{I_1}{2}\frac{(b - a\cos\theta)^2}{\sin^2\theta} + Mgl\cos\theta$$

This equation has the form of an equivalent one dimensional problem in the variable θ, with the effective potential energy function $V'(\theta)$ given by

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

so that, the relavant differential equation for the '*theta*-motion', becomes

$$\dot{\theta}^2 + \left(\frac{b - a\cos\theta}{\sin\theta}\right)^2 + \frac{2Mgl\cos\theta}{I_1} - \frac{2E'}{I_1} = 0$$

$$\Longrightarrow \qquad \dot{\theta}^2 + \frac{2(E' - V'(\theta))}{I_1} = 0 \tag{7.2.19}$$

Thus the dynamics of the symmetric top motion can be understood in principle once we solve this non-linear differential equation in θ, *i.e.*, find θ as a function of time t.

To solve the differential equation (7.2.19), we first write

$$\dot{\theta} = \frac{d\theta}{dt} = \sqrt{\frac{2}{I_1}[E' - V'(\theta)]}$$

or,

$$t(\theta) = \int \frac{d\theta}{\sqrt{\frac{2}{I_1}[E' - V'(\theta)]}}$$

We can then invert $t(\theta)$ to find $\theta(t)$, the angle as a function of time and using this to solve for $\phi(t)$ and $\psi(t)$ from the equations (7.2.17) and (7.2.18). But this method of solutions are associated with elliptic integrals and hence are difficult to solve. On the face of the difficulty, it will be worthwhile to obtain some qualitative features of motion by inspection of the differential equatons, without actually performing the integrals. We can understand the behaviour of the effective potential function $V'(\theta)$ by plotting it against the angle θ. It is seen that the function $V'(\theta)$ is infinite at $\theta = 0,\ \pi$, and finite in between. Further, $V'(\theta)$ is minimum at a value $\theta = \theta_0$ in the range and can be found by setting the first derivative of $V'(\theta)$ equal to zero. So,

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

and hence

$$\left.\frac{dV'}{d\theta}\right|_{\theta=\theta_0} = 0$$

implies

$$-Mgl\sin\theta_0 + \frac{I_1(b - a\cos\theta_0)(a - b\cos\theta_0)}{\sin^3\theta_0} = 0 \tag{7.2.20}$$

Let us define $\chi = I_1(b - a\cos\theta_0)$ so that we can simplify (7.2.20) as

$$\chi^2\cos\theta_0 - \chi I_1 a\sin^2\theta_0 + MglI_1\sin^4\theta_0 = 0$$

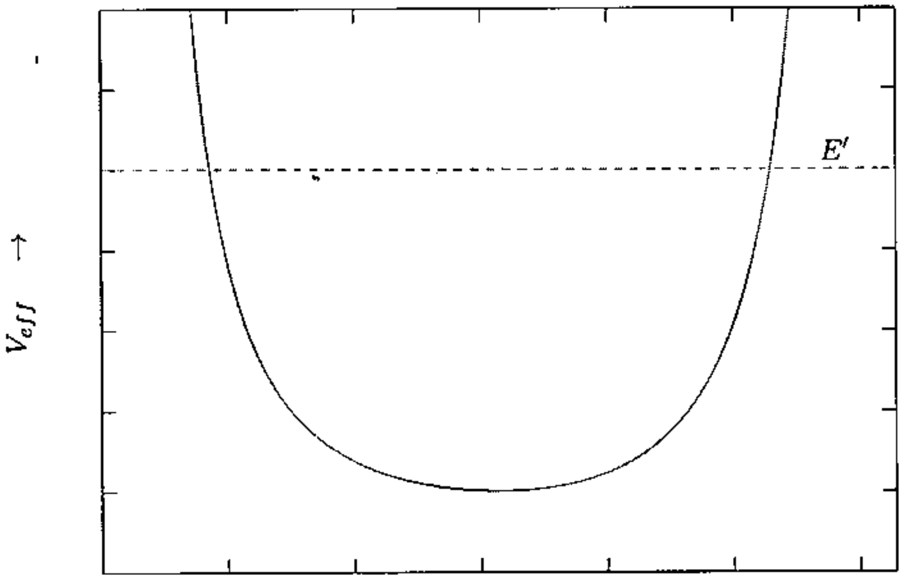


Figure 7.1: Variation of the effective potential

which is a quadratic equation in χ and hence the roots are given by

$$\chi = \frac{I_1 a \sin^2\theta_0}{2\cos\theta_0}\left(1 \pm \sqrt{1 - \frac{4Mgl\cos\theta_0}{I_1 a^2}}\right) \tag{7.2.21}$$

We demand that χ is real, so the quantity inside the radical sign must yield a positive number. Now for any choice of ours as $\theta_0 < \frac{\pi}{2}$ we find

$$1 - \frac{4Mgl\cos\theta_0}{I_1 a^2} \geq 0 \qquad \Longrightarrow \qquad I_1 a^2 \leq 4Mgl\cos\theta_0$$

Referring to the equation (7.2.12), we see that $P_\psi = I_3\omega_3 = I_1 a$. Substituting the value of a, we can see that there exists a lower bound for ω_3:

$$\omega_3 \geq \frac{2}{I_3}\sqrt{MglI_1\cos\theta_0},$$

which suggests that a steady precession is possible at the fixed angle of inclination θ_0, only when the angular velocity of spin exceeds a limiting value.

7.2.5 Precession and Nutation of Earth

Precession

In astronomy, axial precession is a gravity-induced, slow, and continuous change in the orientation of an astronomical body's rotational axis. In particular, it refers to the gradual shift in the orientation of Earth's axis of rotation, which, similar to a wobbling top, traces out a pair of cones joined at their apices in a cycle of approximately 26,000 years.

Precession occurs because:

1. The Earth is rotating,

2. The Earth is not exactly spherical; it has a slight equatorial bulge, and the gravitational fields of the Moon, Sun and planets affect to produce precession.

Nutation

Nutation is a rocking, swaying, or nodding motion in the axis of rotation of a largely axially symmetric object. Nutation takes place because of tidal forces that cause the precession of the equinoxes to vary over time so that the speed of precession is not constant; principal sources of tidal force are the Sun and Moon, which continuously change locations relative to each other and thus cause nutation in Earth's axis.

Precession of the Equinoxes and Satellite

In a broad sense, the earth can be considered a top with the axis precessing about the normal to the ecliptic. The earth is not a perfect sphere, but slightly distorted so that it can be approximated by an oblate spheroid of revolution. It is just the net torque on the resultant equatorial bulge arising from gravitational attraction, chiefly of the sun and the moon, that sets the earth's axis precessing in space.

In order to deduce the kinematics and dynamics of precession of equinoxes, we consider a mass distribution forming a single body, wherein we take a single mass point with mass M. If r_i is the distance between the i-th point in the distribution and the mass point M, then the mutual gravitational potential between the two bodies is

$$V = -\frac{GMm_i}{r_i}$$

It is well known that a simple expansions in terms of Legendre polynomials can be given so that

$$V = -\frac{GM}{r}\sum_{n=0} m_i \left(\frac{r_i'}{r}\right)^n P_n(\cos\psi_i) \tag{7.2.22}$$

Providing r, the distance from the origin to M, is greater than any r_i', we shall make use of only the first three Legendre polynomials that,

$$P_0(x) = 1, \quad P_1(x) = x, \quad P_2(x) = \frac{1}{2}(3x^2 - 1) \quad \text{and so on.}$$

From the orthonormal properties of P_n with respect to P_0, the integral over $\cos\psi$ vanishes except for $n = 0$, which proves the statement.

If a body deviates only slightly from spherical symmetry, as is the case with the earth, one would expect the terms in equation (7.2.22) beyond $n = 0$ to decrease rapidly with increasing n. It will therefore be sufficient to retain only the first non vanishing term in (7.2.22) to the potential for a sphere. Now, the choice of the center of mass as origin cause the $n = 1$ term to vanish identically, since it can be written

$$-\frac{GM}{r^2} m_i r_i' \cos\psi_i = -\frac{GM}{r^3} r m_i r_i'$$

Which is zero. by definition of the center of mass. The next term, for $n = 2$, can be written

$$\frac{GM}{2r^3} m_i r_i'^2 (1 - 3\cos^2\psi_i)$$

It is useful to write $r_i'^2 \cos^2\psi_i$the expression in dyadic form :

$$r_i'^2 \cos^2\psi_i = \frac{r r_i' r_i' r}{r^2}$$

So that, with a little judicious addition and subtraction, the $n = 2$ term in the potential takes the form

$$\frac{2}{3}\frac{GM}{r^5} m_i r \left[r_i'^2 - r_i' r_i'\right] r - \frac{GM}{r^3} m_i r_i'^2$$

The second part is to seen to involve the trace of the inertia tensor. We can therefore write the $n = 2$ term as

$$\frac{3}{2}\frac{GM}{r^5} r.I.r - \frac{GM}{2r^3} TrI$$

And the complete approximation to the non- spherical potential as

$$V = -\frac{GMm}{r} + \frac{GM}{2r^3}[3I_r - TrI]$$

Where m is the mass of the first body and I_r is the moment of inertia about the direction of r. From the diagonal representation of the inertia tensor in the principal axis system, its trace is just the sum of the principal moments of inertia. So that V can be written as

$$V = \frac{GMm}{r} + \frac{GM}{2r^3}[3I_r - (I_1 + I_2 + I_3)] \qquad (7.2.23)$$

Lets now take the z-axis of symmetry to be along the third principal axis, so that $I_1 = I_2$. If α, β, γ are the direction cosines of r relative to the principal axes, then the moment of inertia I_r can be expressed as

$$I_r = I_1(\alpha^2 + \beta^2) + I_3\gamma^2 = I_1 + (I_3 - I_1)\gamma^2$$

With this form for I_r, the potential, equation becomes

$$V = -\frac{GMm}{r} + \frac{GM(I_3 - I_1)}{2r^3}[3\gamma^2 - 1] = -\frac{GMm}{r} + \frac{GM(I_3 - I_1)}{2r^3} P_2(\gamma)$$

Of the terms in equation for the potential, the only one that depends on the orientation of the body, and thus could give rise to torques, is

$$V_2 = \frac{GM(I_3 - I_1)}{2r^3} P_2(\gamma)$$

For example as the earths precession, it should be remembered that γ is the direction cosine between the figure axis of the earth and the radius vector from the earths center to the sun or moon. As these bodies go around their apparent orbits will change. The relation of γ to the more customary astronomical angles can be seen from fig where the orbit of the sun or moon is taken as being in the xy plane, and the figure axis of the body in the xz plane. The angle θ between the figure axis and direction is the obliquity of the figure axis so that

$$\gamma = \sin\theta\cos\eta$$

Hence V_2 can be written

$$V_2 = \frac{GM(I_3 - I_1)}{2r^3}[3\sin^2\theta\cos^2\eta - 1]$$

The averaged potential is then

$$\bar{V_2} = \frac{GM(I_3 - I_1)}{2r^3}\left[\frac{3}{2}\sin^2\theta - 1\right] = \frac{GM(I_3 - I_1)}{2r^3}\left[\frac{1}{2} - \frac{3}{2}\cos^2\theta\right]$$

Or, finally,

$$\bar{V_2} = \frac{GM(I_3 - I_1)}{2r^3}P_2(\cos\theta)$$

where $P_2 = \left[\frac{1}{2} - \frac{3}{2}\cos^2\theta\right]$ is the Legendre polynomial of order 2. It is seen that the torque derived from the above equation is perpendicular to both i-th figure axis and the normal to the orbit. Hence the precession is about the direction of the orbit normal vector. For any symmetric body in which the potential is a function of $\cos\theta$ only, the Lagrange can be written, following, as

$$L = \frac{I_1}{2}(\dot{\theta}^2 + \dot{\phi}^2\sin^2) + \frac{I_3}{2}(\dot{\psi} + \dot{\phi}\cos\theta)^2 - V(\cos\theta)$$

The lagrange equation corresponding to θ is then

$$\frac{\partial L}{\partial\theta} = I_1\dot{\phi}^2\sin\theta\cos\theta - I_3\dot{\phi}\sin\theta(\dot{\psi} + \dot{\phi}\cos\theta) - \frac{\partial V}{\partial\theta} = 0$$

For slow precession, which means basically that $\dot{\phi} \ll w_3$, the $\dot{\phi}^2$ can be neglected, and the rate of uniform precession is given by

$$\dot{\phi} = \frac{1}{I_3 w_3}\frac{\partial V}{\partial(\cos\theta)}$$

With the potential of equation the precession rate is

$$\dot{\phi} = -\frac{3GM}{2w_3 r^3}\frac{I_3 - I_1}{I_3}\cos\theta$$

For the case of the precession due to the sun, this formula can be put in a simpler form, by taking r as the semi major axis of the earths orbit and using Kepler's law, in the form

$$w_0^2 = \left(\frac{2\pi}{\tau}\right)^2 = \frac{GM}{r^3}$$

The precession rate, relative to the orbital angular velocity, w_0 is then

$$\frac{\dot{\phi}}{w_0} = -\frac{3}{2}\frac{w_0}{w_3}\frac{I_3 - I_1}{I_3}\cos\theta \tag{7.2.24}$$

With the value of $\left(\frac{I_3 - I_1}{I_3}\right)$ and $\theta = 23^0 27'$, says that the solarinduced precession would be such as to cause a complete rotation of the figure axis about normal to the ecliptic in about 81,000 years. The moon is so far less massive than the sun but it is also much closer; the net result is that the lunar induced precession rate is over twice that caused by the sun. The two precessions nearly add together arithmetically, and the combined lunisolar precession rate is 50.25year, or one complete rotation in about 26,000 years. This rate of precession is so slow that the approximation of neglecting compared to is abundantly satisfied. Because the sun, moon, and earth are in constant relative motion, and the moons orbit is inclined about to the ecliptic, the precession exhibits irregularities designated as astronomical nutation.

Effects of precession

The precession of the Earth's axis has a number of observable effects. First, the positions of the south and north celestial poles appear to move in circles against the space-fixed backdrop of stars. Secondly, the position of the Earth in its orbit around the Sun at the solstices,equinoxes, or other time defined relative to the seasons, slowly changes. celestial pole, this will change over time, and other stars will become the "north star".

7.3 Summary

In this unit we have studied the inertia tensor and its properties which are useful for further analysis of the rigid body motion. The dynamical equations for the motion of a rigid body, the Euler's equations have been deduced with the help of the vectors under rotations of the coordinates system, or the non-inertial frame that we studied in the previous chapter. These equations then are found to be helpful for exploring the cases of torque-free motion, the motion of symmetrical tops etc. Finally, the use of the Euler's equations are extended to the study of the precessional and nutational motion of the earth that are induced as a result of the spinning motion of the earth. The Euler's equations find its use in the study of the precession of the equinox and the satellite motion.

Self study questions:

1. What is understood by the inertia tensor of a rigid body? Analyse the case of the force free motion of arigid body with the cmponents of inertia related by $I_1 = I_2 = 2I_3$. Obtain the expression for the frequency of precession of the angular velocity about the axis of symmetry.

2. Derive Euler's equations of motion for a rigid body with a fixed point.

3. If a rigid body, with one point fixed, rotates with an angular veloctty $\vec{\omega}$ and has an angular momentum $\vec{L}$, show that the kinetic energy of the rigid body is $\frac{1}{2}\vec{L}\cdot\vec{\omega}$.

4. A rigid body of with its principal components of inertia as (I_1, I_2, I_3) possesses a rotational kinetic energy, say T. If the external torque applied to the rigid body is $\vec{\tau}$ and the resulting angular velocity vector is $\vec{\omega}$, prove that

$$\frac{dT}{dt} = \vec{\tau}\cdot\vec{\omega}.$$

UNIT 8

Theory of small oscillations

Preparatory inputs to this unit

1. Lagrange's equation of motion.
2. Matrix Algebra: Eigen equations, Eigenvalues and Eigenvector.
3. Basics of Integral Calculus and Ordinary differential equations.

Introduction

A particle or a system of particles forming a body is said to be in equilibrium if the vector addition of the forces acting on it is zero, *i.e.*, the resultant force acting on the body vanishes and the body cannot execute any motion. But the mere equilibrium condition does not guarantee a body to continue to be in equilibrium. Three possibilities might arise when the body already under equilibrium is disturbed or perturbed by some small amount.

1. The body might come back to the original equilibrium position or configuration, when the source of perturbation is removed.
2. The body will move away further and further from the equilibrium without any possibility to come back and restore the original equilibrium position or configuration.
3. The body might take up a new equilibrium position or configuration.

The kind of equilibrium in the first case is called the *stable equilibrium*, the second the *unstable equilibrium* and the third is the *neutral equilibrium*.

We will be concerned more here with the stable equilibrium and attempt to understand the mechanism of returning back to equilibrium upon removal of the perturbating agency. Under stable equilibrium, the perturbed system will generate a restoring force and try to regain its original equilibrium, but by the time the system will overshoot the equilibrium and again generates a back restoring force and it goes on. Essentially the system will execute a periodic motion prior to finally settle down to the original equilibrium. If the restoring force is proportional to the displacement from the equilibrium, and also the amplitude of oscillation is small, the system will undergo what is known as the simple harmonic motion, with definite period, amplitude etc. It is possible thus to know the nature of the restoring force along with the parameters related to the internal structure or configuration of the system, by studying the simple harmonic motion. This chapter is mostly dedicated to the discussions in this lines analysing the various situations where simple harmonic motions are possible.

The theory of small oscillations is extremely important in several areas of physics. *e.g.*, molecular spectra, acoustics, vibrations of atoms, coupled mechanical oscillators and electrical circuits etc. Here mainly the motion of the system about the position of stable equilibrium is discussed. Before this, let us stable and unstable equilibrium. For that we consider a conservative system in which potential energy is a function of position only. The system is said to be equilibrium when the generalized force acting on the system vanish:

$$Q_i = -\left(\frac{\partial V}{\partial q_i}\right)_0 = 0, \qquad i = 1, 2, 3, \ldots n$$

and q_i's are generalised co-ordinates. The potential energy does possess an extremum at the equilibrium configuration of the system. We can cite nuerous examples from our day-to-day experience. A pendulum at rest, a suspension of galvanometer at its zero position. As already said, an equilibrium position is classified depending on the behaviour under a small disturbance to the system from equilibrium : whether a small bounded motion about the rest position or an unbound motion ensures. The equilibrium is unstable if an infinitesimal disturbance eventually produces unbounded motion. As an illustration we can take a round bottom bowl and a marble. The marble at the bottom of an upright bowl is at rest. If disturbed slightly, it executes couple of periodic bounded motions about the equilibrium, eventually to be at the original equilibrium again, and is therefore under stable equilibrium. But the same marble, if managed to be kept at the top of the inverted bowl, will move away once a small disturbance is given to it. This system therefore forms an unstable equilibrium system.

8.1 One dimensional oscillator

The motion of the system about the position of stable equilibrium is of great interest in varied branches of science and engineering. As a simple case, we first consider a system to possess one degree of freedom with one generalized co-ordinate, say q. For small displacement from the equilibrium, we can expand the potential energy $V(q)$ in Taylor series about the equilibrium and consider only leading order terms:

$$V(q) = V(q_0) + \left[\frac{\partial V}{\partial q}\right]_0 (q - q_0) + \frac{1}{2}\left[\frac{\partial^2 V}{\partial q^2}\right]_0 (q - q_0)^2 + \cdots$$

At the position of equilibrium $\left(\frac{\partial V}{\partial q}\right)_{q=q_0} = 0$. The first term of the above equation $V(q_0)$ vanishes on shifting the origin of the potential energy curve to be at minimum equilibrium value. Thus

$$V(q) = \frac{1}{2}\left(\frac{\partial^2 V}{\partial q^2}\right)_{q=q_0} (q - q_0)^2$$

If we substitute $\left(\frac{\partial^2 V}{\partial q^2}\right)_0 = k$ and take the origin of q co-ordinate at $q_0 = 0$ then the equation reduces to

$$V(q) = \frac{1}{2}kq^2$$

where $k = \left(\frac{\partial^2 V}{\partial q^2}\right)_0$ is a positive quantitity at the stable equilibrium position. In this case, the potential energy does exclusively depend on the generalised co-ordinates and does not involve time explicity. Further, since in most cases the kinetic energy is a homogenous quadratic function of generalised velocities *i.e.*,

$$T = \frac{1}{2}m(q)\dot{q}^2, \tag{8.1.1}$$

where the co-efficient $m(q)$ is general function of q-co-ordinates and may be expanded in Taylor series about the equilibrium position $q_0 = 0$,

$$m(q) = m(0) + \left(\frac{\partial m}{\partial q}\right)_0 q + \cdots$$

Equation (8.1.1) is already quadratic in $\dot{q}$, the lowest non-vanishing approximation to T is obtained by retaining only the first term in the expansion of $m(q)$; *i.e.*, $m(0)$. Thus for small oscillations, the Lagrangian of the one-dimensional oscillation is given by

$$L = T - V = \frac{1}{2}m(0)\dot{q}^2 - \frac{1}{2}kq^2$$

and the equation of motion is

$$\frac{d}{dt}\left[\frac{\partial L}{\partial \dot{q}}\right] - \frac{\partial L}{\partial q} = 0 \tag{8.1.2}$$

$$m(0)\frac{d^2q}{dt^2} + kq = 0 \tag{8.1.3}$$

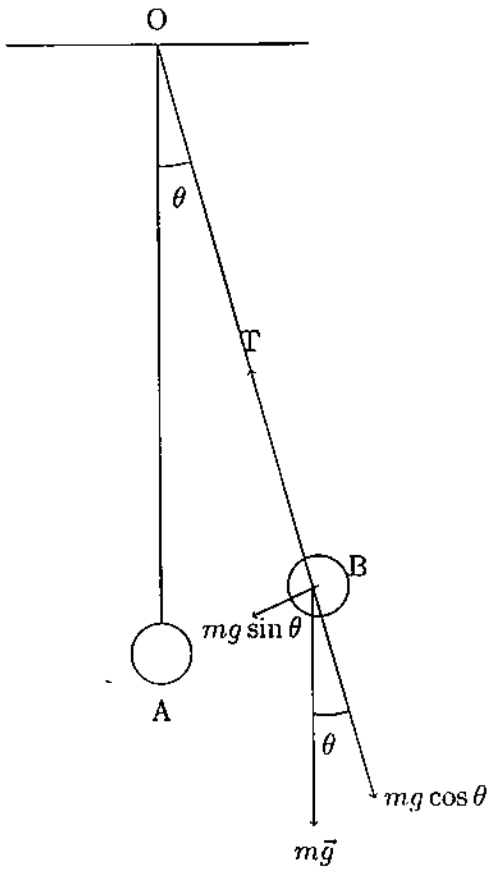
$$\frac{d^2q}{dt^2} + \omega^2 q = 0 \quad \text{where} \quad \omega^2 = \frac{k}{m(0)} \tag{8.1.4}$$

The solution of the above equation is $q = c_1 e^{i\omega t} + c_2 e^{i\omega t}$ or, in real quantities, $q = a\cos(\omega t + \phi)$ where constant a is called the amplitude of oscillation, ω is the angular frequency and ϕ is the initial phase.

8.1.1 Simple Pendulum

In order to proceed systematically, we first discuss the simplest oscillating system of a simple pendulum and gradually extend the concepts towards a more complicated systems oscillating under varied environments. As we have already discussed in the early units of the course, a simple pendulum is a heavy point mass suspended by a weightless, inextensible and a perfect flexible string from a rigid support about which, it can make to and fro motion freely.

Let the bob of the pendulum be displaced from its mean position and release for free motion. Suppose at any instant of time t, the bob has the mass m. The force acting on the bob vertically downward = mg. We resolve mg into two perpendiuclar components:



1. Force along the string $=mg\cos\theta$
2. Force perpendicular to the string $=mg\sin\theta$

Let the tension in the string be F_T. The component $mg\cos\theta$ balances the tension F_T. Hence,

$$mg\cos\theta = F_T$$

Thus the only component of the force acting on the oscillating bob to bring it to the equilibrium position is $-mg\sin\theta$. Therefore,

$$F = -mg\sin\theta$$

(Here, the negative sign signifies that the acceleration is directed towards the mean position.)

Taylor's series expansion of $\sin\theta$ yields

$$\sin\theta = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} \ldots\ldots\ldots\ldots$$

For small angular displacements in θ, we can approximate $\sin\theta$ to θ, *i.e.*, $\sin\theta \approx \theta$ so that.the tangential force is $F = -mg\theta$ and the linear displacement to be $x = l\theta$.

Cosidering the length of the pendulum to be l, the acceleration of the bob can be calculated as

$$a = \frac{d^2x}{dt^2} = l\frac{d^2\theta}{dt^2}$$

and the Force,

$$F = ml\frac{d^2\theta}{dt^2}.$$

From Newton's second law, we have

$$ml\frac{d^2\theta}{dt^2} = -mg\theta,$$

$$\frac{d^2\theta}{dt^2} + \frac{g}{l}\theta = 0 \tag{8.1.5}$$

This equation is similar to the equation of simple harmonic motion

$$\frac{d^2x}{dt^2} + \omega^2 x = 0 \tag{8.1.6}$$

From (8.1.5) and (8.1.6),

$$\omega^2 = \frac{g}{l},$$

or,

$$\omega = \sqrt{\frac{g}{l}}$$

Therefore, the time period of the oscillation is given by

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

Exercises: Simple Pendulum

Example 8.1.1 *Show that the mean kinetic and potential energies of non dissipative simple harmonic vibrating systems are equal.*

Solution: For free vibration in the absence of damping, the displacement at any instant is given by

$$y = a\sin\omega t$$

$$\frac{dy}{dt} = a\omega\cos\omega t$$

$$\begin{aligned}\text{Kinetic energy} &= \frac{1}{2}\times m\left(\frac{dy}{dt}\right)^2\\ &= \frac{1}{2}m(a^2\omega^2\cos^2\omega t)\\ &= \frac{1}{2}Ka^2\cos^2\omega t\end{aligned}$$

Here,

$$K = m\omega^2, \qquad \text{or,} \qquad \omega^2 = \frac{K}{m}$$

and K is the force per unit displacement. Now,

$$\begin{aligned}\text{Potential energy} &= \frac{1}{2}Ky^2 \\ &= \frac{1}{2}Ka^2 sin^2\omega t\end{aligned}$$

Total kinetic energy for one complete cycle is

$$\begin{aligned}\bar{T} &= \int_0^T \frac{1}{2}Ka^2\cos^2\omega t dt \\ &= \frac{1}{4}Ka^2T\end{aligned}$$

Total potential energy for one complete cycle

$$\begin{aligned}\bar{W} &= \int_0^T \frac{1}{2}Ka^2 sin^2\omega t dt \\ &= \frac{1}{4}Ka^2T\end{aligned}$$

Hence the mean kinetic and potential energies are equal.

Example 8.1.2 *A spring is hung vertically and loaded with a mass of 100 grams and allowed to oscillate. Calculate (a) the time period and (b) the frequency of oscillation, when the spring is loaded with 200 grams it extends by 10 cm.*

Solution: Here $M = 100$ grams, $m = 200$ grams, $x = 10$ cm, $g = 980 \ \frac{cm}{s^2}$

(a) Let k be the spring constant of the given spring. When it is loaded with mass m the force is $F = mg$ and the resulting extension is x. So,

$$F = -kx \quad \Longrightarrow \quad mg = -kx \quad \Longrightarrow \quad |k| = \frac{mg}{x}$$

As the spring of mass M oscillates with its characteristic angular frequency ω, we have

$$\omega = \sqrt{\frac{|k|}{M}} \quad \Longrightarrow \quad \frac{2\pi}{T} = \sqrt{\frac{mg}{Mx}}$$

where T is the time period and it is given by

$$T = 2\pi\sqrt{\left(\frac{Mx}{mg}\right)}$$

$$T = 2\pi\sqrt{\left(\frac{100 \times 10}{200 \times 980}\right)} = \frac{2\pi}{14} = 0.449 \ \ s$$

(b) Frequency ν is calculated from

$$\nu = \frac{1}{T} = \frac{1}{0.449} = 2.22 \quad \text{hertz.}$$

Example 8.1.3 *A uniform spring of force constant k is cut into two pieces of equal length. What is the force constant of each piece?*

Solution: Suppose, Force $= F$, Increase in length $= l$ and $k = \frac{F}{l}$
In the second case, we may consider the two halves of the spring pieces to be joined. On application of the force F, the extensions in each of the spring pieces will be $l/2$ so that the total extension in the entire length will be l. Now if k_h be the spring constant of the each half then for the same applied force F, we have $F = -k_h x/2$ for each piece. The enstensions accordingly for each piece is given by $x/2 = -F/k_h$. Adding the extensions of each piece, we must get the same effect of the spring before the cut. So, we get

$$\frac{x}{2} + \frac{x}{2} = -\frac{F}{k_h} - \frac{F}{k_h}$$

so that

$$x = -\frac{2F}{k_h} = -\frac{F}{\frac{k_h}{2}} = -\frac{F}{k},$$

Therefore $k_h = 2k$, which shows that the spring constant of one half of a cut piece from a spring always doubles its original value.

Example 8.1.4 *A body of mass 0.5 kg is suspended from a spring of negligible mass and it stretches the spring by 0.7 m. For a displacement of 0.03 m, it has a downward velocity of 0.4 m/s. Calculate (i) the time period; (ii) the frequency and (iii) amplitude of vibration of the spring.*

Solution: Here $m = 0.5kg$, $x = 0.07m$, $M = 0.5kg$, $g = 9.8m/s^2$
(*i*) Time period

$$T = 2\pi\sqrt{\frac{Mx}{mg}} = 2\pi\sqrt{\frac{0.5 \times 0.07}{05 \times 9.8}} = 0.5311 \text{ sec}$$

(*ii*) Frequency

$$n = \frac{1}{T} = \frac{1}{0.5311} = 1.882 \quad \text{Hz.}$$

Angular frequency

$$\omega = \frac{2\pi}{T} = \frac{2\pi}{0.5311} = 11.8 \quad \text{radians/sec}$$

(*iii*) For amplitude we first consider the solution $y = a\sin\omega t$.

$$\frac{dy}{dt} = a\omega\cos\omega t$$

$$\frac{dy}{dt} = a\omega\sqrt{1 - \frac{y^2}{a^2}} = \omega\sqrt{a^2 - y^2}$$

Now, with $y = 0.03$ m and $w = 11.8$ radians/sec we have,

$$0.4a = 11.8\sqrt{a^2 - (0.03)^2}$$

Simplifying the expression, we have, $a = 0.04526$ m.

8.1.2 Damped harmonic oscillator

We have discussed in the earlier section about a harmonic oscillator which oscillates freely upon the action of an external force. The external force actually generates an internal restoring force which acts in the opposite direction of the acceleration and is proportional to the amount of extension, *i.e.*,

$$\vec{F} = -k\vec{x}$$

where k is a positive constant.

But if the space around the oscillator is filled with materials, such as thick gas, fluid etc. which offers resistance in the form of friction, to the movement, the oscillator will not free oscillate. The oscillator in fact, cannot maintain the oscillation with same amplitude; the amplitude progressively reduces, eventually to stop. Such a motion of oscillator is called the damped harmonic motion. The damping force is normally proportional to the current velocity of the oscillator. The nature of such damping in oscillators crucially depends on the coefficients of friction of the medium in which the oscillator executes its periodic motion. Depending on the friction coefficient the oscillating system can have varieties of motions:

1. Oscillates with a frequency smaller than in the non damped case,and an amplitude decreasing with time (under damped oscillator)
2. Decay to the equilibrium position, without oscillations (over damped oscillator)

In between the boundary of these cases of underdamped oscillator and overdamped oscillator, there also exists a solution which is found for a particular value of the friction coefficient. The oscillators of this type is called critically damped.

We express the damped oscillator mathematically by Newton's 2nd law of motion, as

$$ma + cv + Kx = 0$$
$$m\frac{d^2x}{dt^2} + c\frac{dx}{dt} + Kx = 0$$

This is in the form of a homogeneous second order differential equation and has a solution of the form

$$x = e^{\lambda t}$$

Substituting this form gives an auxiliary equation of λ as:

$$m\lambda^2 + c\lambda + K = 0,$$

The roots of the quadratic auxiliary equation are

$$\lambda = -c \pm \sqrt{\frac{c^2 - 4mK}{2m}} = 0$$

The three resulting cases for the damped oscillator are :

1. *Overdamped* : The motion of the oscillation is said to be overdamped when the condition $c^2 - 4mK > 0$ is satisfied.
2. *critically damped* : In this case the oscillator satisfies the condition $c^2 - 4mK = 0$.
3. *underdamped* : The oscillator executes underdamped motion satisfying the condition $c^2 - 4mK < 0$.

Example

Mass Spring Damper:
An ideal mass spring damper system with mass m, spring constant K and viscous damper of damping coefficient c is subject to a periodic force $F_s = -Kx$ and a damping force $F_d = -cv = -c\frac{dx}{dt}$. The values can be in any consistent system of units; for example, in SI units, m in Kilograms, K in newtons per metre, and c in Newton-seconds per metre or Kilogram per second.

Treating the mass as a free body and applying Newton's second law, the total force F_{tot} on the body is

$$F_{tot} = ma = m\frac{d^2x}{dt^2},$$

where a is the acceleration of the mass and x is the displacement of the mass relative to a fixed point of reference. Since $F_{tot} = F_s + F_d$, we have

$$m\frac{d^2x}{dt^2} = -Kx - c\frac{dx}{dt}$$

This differential equation may be rearranged into

$$\frac{d^2x}{dt^2} + \frac{c}{m}\frac{dx}{dt} + \frac{K}{m}x = 0$$

The following parameters are then defined: as $\omega_0 = \sqrt{\frac{K}{m}}$ and

$$\gamma = \frac{c}{2\sqrt{mK}},$$

the first parameter, ω_0 is called the (undamped) natural frequency of the system. The second parameter γ is called the damping ratio. The natural frequency represents an angular frequency, expressed in radians per second. The damping ratio is a dimensionless quantity
The differential equation now becomes

$$\frac{d^2x}{dt^2} = 2\gamma\omega_0\frac{dx}{dt} + \omega_0^2 x = 0$$

Continuing, we can solve the equation by assuming a solution x such that

$$x = e^{\beta t}$$

where the parameter γ is, in general a complex number.

8.2 Coupled Oscillation

Having discussed about the harmonic motions involving one degree of freedom we now discuss a more complex system of harmonic oscillation with more than one degree of freedom. With more degrees of motion, there is always a possibility that the behavior of each variable influences the motion of the others. This leads to a coupling of the oscillations of the individual degrees of freedom. Coupled Oscillations occur when two or more oscillating systems are connected in such a manner as to allow energy to be exchanged between various degrees of freedom. This phenomenon was first observed by Christian Huygens in 1665.

We define a coupled oscillator as a physical system that contains multiple components of motion connected together and free to move in consistence with the given constraints. The theory and the results of coupled oscillations finds many important applications in molecular physics *viz.*, in studying molecular vibrations of atoms.

A few basic examples are -

1. double pendulums: one pendulum connected to the *bob* of another pendulum.

2. Solids and fluids are another good examples of systems that are beset with coupled oscillations. The molecules mostly oscillate around their equilibrium positions, along with the interactions amongst them to be coupled with each other. If we look at the configuration of a carbon dioxide molecules, two outer oxygen atoms are bound with forces of electrostatic origin to the central carbon atom and the oscillation of any one component does influence that in other components. Another example of such a coupled system is a crystalline solid in which the atoms constituting the crystal interacts with other via interatomic forces.

In order to illustrate the mechanics of coupled oscillations, and provide a framework for subsequent extensions to oscillating systems with multiple degrees of freedom, we work out here a system of two objects connected end to end by three springs. The spring constants of the three springs are assumed to be k_1, k_2 and k_3 with their unstretched lengths l_1, l_2 and l_3 respectively. The two masses, m_1 and m_2 are connected in between, as shown in the figure. Two extreme ends of the springs are fixed in rigid walls on two sides.

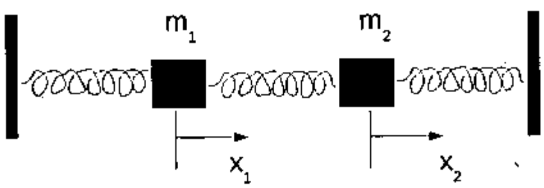


Figure 8.1: Coupled oscillator

The objects are allowed to move on a frictionless horizontal surface in a straight line along the layout of the springs. If the extension in the leftmost spring is x_1 and the middle spring is x_2, then the kinetic energy of the system can be written as

$$T = \frac{1}{2}(m_1\dot{x}_1^2 + m_2\dot{x}_2^2)$$

where the dots represent the derivatives with respect to time. Here the extensions are considered as the generalised coordinates and their dot derivatives the generalised velocities.

The potential energy of the system is written as

$$V = \frac{1}{2}k_1x_1^2 + \frac{1}{2}k_2x_2^2 + \frac{1}{2}k_3(x_2 - x_1)^2$$

so that the Lagrangian of the system is given by

$$L = T - V = \frac{1}{2}(m_1\dot{x}_1^2 + m_2\dot{x}_2^2) - \frac{1}{2}\left[k_1x_1^2 + k_2x_2^2 + k_3(x_2 - x_1)^2\right]$$

We now calculate the partial derivatives

$$\frac{\partial L}{\partial \dot{x}_1} = m_1\dot{x}_1, \qquad \frac{\partial L}{\partial x_1} = -k_1x_1 + k_3(x_2 - x_1), \qquad \frac{\partial L}{\partial \dot{x}_2} = m_2\dot{x}_2; \qquad \frac{\partial L}{\partial x_2} = -k_2x_2 + k_3(x_2 - x_1)$$

The Lagrange's equations for the system will then follow as

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_1}\right) - \frac{\partial L}{\partial x_1} = 0, \qquad \Longrightarrow \qquad m_1\ddot{x}_1 + k_1x_1 - k_3(x_2 - x_1) = 0 \tag{8.2.1a}$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_2}\right) - \frac{\partial L}{\partial x_2} = 0, \qquad \Longrightarrow \qquad m_2\ddot{x}_2 + k_2x_2 + k_3(x_2 - x_1) = 0 \tag{8.2.1b}$$

Equations (8.2.1a) and (8.2.1b) form a set of simultaneous linear differential equations with constant coefficients and therefore, we can proceed to solve them through trial solutions

$$x_1 = A\exp\lambda t, \qquad x_2 = B\exp\lambda t$$

with A and B as constants. Substituting these into (8.2.1a) and (8.2.1b) we obtain two homogeneous simultaneous equations for the determination of the constants A and B:

$$(\lambda^2 m_1 + k_1 + k_3)A - k_3B = 0 \tag{8.2.2a}$$

$$-k_3A + (\lambda^2 m_2 + k_2 + k_3)B = 0 \tag{8.2.2b}$$

This set of equations will have a solution if and only if the determinant formed by the coefficients of A and B vanishes, *i.e.*,

$$\begin{vmatrix} \lambda^2 m_1 + k_1 + k_3 & -k_3 \\ -k_3 & \lambda^2 m_2 + k_2 + k_3 \end{vmatrix} = 0 \tag{8.2.3}$$

At this point we may simplify the problem by considering the masses to be equal, *i.e.*, $m_1 = m_2 = m$ (say) and the spring constants of the three springs are also identical so that $k_1 = k_2 = k_3 = k$ (say). So, the determinant, with further substitution of $\frac{k}{m} = \omega_0^2$ is simplified to

$$\begin{vmatrix} \lambda^2 + 2\omega_0^2 & -\omega_0^2 \\ -\omega_0^2 & \lambda^2 + 2\omega_0^2 \end{vmatrix} = 0 \tag{8.2.4}$$

We expand the determinant to solve for λ and get the roots as

$$\lambda = \pm i\omega_0, \qquad \pm i\sqrt{3}\omega_0$$

As the differential equations were linear, the superposition principle holds, so that the sum of the four solutions, will also be solution and will contain four arbitrary constants:

$$x_1 = a_1 e^{i\omega_0 t} + a_2 e^{-i\omega_0 t} + a_3 e^{i\sqrt{3}\omega_0 t} + a_4 e^{-i\sqrt{3}\omega_0 t} \tag{8.2.5}$$

$$x_2 = b_1 e^{i\omega_0 t} + b_2 e^{-i\omega_0 t} + b_3 e^{i\sqrt{3}\omega_0 t} + b_4 e^{-i\sqrt{3}\omega_0 t} \tag{8.2.6}$$

where a's and b's are each four arbitrary constants consistent with the conditions (8.2.2a) and (8.2.2b) above. These conditions in fact reduce the number of arbitrary constants to be involved in the solutions to four corresponding to two second order differential equations (8.2.1a) and (8.2.1b).

It is obvious that these solutions are represented in complex number form. Trigonometric representations of the solutions are also in fact possible. For this, let

$$a_1 = \frac{A_1}{2} e^{i\phi_1}, \qquad a_2 = \frac{A_1}{2} e^{-i\phi_1}$$

With this, the first two terms of the solution (8.2.5) can be rewritten as

$$\begin{aligned} x_1(12) &= \frac{A_1}{2} e^{i\omega_0 t} e^{i\phi_1} + \frac{A_1}{2} e^{-i\omega_0 t} e^{-i\phi_1} \\ &= \frac{A_1}{2} e^{i(\omega_0 t + \phi_1)} + \frac{A_1}{2} e^{-i(\omega_0 t + \phi_1)} \\ &= A_1 \cos(\omega_0 t + \phi_1) \end{aligned}$$

Similarly the last two terms of the solution (8.2.5) can be written as

$$x_1(34) = A_2 \cos(\sqrt{3}\omega_0 t + \phi_2),$$

and following the same procedure the first two terms of the second solution (8.2.6) reduce to

$$x_2(12) = B_1 \cos(\omega_0 t + \phi_1),$$

and the last two terms to

$$x_2(34) = B_2 \cos(\sqrt{3}\omega_0 t + \phi_2).$$

Combining all, we can have the trigonometric representation of the solutions, with four arbitrary constants as

$$\begin{aligned} x_1 &= A_1 \cos(\omega_0 t + \phi_1) + A_2 \cos(\sqrt{3}\omega_0 t + \phi_2) \\ x_2 &= B_1 \cos(\omega_0 t + \phi_1) + B_2 \cos(\sqrt{3}\omega_0 t + \phi_2) \end{aligned}$$

The new arbitrary constants $A_1, A_2, B_1, B_2, \phi_1$ and ϕ_2 are related to each other through (8.2.2a) and (8.2.2b):

$$B_1 = \frac{2\omega_0^2 + \lambda_1^2}{\omega_0^2} A_1, \qquad \text{and} \qquad B_2 = \frac{2\omega_0^2 + \lambda_2^2}{\omega_0^2} A_2,$$

where $\lambda_1 = \pm\omega_0$ and $\lambda_2 = \pm\sqrt{3}\omega_0$. The two conditions above relates the constants as $B_1 = A_1$ and $B_2 = -A_1$. Thus the general solution assumes the form

$$\left.\begin{aligned} x_1 &= A_1\cos(\omega_0 t + \phi_1) + A_2\cos(\sqrt{3}\omega_0 t + \phi_2) \\ x_2 &= A_1\cos(\omega_0 t + \phi_1) - A_2\cos(\sqrt{3}\omega_0 t + \phi_2) \end{aligned}\right\} \tag{8.2.7}$$

Equations (8.2.7) reveal that the motion of each coordinate is a superposition of two hamonic vibrations of frequencies ω_0 and $\sqrt{3}\omega_0$. The frequencies of oscillation are the same for both x_1 and x_2 ; only their relative amplitudes are different.

Now if any one of A_1 and A_2 is zero, we will have only one frequency of vibration in the system. If $A_1 = 0$ we have

$$\begin{aligned} x_1 &= A_2\cos(\sqrt{3}\omega_0 t + \phi_2) \\ x_2 &= -A_2\cos(\sqrt{3}\omega_0 t + \phi_2) \end{aligned}$$

and the two vibrations are of opposite phase, *i.e.*, the two bodies will move opposite to each other. On the other hand, if $A_2 = 0$, then

$$\begin{aligned} x_1 &= A_1\cos(\omega_0 t + \phi_1) \\ x_2 &= A_1\cos(\omega_0 t + \phi_1) \end{aligned}$$

the two vibrations will be in phase and the bodies will move in the same direction.

Thus we see that there are two modes of motion in the considered mass-spring system involving a single frequency. These modes of vibrations are known as the *normal modes of vibration*. It is also possible to find a system of coordinates in which oscillation of a single coordinate contains one of the frequencies of vibration, irrespective of the initial phase or conditions.

It is also noteworthy that an interchange of the two oscillating bodies does not affect the equations of motion. Therefore, we can look for a coordinate system $(q_1,\ q2)$ where q_1 is symmetric with respect to an interchange of the masses and q_2 is antisymmetric in this interchange :

$$q_1 = \frac{1}{2}(x_1 + x_2) \qquad \text{and} \qquad q_2 = \frac{1}{2}(x_1 - x_2)$$

$$\text{or,} \qquad x_1 = q_1 + q_2 \qquad \text{and} \qquad x_2 = (q_1 - q_2)$$

Substituting these to the Lagrange's equations of motion (8.2.1a), (8.2.1b) yields

$$m(\ddot{q}_1 + \ddot{q}_2) + m\omega_0^2(q_1 + q_2) - m\omega_0^2(-2q_2) = 0$$

$$\text{and} \qquad m(\ddot{q}_1 - \ddot{q}_2) + m\omega_0^2(q_1 - q_2) + m\omega_0^2(-2q_2) = 0$$

Addition and subtraction gives

$$\left.\begin{aligned} &\ddot{q}_1 + \omega_0^2 q_1 = 0 \\ \text{and} \qquad &\ddot{q}_2 + 3\omega_0^2 q_2 = 0 \end{aligned}\right\} \tag{8.2.8}$$

which are two independent and simpler equations. We may now contrast equations (8.2.8) with (8.2.7). We observe that in the set of solutions (8.2.8), q_1 involves only a single frequency ω_0 and q_2 only $\sqrt{3}\omega_0$, whereas in equations (8.2.7) both the coordinates x_1 and x_2 consist of both the frequencies ω_0 and $\sqrt{3}\omega_0$.

8.3 General theory of small oscillation

Having discussed the systems with degrees of freedom restricted to one or two, we are now going to discuss its extension to systems with any number say n degrees of freedom and analyse the oscillations in the system. Further, as the title suggests, the basic assumption that we adopt in the discussion to follow is the assumption of small displacements of the associated particles from the position of stable equilibrium and so also the amplitude of corresponding oscillation amplitudes. Alongside, we will consider that the system for study is a conservative system so that the potential energy associated with the system is dependent only on n generalised co-ordinates $q_1, q_2, \ldots, q_n$, represented as

$$V = (q_1, q_2, \ldots, q_n)$$

Let us denote the equilibrium positions by q_i^0 so that the displacements of individual particles from equilibrium position by u_i, *i.e.*,

$$q_i = q_i^0 + u_i$$

which are the generalised coordinates of the system. Expanding the potential energy about the position of equilibrium, we obtain

$$V(q_1, q_2, \ldots, q_n) = V(q_1^0, q_2^0, \ldots, q_n^0) + \sum_{i=1}^{n} \left[\frac{\partial V}{\partial q_i}\right]_0 (q_i - q_i^0) + \frac{1}{2!} \sum_{i=1}^{n} \sum_{j=1}^{n} \left[\frac{\partial^2 V}{\partial q_i q_j}\right]_0 (q_i - q_i^0)(q_j - q_j^0) + \ldots$$

Here $(\frac{\partial V}{\partial q_i})_0 = 0$ because this represents the force applied to the system under equilibrium which is zero. Also the first term $V(q_1^0, q_2^0, \ldots, q_n^0)$ in the expansion represents the potential energy in the equilibrium position and is constant for the system. As we know the potential is associated with an additive constant, we can always adjust this constant so that the equilibrium potential energy function is zero without affecting dynamics of the system. This adjustment leads the potential energy in equilibrium to zero as said, and writing $u_i = q_i - q_i^0$ etc., we get the leading order potential energy of the displaced configuration as

$$V = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} V_{ij} u_i u_j$$

where $V_{ij} = \left[\frac{\partial^2 V}{\partial q_i q_j}\right]_0 = \left[\frac{\partial^2 V}{\partial u_i u_j}\right]_0$ =constant which is to be evaluated at $q_i = q_i^0$ and $q_j = q_j^0$. The constant $V_{ij} = V_{ji}$ form a symmetric $n \times n$ matrix V,

$$V = \begin{pmatrix} V_{11} & V_{12} & \ldots & V_{1n} \\ V_{21} & V_{22} & \ldots & V_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ V_{n1} & V_{n2} & \ldots & V_{nn} \end{pmatrix}$$

The kinetic energy of the system associated with the displacements is given by

$$T = \sum_i \sum_j \frac{1}{2} m_{ij}.\dot{q}_i.\dot{q}_j = \sum_i \sum_j m_{ij}.\dot{u}_i \dot{u}_j \tag{8.3.1}$$

The co-efficients m_{ij} are in general, functions of generalised co-ordinates. Expanding m_{ij} around the equilibrium position, in Taylor's series, we get

$$m_{ij}(q_1, \ldots, q_n) = m_{ij}^0(q_0^1, \ldots, q_0^n) + \sum_{k=1}^{n} \left[\frac{\partial m_{ij}}{\partial q_k}\right]_0 u_k + \ldots\ldots \tag{8.3.2}$$

In order to keep parity with the quadratic form of potential energy function V, the kinetic energy function T is also to be considered in the quadratic form. But the expression for T in Equation (8.3.1) is already quadratic in $\dot{q}_i$'s, therefore it is sufficient to consider only the first term in the equation (8.3.2).

The kinetic energy is given by

$$T = \frac{1}{2} \sum_i \sum_j m_{ij}^0 \dot{u}_i \dot{u}_j = \frac{1}{2} \sum_i \sum_j T_{ij}.v_i.v_j$$

where the constant m_{ij}^0's are denoted by T_{ij}s. One can easily discover that the constants T_{ij} are in fact the elements of the $n \times n$ symmetric matrix T.

Note !

At this point it is recalled that a general expression for the kinetic energy of a system of N particles with masses m_i, having n degrees of freedom and with generalised coordinates $(q_1, q_2, \ldots, q_n)$ can be derived as consisting of three terms: one independent of the generalised coordinates one involving generalised velocity and one term associated with the square of the generalised velocity as

$$T = \frac{1}{2}\sum_{j,k} m_{jk}\dot{q}_j\dot{q}_k + \sum_j a_j\dot{q}_j + b$$

where

$$\left.\begin{aligned} m_{jk} &= \sum_i m_i \frac{\partial \vec{r}_i}{\partial q_j}.\frac{\partial \vec{r}_i}{\partial q_k} \\ a_j &= \sum_i m_i \frac{\partial \vec{r}_i}{\partial q_j}.\frac{\partial \vec{r}_i}{\partial t} \\ b &= \frac{1}{2}\sum_i m_i \left(\frac{\partial \vec{r}_i}{\partial t}\right)^2 \end{aligned}\right\}$$

Here $\vec{r}_i \equiv \vec{r}_i(q_1, q_2, \ldots, q_n, t)$ is the position vector of the i-th particle of the system. If the position vector is independent of time explicitly, the terms a_j and b above will vanish and the kinetic energy is a homogeneous quadratic function of the generalised velocities

$$T = \frac{1}{2}\sum_{j,k} m_{jk}\dot{q}_j\dot{q}_k$$

The Lagrangian of the sysem can be written as

$$L = T - V = \frac{1}{2}\sum_i\sum_j [T_{ij}.\dot{u}_i\dot{u}_j - V_{ij}u_iu_j]$$

Using u's as the generalized co-ordinates, the Lagrange's equations

$$\frac{d}{dt}\left[\frac{\partial L}{\partial \dot{q}_i}\right] - \frac{\partial L}{\partial q_i} = 0$$

will take the form

$$\sum_{j=1} [T_{ij}\ddot{u}_j + V_{ij}u_j] = 0 \qquad (8.3.3)$$

for $i = 1, 2, \cdots, n$. The expression (8.3.3) actually represents a set of n second order differential equations which are to be solved to obtain the motion near the equilibrium position.

8.3.1 Secular equation and eigen value equation

Equations (8.3.3) are second order differential equations with constant coefficients. So, it is worthwhile to proceed with a trial solution of the form

$$u_i = a_i e^{i\omega t} \tag{8.3.4}$$

Where a_i are the complex amplitudes of the oscillations for each the co-ordinate u_i. Substituting (8.3.4) into equation (8.3.3) we get the auxillary equations

$$\sum_{j=1}^{n} \left(V_{ij} - \omega^2 T_{ij}\right) a_i = 0 \tag{8.3.5}$$

Or, in a matrix form

$$\mathbf{V}.\mathbf{a} - \omega^2 \mathbf{T}.\mathbf{a} = \mathbf{0} \tag{8.3.6}$$

where $\mathbf{V}$, $\mathbf{T}$ are $n \times n$ matrices and $\mathbf{a}$ forms the $n \times 1$ column vector.

(8.3.7)

Equation (8.3.5) represents a linear, homogeneous, algebraic equation is a_j and ω, *i.e.*,

$$\left.\begin{array}{l}\left(V_{11} - \omega^2 T_{11}\right) a_1 + \left(V_{12} - \omega^2 T_{12}\right) a_2 + \cdots\cdots + \left(V_{1n} - \omega^2 T_{1n}\right) a_n = 0 \\ \quad\vdots \qquad\qquad\qquad \vdots \qquad\qquad\qquad \vdots \qquad\qquad \vdots \\ \left(V_{n1} - \omega^2 T_{n1}\right) a_1 + \left(V_{n2} - \omega^2 T_{n2}\right) a_2 + \cdots\cdots + \left(V_{nn} - \omega^2 T_{nn}\right) a_n = 0\end{array}\right\} \tag{8.3.8}$$

Assuming that T^{-1} exists, we multiply the equation (8.3.6) by $\mathbf{T}^{-1}$

$$\begin{aligned}\mathbf{T}^{-1}\mathbf{V}\mathbf{a} - \omega^2\mathbf{a} &= 0 \\ \mathbf{P}\mathbf{a} - \omega^2\mathbf{I}\mathbf{a} &= 0, \qquad \text{with} \quad \mathbf{P} = \mathbf{T}^{-1}\mathbf{V} \\ (\mathbf{P} - \omega^2\mathbf{I})\mathbf{a} &= 0\end{aligned} \tag{8.3.9}$$

Equation (8.3.9) is the well known *eigenvalue equation* or the *characteristic equation* or the *secular equation* of the system. It is an equation of degree n in ω^2 and hence there exists n roots in ω^2.
For the non-trivial solution to equations (8.3.8) to exist, the determinant of the coefficients of a_i's in (8.3.8) or (8.3.9) must be zero, *i.e.*,

$$\left|(\mathbf{P} - \omega^2\mathbf{I})\right| = 0$$

$$\text{or,} \qquad \left|(\mathbf{V} - \omega^2\mathbf{T})\right| = 0$$

which gives the eigenvalues of the corresponding to the eigen equations. These eigen values are the frequencies ω_r, also called the *eigenfrequencies* or the *normal modes* of the system.

The matrices $\mathbf{V}$ and $\mathbf{T}$ being symmetric, the eigen values ω^2 are all real.
Again to every eigenfrequency ω_r there must exist an eigen vector $\mathbf{a}_r$ of the system which does not change in direction but in magnitude only and consequently the mode consists of the simultaneous

oscillations of several degrees of freedom. In case two or more frequencies are equal, the system is said to be degenerate (*i.e.*, there is an arbitrary choice of normal modes which corresponds to same eigen frequencies.

Now, if we use the symbol a_{jr} to represent the j-th component of the r-th eigenvector corresponding to eigenfrequency ω_r, then by using the principle of superposition, the general solution for the displacements q_j can be written as a function of time as,

$$q_j(t) = \sum_r a_{jr} \exp\left[\mathrm{i}(\omega_r t - \delta_r)\right]$$

or, $$q_j(t) = \sum_r \cos(\omega_r t - \delta_r) \tag{8.3.10}$$

8.3.2 Small Oscillations of Particles on String

We consider a system of n particles each of equal mass m connected linearly with equal spacings l by small, light and identical springs each with spring constant k. Naturally the system will form a string, the end points of which are attached to fixed points. Let the string be stretched in the transverse direction with force F, rendering the different masses to be possess different transverse displacements.

Let q_i be the transverse displacement of the i-th particle from the equilibrium position. The system will then possess a kinetic energy given by

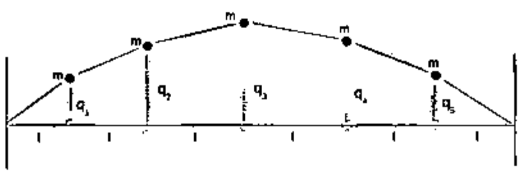


Figure 8.2: String of particles displaced from their equilibrium positions.

$$T = \frac{1}{2} m \sum_j \dot{q}_j^2$$

We can see that the total length of the string in equilibrium position is $L = (n+1)l$. It requires to find out the displacement of the j-th particle, in order to calculate the potential energy associated with it. We can figure out that there is a change in the length between the j-th and $(j+1)$-th particle from its equilibrium length l. We can write the expression for the new displaced length between the pair by

$$l + \delta l = \sqrt{l^2 + (q_{j+1} - q_j)^2}$$
$$\simeq l \left\{ 1 + \frac{(q_{j+1} - q_j)^2}{2l^2} \right\}$$

where the change of the length is denoted by δl. Here we have neglected the higher order terms in the expansion of the quantity under the radical sign.

Thus, the work done associated with the displacement $\delta l = (q_{j+1} - q_j)^2/2l$, is $F.\delta l$. Here of course, we have assumed that the tension in the string in equilibrium position and in a stretched position are equal. The work done in displacing the mass points from their equilibrium position is stored in the system as the potential energy V and so we write the expression for V as

$$\begin{aligned} V &= \frac{F}{2l}\left[q_1^2 + (q_2 - q_1)^2 + (q_3 - q_2)^2 + \cdots + (q_{n-1} - q_n^2) + q_n^2\right] \\ &= \frac{F}{2l}\sum_{j=1}^{n+1}(q_{j-1} - q_j)^2 \end{aligned}$$

Thus, the Lagrangian of the system is

$$L = \sum_{j=1}^{n+1}\left[\frac{1}{2}m\dot{q}_j^2 - \frac{F}{2l}(q_{j-1} - q_j)^2\right]$$

Using Lagrange's equations for q_j, we get

$$m\ddot{q}_j - \frac{F}{l}(q_{j-1} - 2q_j + q_{j+1}) = 0$$

$$\text{or,} \qquad \ddot{q}_j = \omega_0^2(q_{j-1} - 2q_j + q_{j+1})$$

Here we have replaced $\frac{F}{ml} = \omega_0^2$. The equation of motion for the particle here reveals that it depends on displacements q_{j-1} and q_{j+1} of its neighboring particles only.

As the obtained differential equation is of 2nd order with constant coefficients, the solution can be obtained thruogh a trial solution of the type $q_j = a_j \exp(i\omega t)$, we get the set of following equations:

$$\begin{aligned} (2\omega_0^2 - \omega^2)a_1 - \omega_0^2 a_2 &= 0 \\ -\omega_0^2 + (2\omega_0^2 - \omega^2)a_2 - \omega_0^2 a_3 &= 0 \\ \vdots \qquad \vdots \qquad \vdots \qquad &\vdots \\ -\omega_0^2 a_{n-1} + (2\omega_0^2 - \omega^2)a_n &= 0 \end{aligned}$$

We will look for the solution of the secular equation for small values of n. For $n = 1$, there is only one particle between the fixed points and hence only the first term of the first equation above is relevant, *i.e.*, $2\omega_0^2 - \omega^2 = 0$ and we have only one normal mode with frequency $\omega^2 = 2\omega_0^2$. For $n = 2$, the characteristic equation is the determinant of

$$\begin{pmatrix} 2\omega_0^2 - \omega^2 & -\omega_0^2 \\ -\omega_0^2 & 2\omega_0^2 - \omega^2 \end{pmatrix} = (2\omega_0^2 - \omega^2)^2 - \omega_0^4 = 0$$

and accordingly we have two normal mode solutions :

$$\begin{aligned} &\omega_1^2 = \omega_0^2 \qquad \text{with} \quad a_1 = a_2 \\ \text{and} \qquad &\omega_2^2 = 3\omega_0^2 \qquad \text{with} \quad a_1 = -a_2 \end{aligned}$$

For $n = 3$, three particles are involved and the secular equation is the given by the determinant

$$\begin{vmatrix} 2\omega_0^2 - \omega^2 & \omega_0^{-2} & 0 \\ -\omega_0^2 & 2\omega_0^2 - \omega^2 & -\omega_0^2 \\ 0 & -\omega_0^2 & 2\omega_0^2 - \omega^2 \end{vmatrix} = (2\omega_0^2 - \omega^2)^3 - 2\omega_0^4(2\omega_0^2 - \omega^2) = 0$$

which has the roots $2\omega_0^2$ and $(2 \pm \sqrt{2})\omega_0^2$. The corresponding normal modes are

$$\begin{aligned} \omega_1^2 = (2 - \sqrt{2})\omega_0^2, &\qquad \frac{a_1}{a_2} = \frac{1}{\sqrt{2}} = \frac{a_3}{a_2} \\ \omega_2^2 = 2\omega_0^2, &\qquad \frac{a_1}{a_3} = -1, \qquad a_2 = 0 \\ \omega_3^2 = (2 + \sqrt{2})\omega_0^2, &\qquad \frac{a_1}{a_2} = -\frac{1}{\sqrt{2}} = \frac{a_3}{a_2} \end{aligned}$$

Proceeding in the similar manner it is possible to find out the modes for $n = 4, 5, \ldots$ and so on. It is worthwhile to note here that for any value of n, the slowest mode is the one in which all the particles are oscillating in the same direction and the fastest mode has the adjacent particles oscillating in opposite directions. In the limit of a very high value of n, the separation between the particles will be infinitesimal and the corresponding normal modes approach those of a continuous stretched string.

Example 8.3.1 *Find the normal frequencies and normal coordinates of the system whose Lagrangian is given by*

$$L = \frac{1}{2}\left(x'^2 + y'^2\right) - \frac{1}{2}\left(\omega_1^2 x^2 + \omega_2^2 y^2\right) + \alpha xy.$$

Solution: Here, the lagrangian is given. The Lagrange's equation for both the coordinates are

$$x'' + \omega_1^2 x - \alpha y = 0 \tag{8.3.11}$$

$$y'' + \omega_2^2 y - \alpha x = 0 \tag{8.3.12}$$

Now, we seek the solutions in which the system is vibrating in simple harmonic motion at normal frequencies ω, *i.e.*, we seek solutions of the form

$$x'' = -\omega^2 x \quad (8.3.13)$$
$$and \quad y'' = -\omega^2 y \quad (8.3.14)$$

Using these equations (8.3.11) and (8.3.12) reduce to

$$x(\omega_1^2 - \omega^2) = \alpha y, \qquad and \qquad y(\omega_2^2 - \omega^2) = \alpha x$$

$$i.e., \quad \frac{x}{y} = \frac{\alpha}{(\omega_1^2 - \omega^2)}, \qquad and \qquad \frac{x}{y} = \frac{(\omega_2^2 - \omega^2)}{\alpha}$$

These are equal and by equating the right hand sides,we obtain the following equation for the normal frequencies

$$\omega^4 - \omega^2(\omega_1^2 + \omega_2^2) + (\omega_1^2\omega_2^2 - \alpha^2) = 0$$

This gives

$$\omega^2 = \frac{(\omega_1^2 + \omega_2^2) \pm \sqrt{(\omega_1^2 - \omega_2^2)^2 + 4\alpha^2}}{2} \quad (8.3.15)$$

If $\alpha = 0$ then the normal frequencies are ω_1 and ω_2.

The normal coordinates x and y are obtained by solving equation (8.3.13) and (8.3.14). They are

$$x = A_1 \cos(\omega t + \epsilon_1)$$
$$and \quad y = A_2 \cos(\omega t + \epsilon_2),$$

where $x = A_1$, A_2, ϵ_1, ϵ_2 are arbitrary constants. The values of ω are given by (8.3.15).

8.4 Summary

In this section we have discussed a relative complex problem of oscillations induced in the particles of of a system. As the mathematical treatment and the corresponding solution will be intractable for any arbitrary oscillations in the first place, the problem is simplified by the assumption that the magnitude of the oscillations are small compared to the characteristic length considerations in the system. The corresponding kinetic energy, the potential energy are found out to write down the Lagrangian of the system. The corresponding Euler equations of motion for the particles will give the differential equations governing the motion of each particle. The normal mode solutions are then found out to find various modes of oscillations that are at work in such a system.

Self Study Questions:

1. Determine the dynamics of the forced small oscillations of a system where the force $F(t)$ associated with the system at the initial moment of time is given by

   (a) $F = F_0 =$ constant,

   (b) $F = \alpha t$, where α is a constant,

(c) $F = F_0 e^{-\mu t}$, where μ is a constant,

(d) $F = F_0 e^{-\mu t} \cos \nu t$, where ν is a constant.

2. What is a coupled oscillation? Find the equations motion in the coupled pendulum.

3. Obtain the normal modes of vibration for the double pendulum, where the bob of the first pendulum is the point of suspension of the second pendulum.

4. The Lagrangian L of a system with two degrees of freedom, is given by

$$L = \frac{1}{2}\left[(\dot{x}^2 + \dot{y}^2) - \omega_0^2(x^2 + y^2)\right] + \alpha xy,$$

where ω_0 is the natural frequency of the system and α is a constant. Determine the normal coordinates and eigenfrequencies of the system.

UNIT 9

Canonical Transformation

Preparatory inputs to this unit

1. Hamilton's principle and Canonical equations of motion.
2. Basics of Ordinary and Partial differential equations.
3. Elementary ideas of coordinate transformations.

9.1 Canonical Transformation

9.1.1 Introduction

Canonical transformations are considered the heart of the philosophy behind the formalism of the Hamiltonian mechanics which allows us to regard the coordinates and momenta in the equal footing. So far we have considered q's as the generalised coordinates and p's as the generalised momenta and use the Hamilton's Canonical equations to solve a dynamical problem. Now suppose we raise a question here. While solving a given dynamical problem, can we think of q's and p's to interchange their roles, *i.e.*, can we consider the quantities q's as the generalised momenta and p's as the generalised coordinates without affecting the outcome of the solution and the results are still identical in both the situations?

Lagrangian dynamics does not have direct role of the generalised momenta in the equations. Generalised momenta are the outcomes of the equations. Hence the Lagrangian dynamics does not seem to answer the outcomes of the exchanging of the roles in the generalised coordinates and momenta. In the Hamiltonian formalism, for a set of the conjugate pairs of the generalised coordinates and generalised momenta $(q_i,\ p_i)$ the canonical equations are

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \qquad\qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}$$

If we consider a transformation that the old coordinates q transform to a new momentum P and the old momenta p transform to new coordinates Q so that

$$q \to -P, \qquad \text{and} \qquad p \to Q$$

then the Hamilton's equations in the new situation will be given by

$$\dot{Q}_i = \frac{\partial H}{\partial P_i} \qquad\qquad \dot{P}_i = -\frac{\partial H}{\partial Q_i}$$

i.e., the new set of conjugate pairs $(Q_i,\ P_i)$ the equations are indistinuguishable from those of the old pairs, which would give the same result even we consider the first of the conjugate pair as the generalised coordinates whereas it was the momenta in the old system.

Anther requirement in view of the solution of dynamical problems is that if a coordinate is cyclic in the Hamiltonian or the Lagrangiaqn, the corresponding conjugate momentum is a constant of motion and this leads to an easy solution in relation to the cyclic coordinate. If there are more cyclic coordinates in the Hamiltonian, the solution becomes progressively easier and in the eventual case, if all the coordinates are cyclic in the Hamiltonian, it will be very easy to solve the problem. This calls for finding some transformations which will render the Hamiltonian to have all the coordinates to be cyclic. The condition to be imposed on such transformation is that the the dynamics should not be altered by the transformation so that the form of the Hamilton's canonical equations remains the same.

In view of the above requirements, a special kind of transformations in the generalised coordinates and generalised momenta can be defined in such a way that the form of the Hamilton's canonical equations retains the form, although the New Hamiltonian might have a different dependence on

the new generalised coordinates and generalised momenta. Such a transformation is called Canonical Transformation.
There are many equivalent definitions of canonical transformations of variables, but they all boil down to the single essence that a set of variables $(P,\ Q)$ are canonical transforms of the variables (p, q) if they preserve the structure of the equations of motion.

It is imperative to look for some generating functions to obtain canonical transformations such that operating these generating functions yields the new set of canonical variables in terms of the old set. It must of course be remembered that no new physical informations are derivable from any choice of coordinate system. The canonical variables and transformations are purely lead by a convenience, although a very useful, in describing dynamical systems.

Formal definition of Canonical Transformation: When the set of canonically conjugate pairs of generalised coordinates and generalised momenta $(q_i,\ p_i)$ corresponding to a Hamiltonian $H(q_i,\ p_i, t)$ and satisfying the Hamilton's canonical equations

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}$$

are transformed to a new set $(Q_i,\ P_i)$ such that

$$\begin{aligned} Q_i &= Q_i(q_1, q_2, \dots, q_n;\ p_1, p_2, \dots, p_n;\ t) \\ P_i &= P_i(q_1, q_2, \dots, q_n;\ p_1, p_2, \dots, p_n;\ t) \end{aligned}$$

or, in short,

$$Q_i = Q_i(q_j, p_j, t) \qquad P_i = P_i(q_j, p_j, t)$$

and the new hamiltonian $K(Q_i,\ P_i, t)$ satisfies

$$\dot{Q}_i = \frac{\partial K}{\partial P_i}, \qquad \dot{P}_i = -\frac{\partial K}{\partial Q_i}$$

the transformations are known as *Canonical Transformations* or *Contact Transformations.*

The Hamiltonian functions H and K in the two sets of coordinates are related by

$$H = \sum_k p_k \dot{q}_k - L(q_i, \dot{q}_i, t) \qquad \text{and} \qquad K = \sum_k P_k \dot{Q}_k - \mathcal{L}(Q, \dot{Q}, t)$$

where $\mathcal{L}$ is the transformed Lagrangian.

Requirements for Canonical transformations: Generating function

If Q_i and P_i are canonical coordinates, they must satisfy a modified Hamilton's principle of the form

$$\delta \int_{t_1}^{t_2} \left(\sum_i P_i \dot{Q}_i - K(Q, P, t) \right) dt = 0 \tag{9.1.1}$$

Similarly, the old canonical coordinates too, satisfy

$$\delta \int_{t_1}^{t_2} \left(\sum_i p_i \dot{q}_i - H(q, p, t) \right) dt = 0 \tag{9.1.2}$$

These two equations are simultaneously valid only when the two integrands differ at best by a total time derivative of an arbitrary function F, *i.e.*,

$$\delta \int_{t_1}^{t_2} (L - \pounds) dt = 0 \tag{9.1.3}$$

with $L = \sum_i p_i \dot{q}_i - H(q, p, t)$ and $\pounds = \sum_i P_i \dot{Q}_i - K(Q, P, t)$, lead to

$$L - \pounds = \frac{dF}{dt}$$

This is because of the general form of the modified Hamilton's principle which has zero variation at the end points. Both the statements will be satisfied only if the integrands are related by

$$\lambda \left(\sum_i p_i \dot{q}_i - H(q, p, t) \right) = \left(\sum_i P_i \dot{Q}_i - K(Q; P, t) \right) + \frac{dF}{dt} \tag{9.1.4}$$

Here λ is a multiplicative constant independent of canonical coordinates and related to a simple type of transformation of canonical coordinates, known as a scale transformation. A transformation of canonical coordinates for which $\lambda \neq 1$ will be called as an extended canonical transformation. If the transformation equations do not contain the time explicitly then they will be called restricted canonical transformations.

The action integral ((9.1.1), (9.1.2) or (9.1.3)) will vanish for any form of the function; irrespective of whether F is a function of (q, p, t) only, or (Q, P, t) only, or any compbination of the new and old phase space coordinates since all these will have zero variations at the end points.

We can use the transformation equations

$$Q_i = Q_i(q, p, t)$$
$$P_i = P_i(q, p, t)$$

And their inverses

$$q_i = q_i(Q, P, t)$$
$$p_i = p_i(Q, P, t)$$

to express the function F in terms of a part of old set of variables and partly of the new. In fact the function F can be considered as a bridge between the two sets of canonical variables and is called the generating function of transformation.

9.1.2 Generating functions for Canonical Transformation

Depending on the choice of old and new variables we can consider four basic types of generating functions. The generating function will serve to generate a transformation from old set to new set of canonical variables so that the generated transformation $(q, p) \to (Q, P)$ is guaranteed to be a canonical transformation.

Type 1 generating function

The type 1 generating function, denoted by G_1, involves only the generalized coordinates-both old and new, and of course the the time, *i.e.*, $G = G_1(q_i, Q_i, t)$.

To deduce the relations of the transformations, with the generating function we write the total time derivative of the generating function as

$$\frac{dG_1}{dt} = \sum_i \left(\frac{\partial G_1}{\partial q_i}\dot{q}_i + \frac{\partial G_1}{\partial Q_i}\dot{Q}_i \right) + \frac{\partial G_1}{\partial t}$$

So, the requirement of canonical transformation

$$\sum_i p_i\dot{q}_i - H = \sum_i P_i\dot{Q}_i - K + \frac{dG_1}{dt}$$

or
$$\sum_i p_i\dot{q}_i - H = \sum_i P_i\dot{Q}_i - K + \sum_i \left(\frac{\partial G_1}{\partial q_i}\dot{q}_i + \frac{\partial G_1}{\partial Q_i}\dot{Q}_i \right) + \frac{\partial G_1}{\partial t}$$

or
$$\sum_i \left(p_i - \frac{\partial G_1}{\partial q_i} \right) \dot{q}_i - \sum_i \left(P_i + \frac{\partial G_1}{\partial Q_i} \right) \dot{Q}_i + K - H - \frac{\partial G_1}{\partial t} = 0$$

Collecting the coefficients of $\dot{q}_i$ and $\dot{Q}_i$ is possible for both being independent of each other, *i.e.*,

$$p_i = \frac{\partial G_1}{\partial q_i} \qquad P_i = -\frac{\partial G_1}{\partial Q_i} \qquad K = H + \frac{\partial G_1}{\partial t}$$

Thus it is seen that once the generating function $G(q, Q, t)$ as a function of the old and new generalised coordinates, its partial derivative with respect to the old coordinates yields the old momenta and the negative of the partial derivative of G with respect to the new coordinate results in the new momenta.

Type 2 generating function

The type 2 generating function G_2 depends on the old generalised coordinates and the new generalized momenta $G_2 \equiv G_2(q_i, P_i, t)$.

In fact there is a relation between the type 1 and type 2 generating functions. Legendre dual transformation is a special type of transformation which links between these two types. Without going through the details of the Legendre transformation we would write down the relation which goes as

$$G_2(q_i, P_i, t) = \sum_i P_i Q_i + G_1(q_i, Q_i, t)$$

The Canonicality requirement

$$\sum_i p_i\dot{q}_i - H = \sum_i P_i\dot{Q}_i - K + \frac{dG_1}{dt}$$
$$= \sum_i P_i\dot{Q}_i - K + \frac{d}{dt}\left[G_2(q_i, P_i, t) - \sum_i P_iQ_i\right]$$
$$= \sum_i P_i\dot{Q}_i - K + \frac{dG_2}{dt} - \sum_i P_i\dot{Q}_i - \sum_i \dot{P}_iQ_i$$

or,
$$\frac{dG_2}{dt} = \sum_i p_i\dot{q}_i + \sum_i \dot{P}_iQ_i + K - H$$

Expanding the total derivative on the left hand side,

$$\sum_i \frac{\partial G_2}{\partial q_i}\dot{q}_i + \sum_i \frac{\partial G_2}{\partial P_i}\dot{P}_i + \frac{\partial G_2}{\partial t} = \sum_i p_i\dot{q}_i + \sum_i \dot{P}_iQ_i + K - H$$

Since the old coordinates and new momenta are each independent, the following equations must hold:

$$\boxed{p_i = \frac{\partial G_2}{\partial q_i}} \qquad \boxed{Q_i = \frac{\partial G_2}{\partial P_i}} \qquad \boxed{K = H + \frac{\partial G_2}{\partial t}}$$

Type 3 generating function

The type 3 generating function G_3 is a function of the old generalized momenta and the new generalized coordinates, *i.e.*, $G_3 \equiv G_3(p_i, Q_i, t)$ and it is connected to the type 1 generating function through the Legendre transformation as

$$G_3(p_i, Q_i, t) = G_1(q_i, Q_i, t) - \sum_i p_iq_i \tag{9.1.5}$$

We again meet the requirement of Canonicality as

$$\sum_i p_i\dot{q}_i - H = \sum_i P_i\dot{Q}_i - K + \frac{dG_1}{dt}$$

Substitution of the last term through (9.1.5) yields

$$= \sum_i P_i \dot{Q}_i - K + \frac{d}{dt}\left[G_3(p_i, Q_i, t) + \sum_i p_i q_i\right]$$

$$= \sum_i P_i \dot{Q}_i - K + \frac{dG_3}{dt} + \sum_i p_i \dot{q}_i + \sum_i \dot{p}_i q_i$$

or,
$$\frac{dG_3}{dt} = -\sum_i q_i \dot{p}_i - \sum_i P_i \dot{Q}_i + K - H$$

Expanding the total derivative of the generating function on the left hand side,

$$\sum_i \frac{\partial G_3}{\partial p_i}\dot{p}_i + \sum_i \frac{\partial G_3}{\partial Q_i}\dot{Q}_i + \frac{\partial G_3}{\partial t} = -\sum_i q_i \dot{p}_i - \sum_i P_i \dot{Q}_i + K - H$$

As before, the generalised quantities are independent of each other, we can equate the coefficients so that

$$q_i = -\frac{\partial G_3}{\partial p_i} \qquad P_i = -\frac{\partial G_3}{\partial Q_i} \qquad K = H + \frac{\partial G_3}{\partial t}$$

Type 4 generating function

The type 4 generating function G_4 involves both the old and new generalised momenta, so that $G_4 \equiv G_4(p_i, P_i, t)$, so that

$$\frac{dG_4}{dt} = \sum_i \frac{\partial G_4}{\partial p_i}\dot{p}_i + \sum_i \frac{\partial G_4}{\partial P_i}\dot{P}_i + \frac{\partial G_4}{\partial t}$$

Further, the function G_4 is also related to G_1 via Legendre Transformation in the manner as,

$$G_4(p_i, P_i, t) = \sum_i P_i Q_i - \sum_i p_i q_i + G_1(q_i, Q_i, t)$$

Using the above result, we can now write down the requirement of Canonicality as

$$\sum_i p_i \dot{q}_i - H = \sum_i P_i \dot{Q}_i - K + \frac{dG_1}{dt}$$

$$= \sum_i P_i \dot{Q}_i - K + \frac{d}{dt}\left[G_4(p_i, Q_i, t) + \sum_i p_i q_i - \sum_i P_i Q_i\right]$$

$$= \sum_i P_i \dot{Q}_i - K + \frac{dG_4}{dt} + \sum_i p_i \dot{q}_i + \sum_i \dot{p}_i q_i - \sum_i P_i \dot{Q}_i - \sum_i \dot{P}_i Q_i$$

or,
$$\frac{dG_4}{dt} = -\sum_i \dot{p}_i q_i + \sum_i \dot{P}_i Q_i + K - H$$

Expanding the total derivative of the generating function on the left hand side,

$$\sum_i \frac{\partial G_4}{\partial p_i}\dot{p}_i + \sum_i \frac{\partial G_4}{\partial P_i}\dot{P}_i + \frac{\partial G_4}{\partial t} = -\sum_i \dot{p}_i q_i + \sum_i \dot{P}_i Q_i + K - H$$

As before, the generalised quantities are independent of each other, we can equate the coefficients so that

$$q_i = -\frac{\partial G_4}{\partial p_i} \qquad Q_i = \frac{\partial G_4}{\partial P_i} \qquad K = H + \frac{\partial G_4}{\partial t}$$

It is interesting to note here that in all the four types of generating functions, the relation between the old and new hamiltonians are of the same form, the connecting generating functions appear in explicit time derivatives. So, it is obvious that if the generating functions are explicitly independent of time, the hamiltonian does not change on the canonical transformation in question, *i.e.*, $K = H$.

It will be worthwhile here to demonstrate the role canonical transformations towards simplifying a problem. We consider the problem of the simple harmonic oscillator. Since the system is conservative, its Hamiltonian is given by the sum of the kinetic and potential energies.

$$H = \frac{p^2}{2m} + \frac{1}{2}kq^2 = \frac{1}{2m}\left[p^2 + m^2\omega^2 q^2\right] \tag{9.1.6}$$

Our next objective is to transform the Hamiltonian to a new set of coordinates so as to increase the number of cyclic variables. As the Hamiltonian here is quadratic in the momentum and the spatial coordinates, we attempt a transformation of the form

$$p = f(P)\cos Q, \qquad q = \frac{f(P)}{m\omega}\sin Q \tag{9.1.7}$$

which leads to

$$H = K = \frac{f^2(P)}{2m} \tag{9.1.8}$$

The next step is to choose $f(P)$ such that the transformation becomes canonical. We select the generating function as

$$F_1(q,Q) = \frac{m\omega q^2}{2}\cot Q \tag{9.1.9}$$

This gives us the canonical momenta

$$p = \frac{\partial F_1}{\partial q} = m\omega q \cot Q, \qquad P = -\frac{\partial F_1}{\partial Q} = \frac{m\omega q^2}{2\sin^2 Q} \tag{9.1.10}$$

From these two equations we can solve for q and p

$$q = \sqrt{\frac{2P}{m\omega}}\sin Q, \qquad p = \sqrt{2Pm\omega}\cos Q \tag{9.1.11}$$

which implies

$$f(P) = \sqrt{2m\omega P} \tag{9.1.12}$$

Therefore the Hamiltonian is

$$K = H = \omega P \quad \Longrightarrow \quad P = \frac{E}{\omega} \tag{9.1.13}$$

where we have used the fact that a conservative system yeilds a Hamiltonian independent of time and equals the total energy of the system. We can see that the new Hamiltonian is cyclic and hence is amenable to easy solution, that since the variable Q is cyclic, the momentum is a constant of motion. The only equation of motion that needs to be solved is

$$\dot{Q} = \frac{\partial H}{P} = \omega \quad \Longrightarrow \quad Q = \omega t + \alpha \tag{9.1.14}$$

Substituting this back into equation (9.1.7), we find

$$q = \sqrt{\frac{2E}{m\omega^2}} \sin(\omega t + \alpha), \qquad p = \sqrt{2mE}\cos(\omega t + \alpha) \tag{9.1.15}$$

9.1.3 Conditions for Canonical Transformation

Let us consider a restricted canonical transformation *i.e.*, one in which time does not appear in the equations of transformations.

$$Q_i = Q_i(q,p) \qquad \text{and} \qquad P_i = P_i(q,p)$$

We have already seen that the hamiltonian remains unchanged in a restricted canonical transformation. So, the time derivative of Q_i and P_i can be written as

$$\dot{Q}_i = \frac{\partial Q_i}{\partial \dot{q}_j}\dot{q}_j + \frac{\partial Q_i}{\partial \dot{p}_j}\dot{p}_j \tag{9.1.16a}$$

$$\dot{P}_i = \frac{\partial P_i}{\partial \dot{q}_j}\dot{q}_j + \frac{\partial P_i}{\partial \dot{p}_j}\dot{p}_j \tag{9.1.16b}$$

But from Hamilton's equations of motion

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

$$\dot{p}_j = -\frac{\partial H}{\partial q_j}$$

Substituting $\dot{q}_j$ and $\dot{p}_j$ to the equations (9.1.16a), (9.1.16b) above, we get

$$\dot{Q}_i = \frac{\partial Q_i}{\partial q_j}\frac{\partial H}{\partial p_j} - \frac{\partial Q_i}{\partial p_j}\frac{\partial H}{\partial q_j} \tag{9.1.17}$$

$$\dot{P}_i = \frac{\partial P_i}{\partial q_j}\frac{\partial H}{\partial p_j} - \frac{\partial P_i}{\partial p_j}\frac{\partial H}{\partial q_j} \tag{9.1.18}$$

The inverse of the transformation equations are

$$q_i = q_i(Q,P)$$

$$p_i = p_i(Q, P)$$

All these equations enable us to consider $H(q, p, t)$ as a function of Q and P. Differentiating H with respect to Q and P we get

$$\frac{\partial H}{\partial P_i} = \frac{\partial H}{\partial p_j}\frac{\partial p_j}{\partial P_i} + \frac{\partial H}{\partial q_j}\frac{\partial q_j}{\partial P_i} \tag{9.1.19}$$

$$\frac{\partial H}{\partial Q_i} = \frac{\partial H}{\partial p_j}\frac{\partial p_j}{\partial Q_i} + \frac{\partial H}{\partial q_j}\frac{\partial q_j}{\partial Q_i} \tag{9.1.20}$$

But the Hamilton's equations of motion for Q and P are

$$\dot{Q}_i = \frac{\partial H}{\partial P_i} \tag{9.1.21}$$

$$\dot{P}_i = -\frac{\partial H}{\partial Q_i} \tag{9.1.22}$$

Comparing Equations (46) and (48) with equations (49) and (50) we can write

$$\frac{\partial Q_i}{\partial q_j}\frac{\partial H}{\partial p_j} - \frac{\partial Q_i}{\partial p_j}\frac{\partial H}{\partial q_j} = \frac{\partial H}{\partial p_j}\frac{\partial p_j}{\partial P_i} + \frac{\partial H}{\partial q_j}\frac{\partial q_j}{\partial P_i} \tag{9.1.23}$$

$$\frac{\partial P_i}{\partial q_j}\frac{\partial H}{\partial p_j} - \frac{\partial P_i}{\partial p_j}\frac{\partial H}{\partial q_j} = -\left[\frac{\partial H}{\partial p_j}\frac{\partial p_j}{\partial Q_i} + \frac{\partial H}{\partial q_j}\frac{\partial q_j}{\partial Q_i}\right] \tag{9.1.24}$$

Equating the coefficients on both sides of the above two equations we get,

$$\left(\frac{\partial Q_i}{\partial q_j}\right)_{(q,p)} = \left(\frac{\partial p_j}{\partial P_i}\right)_{(Q,P)} \tag{9.1.25}$$

$$\left(\frac{\partial Q_i}{\partial p_j}\right)_{(q,p)} = -\left(\frac{\partial q_j}{\partial P_i}\right)_{(Q,P)} \tag{9.1.26}$$

$$\left(\frac{\partial P_i}{\partial q_j}\right)_{(q,p)} = -\left(\frac{\partial p_j}{\partial Q_i}\right)_{(Q,P)} \tag{9.1.27}$$

$$\left(\frac{\partial P_i}{\partial p_j}\right)_{(q,p)} = \left(\frac{\partial q_j}{\partial Q_i}\right)_{(Q,P)} \tag{9.1.28}$$

The above four sets of equations are known as the direct conditions for a restricted canonical transformation.

Another version of the condition for canonically can be deduced from the considerations from the explicit time independence in the generating function, $F_1(q_i, Q_i)$ which is a function of the old and new generalised coordinates. If $F_1(q_i, Q_i)$ is given then we can write down its total differential as

$$dF = \sum_i \frac{\partial F_1}{\partial q_i}dq_i + \sum_i \frac{\partial F_1}{\partial Q_i}dQ_i$$

But from the type 1 generating function, we find that

$$p_i = \frac{\partial F_1}{\partial q_i}, \qquad P_i = -\frac{\partial F_1}{\partial Q_i}$$

Therefore, we can write the form of the total differential dF_1 as

$$dF_1 = \sum_i p_i dq_i - \sum_i P_i dQ_i$$

We see that in the canonical transformation from (q, p) to (Q, P), the expression $\sum_i p_i dq_i - \sum_i P_i dQ_i$ turns out to be a perfect differential. So, we conclude that

If the expression $\sum_i p_i dq_i - \sum_i P_i dQ_i$ can be written as a perfect differential, or in other words, if the expression $\sum_i p_i dq_i - \sum_i P_i dQ_i$ is exact differential or perfect differential, the transformation from (q, p) to (Q, P) is canonical.

Now we shall take up some canonical transformations and discuss the role of the generating functions through some simple examples.

Example 9.1.1 *Show that the transformation*

$$Q = \tan^{-1}\frac{q}{p}$$

$$P = \frac{1}{2}(q^2 + p^2)$$

is canonical. Find the new form of the Hamiltonian obtained after the said transformation when the old one is given by

$$H = \frac{1}{2}(q^2 + p^2)$$

Solution: In order to establish that the given transformation is canonical, we consider the following differential form

$$\sum_{i=1}^{n} p_i dq_i - \sum_{i=1}^{n} P_i dQ_i = dF$$

with

$$Q = \tan^{-1}\frac{q}{p}, \qquad P = \frac{1}{2}(q^2 + p^2)$$

so that

$$dQ = \frac{p^2}{p^2 + q^2} d\left(\frac{q}{p}\right) = \frac{pdq - qdp}{p^2 + q^2}.$$

The above expression in terms of the old variables is therefore

$$pdq - PdQ = pdq - \frac{p^2+q^2}{2}.\frac{pdq-qdp}{p^2+q^2}$$
$$= pdq - \frac{pdq-qdp}{2} = \frac{pdq+qdp}{2}$$

or,
$$pdq - PdQ = d\left(\frac{pq}{2}\right)$$
$$= dF = \text{perfect differential}$$

Hence the transformation is canonical. The corresponding generating function is $F = \frac{pq}{2}$, which can be expressed in terms of the type 1 as the following.

From the given transformation,

$$Q = \tan^{-1}\frac{q}{p}$$

or,
$$\tan Q = \frac{q}{p}$$
$$\Longrightarrow \quad p = q\cot Q$$
$$\therefore \quad F = \frac{pq}{2} = \frac{q^2}{2}\cot Q \equiv F_1$$

So the generating function is of type 1 with $F_1 = \frac{q^2}{2}\cot Q$.

As the generating function is explicitly independent of time, $\frac{\partial F_1}{\partial t} = 0$. So, the New hamiltonian equals the old hamiltonian, *i.e.*,

$$K = H = \frac{1}{2}(p^2+q^2).$$

Example 9.1.2 *Prove that the rheonomic transformation*

$$Q = \sqrt{2q}e^t\cos p, \qquad P = \sqrt{2q}e^{-t}\sin p$$

is canonical.

Solution: Given the transformations

$$Q = \sqrt{2q}e^t\cos p, \qquad P = \sqrt{2q}e^{-t}\sin p$$

from which we first obtain

$$dQ = \left(\frac{\cos p dq}{\sqrt{2q}} - \sqrt{2q}\sin p dp\right)e^t = \frac{\cos p dq - 2q\sin p dp}{\sqrt{2q}}e^t$$
$$\therefore \quad PdQ = \sin p\cos p dq - 2q\sin^2 p dp$$

so that

$$pdq - PdQ = (p - \sin p \cos p)dq + 2q \sin^2 p dp = Mdq + Ndp \qquad \text{(say)},$$

$$\begin{aligned} \text{with} \qquad & M = p - \sin p \cos p \\ \Longrightarrow \qquad & \frac{\partial M}{\partial p} = 1 - \cos 2p = 2\sin^2 p \end{aligned}$$

$$\begin{aligned} \text{and} \qquad & N = 2q \sin^2 p \\ \Longrightarrow \qquad & \frac{\partial N}{\partial q} = 2\sin^2 p. \end{aligned}$$

The exactness check condition of

$$\frac{\partial M}{\partial p} = \frac{\partial N}{\partial q}$$

indicates that the expression $pdq - PdQ$ is a perfect differential and hence the transformation in question is canonical.

Example 9.1.3 *For a given canonical transformation, the following are supplied:*

$$Q = \sqrt{(q^2 + p^2)}, \qquad F = \frac{1}{2}(q^2 + p^2) \tan^{-1} \frac{q}{p} + \frac{1}{2} qp$$

From the information find the transformation $P(q, p)$ and the generating function $F(q, Q)$ as a function of the old and the new coordinates.

Solution: First we solve for p from the given expression of Q and substitute it in the given F to reexress it in terms of $F(q, Q)$, *i.e.*,

$$p = \sqrt{Q^2 - q^2}$$

so that

$$\begin{aligned} F &= \frac{Q^2}{2} \tan^{-1} \frac{q}{\sqrt{Q^2 - q^2}} + \frac{q}{2}\sqrt{Q^2 - q^2} \\ &= \frac{Q^2}{2} \sin^{-1} \frac{q}{Q} + \frac{q}{2}\sqrt{Q^2 - q^2} \end{aligned}$$

As this depends on the old and the new generalised coordinates, it is the type 1 generating function. So, $F \equiv F_1(q, Q)$. Hence

$$
\begin{aligned}
P &= -\frac{\partial F_1}{\partial Q} \\
&= -\frac{\partial}{\partial Q}\left[\frac{Q^2}{2}\sin^{-1}\frac{q}{Q} + \frac{q}{2}\sqrt{Q^2 - q^2}\right] \\
&= -\left[Q\sin^{-1}\frac{q}{Q} + \frac{Q^2}{2}\frac{(-q)}{Q\sqrt{Q^2-q^2}} + \frac{q}{2}\left(\frac{2\dot{Q}}{2\sqrt{Q^2-q^2}}\right)\right] \\
&= -\left[Q.\sin^{-1}\frac{q}{Q} - \frac{1}{2}\frac{qQ}{\sqrt{Q^2-q^2}} + \frac{1}{2}\frac{qQ}{\sqrt{Q^2-q^2}}\right] \\
\text{or,} \qquad P(q, Q) &= -Q\sin^{-1}\frac{q}{Q}
\end{aligned}
$$

Now for obtaining $P(q, p)$, we need to substitute Q for $\sqrt{q^2 + p^2}$ to get

$$
\begin{aligned}
P(q, p) &= -\sqrt{(q^2+p^2)}\sin^{-1}\frac{q}{\sqrt{(q^2+p^2)}} \\
\text{or,} \qquad P(q, p) &= -\sqrt{(q^2+p^2)}\tan^{-1}\frac{q}{p}.
\end{aligned}
$$

which is the required transformation.

9.1.4 Some Special Canonical Transformations

1. **Identity Transformation**

   Let us consider a generating function of the type

$$F \equiv F_2 = \sum_i q_j P_j$$

   Utilizing the relations for the generating function of type 2

$$
\begin{aligned}
p_i &= \frac{\partial F_2}{\partial q_i} \\
&= \frac{\partial}{\partial q_i}\left(\sum_j q_j P_j\right) \\
&= \sum_j \frac{\partial q_j}{\partial q_i}P_j + \sum_j q_j\frac{\partial P_j}{\partial q_i} \\
&= \sum_j \delta_{ji}P_j + 0, \qquad \left[\because \quad \frac{\partial q_j}{\partial q_i} = \text{Kronecker delta} = \delta_{ji} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}\right] \\
\text{or,} \qquad p_i &= P_i
\end{aligned}
$$

Similarly,

$$\begin{aligned}Q_i &= \frac{\partial F_2}{\partial P_i} \\ &= \frac{\partial}{\partial P_i}\left(\sum_j q_j P_j\right) \\ &= \sum_j \frac{\partial q_j}{\partial P_i} P_j + \sum_j q_j \frac{\partial P_j}{\partial P_i} \\ &= 0 + \sum_j \delta_{ji} q_j; \qquad [\because \; q_i, \; P_i \text{ are mutually independent.}]\end{aligned}$$

or, $$Q_i = q_i$$

Thus we see that the generator $F \equiv F_2 = \sum_i q_j P_j$ generates the new transformation such that the old and new values are identical, *i.e., identity transformation.* Further since the generator is explicitly independent of time t, the hamiltonian will also remain unchanged upon this transformation.

Summarising,

$q_i = Q_i, \quad p_i = P_i, \quad K = H$	for the generating function	$F = \sum_i q_i P_i$

2. **Coordinate Swap**
We now consider the coordinate swap or the exchange transformation in which the new position coordinates are the old momenta, and the new momenta turn out to be the old position coordinates (except for a negative sign, as we can see below). For this we consider the generating function of the type

$$F \equiv F_1 = \sum_j q_j Q_j$$

which is the generator of type 1. In this case we find that

$$\begin{aligned}p_i &= \frac{\partial F_1}{\partial q_i} \\ &= \frac{\partial}{\partial q_i}\left(\sum_j q_j Q_j\right) \\ &= \sum_j \delta_{ji} Q_j + 0\end{aligned}$$

or, $$p_i = Q_i$$

and

$$P_i = -\frac{\partial F_1}{\partial Q_j}$$
$$= -\frac{\partial}{\partial Q_j}\left(\sum_j q_j Q_j\right)$$
$$= -\sum_j \delta_{ji} q_j + 0$$

or, $$P_i = -q_i$$

Thus we see that the given generator serves to generate a transformation which causes an exchange of the role of the canonical variables: momenta behaves like coordinates and the coordinates behave like momenta. This is the *exchange transformation.* Further the generator being explicitly independent of time, the hamiltonian also does not change upon transformation.

So, we see that

$p_i = Q_i, \quad P_i = -q_i, \quad K = H$	for the generating function	$F = \sum_i q_i Q_i$

9.2 Lagrange Bracket and Poisson Bracket

Apart from the test of canonicality as laid in the previous sections, there exists another viewpoint of testing the canonicality of the dynamical variables and translations in relation to the volume defined by the canonically conjugate variables. For a dynamical system with n degrees of freedom we have already seen that Hamilton's canonicalequations admit $2n$ canonically conjugate variables of which n numbers are the generalised coordinates and rest n the canonically conjugate momentum coordinates. It is then possible to conceive of a 'space' spanned by these $2n$ coordinates in relation to the given dynamical system. This space of coordinates and momenta is also called a *phase space* corresponding to the given system and every point of this space defines a particular state of the dynamical system.

There exists two known classes of relations between the dynamical variables which can be used to test the canonicality of a given transformation. One class is the *Lagrange Bracket* relations and the other is the *Poisson Bracket* relation. There also exists a mutual relationship in between these two classes. We start our discussions by defining the bracket relations and get along with finding their relationships and utilities.

9.2.1 Lagrange Bracket

This class of relationship in the dynamic variables was introduced by Joseph Lagrange in 1808. For any two independent dynamical variables $u(q, p, t)$ and $v(q, p, t)$ of a dynamical system with n degrees of freedom, the Lagrange Bracket is defined as

$$(u, v)_{(q,p)} = \sum_{i=1}^{n} \left(\frac{\partial q_i}{\partial u}\frac{\partial p_i}{\partial v} - \frac{\partial q_i}{\partial v}\frac{\partial p_i}{\partial u} \right) = \sum_{i=1}^{n} \frac{\partial(q_i, p_i)}{\partial(u, v)}$$

From the definition itself it is clear that the Lagrange Bracket antisymmetric with respect to the variables u, v, *i.e.*,

$$(u, v) = -(v, u)$$

Properties of Lagrange Bracket

Some of the properties of the Lagrange Bracket are listed below.

1. $(u \pm a, v \pm b) = (u, v)$.
2. $(au, bv) = \frac{1}{ab}(u, v)$.
3. $(q_k, q_l) = 0 = (p_k, p_l)$.
4. $(q_i, p_j) = \delta_{ij}$.
5. $(u, v)_{(q \pm a, p \pm b)} = (u, v)_{(q,p)}$.
6. $(u, v)_{(aq, bp)} = ab(u, v)_{(q,p)}$.

9.2.2 Poisson Brackets

The Poisson Brackets are important algebraic structures available for any Hamiltonian system. The properties of Poisson Brackets serve to offer an elegant transition from the classical mechanics to quantum mechanics where the Poisson bracket in the dynamical variables transits to the commutators in the quantum mechanical operators.

Consider any two functions on phase space, $f(q_i, p_i)$ and $g(q_i, p_i)$, such as the components of the linear or angular momentum, or the energy etc. The Poisson bracket of this pair, denoted as $[f, g]$, is defined as

$$[f, g] = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} \right) \tag{9.2.1}$$

The Poisson bracket has a number of important properties, as discussed below.

Some properties of Poisson Bracket

1. The Poisson bracket of any two dynamical variables is antisymmetric, *i.e.*,

$$[u, v] = -[v, u]$$

and the corollary

$$[u,\ u] = [u,\ u] = 0$$

The Poisson bracket of u with itself is zero.

2. The poisson bracket of u with any constant c is zero, *i.e.*, $[u,\ c] = 0$.

3. If c is a constant, independent of the canonical variables q, p and the time t, then

$$[cu,\ v] = [u,\ cv] = c[u,\ v]$$

4. The Poisson brackets satisfy the distributive property

$$\begin{aligned}
[u \pm v,\ w] &= [u,\ w] \pm [v,\ w] \\
[u,\ v \pm w] &= [u,\ v] \pm [u,\ w] \\
[uv,\ w] &= [v,\ w]u + v[u,\ w] \\
[u,\ vw] &= [u,\ v]w + v[u,\ w]
\end{aligned}$$

5. The partial and the total derivative of any Poisson bracket relation satisfy the following relations

$$\frac{\partial}{\partial t}[u,\ v] = \left[\frac{\partial u}{\partial t},\ v\right] + \left[u,\ \frac{\partial v}{\partial t}\right]$$

and

$$\frac{d}{dt}[u,\ v] = \left[\frac{du}{dt},\ v\right] + \left[u,\ \frac{dv}{dt}\right]$$

6. Poisson brackets of three variables X, Y and Z; $[X, Y], [Y, Z]$ and $[Z, X]$ are related by the following identity,which are known as Jacobi identity.

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$$

From a mathematical point of view, these properties of the Poisson Brackets mean that the set of phase space functions is endowed with the structure of a Lie algebra. From a more practical point of view, these properties can be used to simplify the computations in poisson brackets, once we have the poisson brackets of the canonically conjugate variables, *i.e.*, $(q_i,\ p_i)$, called the fundamental Poisson Brackets:

$$[q_i,\ p_j] = \delta_{ij}, \qquad \text{and} \qquad [q_i,\ q_j] = 0 = [p_i,\ p_j]$$

For example, to compute the Poisson Bracket of $f = qp$ with $q = pe^q$ we get

$$\begin{aligned}
[f,\ g] &= [qp,\ pe^q] \\
[f,\ g] &= q[p,\ pe^q] + p[q,\ pe^q] \\
&= qp[p,\ e^q] + qe^q[p,\ p] + p^2[q,\ e^q] + pe^q[q,\ p] \\
&= -qpe^q + pe^q.
\end{aligned}$$

Here we used the fact that

$$[g(q),\ f(q)] = 0 = [h(p),\ k(p)],$$
$$[q_i,\ f(p)] = \frac{\partial f(p)}{\partial p_i},$$
and
$$[p_i,\ f(q)] = -\frac{\partial f(q)}{\partial q^i}$$

We see that the Poisson brackets have a fundamental role to play in mechanics relative to canonical transformations.

Example 9.2.1 *Prove that the distributive law*

$$[F, G+K] = [F, G] + [F, K]$$

holds good for poisson brackets.

Proof:

$$\begin{aligned}
[F, G+k] &= \sum_k \left(\frac{\partial F}{\partial q_k}\frac{\partial (G+K)}{\partial p_k} - \frac{\partial F}{\partial p_k}\frac{\partial (G+K)}{\partial q_k}\right)\\
&= \sum_k \left(\frac{\partial F}{\partial q_k}\frac{\partial G}{\partial p_k} + \frac{\partial F}{\partial q_k}\frac{\partial K}{\partial p_k}\right) - \sum_k \left(\frac{\partial F}{\partial p_k}\frac{\partial G}{\partial q_k} + \frac{\partial F}{\partial p_k}\frac{\partial K}{\partial q_k}\right)\\
&= \sum_k \left(\frac{\partial F}{\partial q_k}\frac{\partial G}{\partial p_k} - \frac{\partial F}{\partial p_k}\frac{\partial G}{\partial q_k}\right) + \sum_k \left(\frac{\partial F}{\partial q_k}\frac{\partial K}{\partial p_k} - \frac{\partial F}{\partial p_k}\frac{\partial K}{\partial q_k}\right)\\
&= [F, G] + [F, K]
\end{aligned}$$

Example 9.2.2 *If $[\phi, \psi]$ be the Poisson bracket, then prove that*

$$\frac{\partial}{\partial t}[\phi, \psi] = \left[\frac{\partial \phi}{\partial t}, \psi\right] + \left[\phi, \frac{\partial \psi}{\partial t}\right]$$

Proof: From the definition of Poisson bracket,

$$\begin{aligned}
[\phi, \psi] &= \sum_k \left(\frac{\partial \phi}{\partial q_k}\frac{\partial \psi}{\partial p_k} - \frac{\partial \phi}{\partial p_k}\frac{\partial \psi}{\partial q_k}\right)\\
\frac{\partial}{\partial t}[\phi, \psi] &= \sum_k \left[\frac{\partial}{\partial q_k}\left(\frac{\partial \phi}{\partial t}\right)\frac{\partial \psi}{\partial p_k} + \frac{\partial \phi}{\partial q_k}\frac{\partial}{\partial p_k}\left(\frac{\partial \psi}{\partial t}\right) - \frac{\partial}{\partial p_k}\left(\frac{\partial \phi}{\partial t}\right)\frac{\partial \psi}{\partial q_k} - \frac{\partial \phi}{\partial p_k}\frac{\partial}{\partial q_k}\left(\frac{\partial \psi}{\partial t}\right)\right]\\
&= \sum_k \left[\frac{\partial}{\partial q_k}\left(\frac{\partial \phi}{\partial t}\right)\frac{\partial \psi}{\partial p_k} - \frac{\partial}{\partial p_k}\left(\frac{\partial \phi}{\partial t}\right)\frac{\partial \psi}{\partial q_k}\right] + \sum_k \left[\frac{\partial \phi}{\partial q_k}\frac{\partial}{\partial p_k}\left(\frac{\partial \psi}{\partial t}\right) - \frac{\partial \phi}{\partial p_k}\frac{\partial}{\partial q_k}\left(\frac{\partial \psi}{\partial t}\right)\right]\\
&= \left[\frac{\partial \phi}{\partial t}, \psi\right] + \left[\phi, \frac{\partial \psi}{\partial t}\right]
\end{aligned}$$

Jacobi Identity

For any three functions F, G and K of p_k and q_k, the following identity holds true :

$$[F,[G,K]]+[G,[K,F]]+[K,[F,G]]=0.$$

This relation is known as *Jacobi Identity.*

Proof of Jacobi identity

Let us consider the following expression:

$$\begin{aligned}
[F,[G,K]]+[G,[K,F]] &= \left[F,\sum_k\left(\frac{\partial G}{\partial q_k}\frac{\partial K}{\partial p_k}-\frac{\partial G}{\partial p_k}\frac{\partial K}{\partial q_k}\right)\right]-\left[G,\sum_k\left(\frac{\partial F}{\partial q_k}\frac{\partial K}{\partial p_k}-\frac{\partial F}{\partial p_k}\frac{\partial K}{\partial q_k}\right)\right]\\
&= \left[F,\sum_k\left(\frac{\partial G}{\partial q_k}\frac{\partial K}{\partial p_k}\right)\right]-\left[F,\sum_k\left(\frac{\partial G}{\partial p_k}\frac{\partial K}{\partial q_k}\right)\right]\\
&\quad -\left[G,\sum_k\left(\frac{\partial F}{\partial q_k}\frac{\partial K}{\partial p_k}\right)\right]+\left[G,\sum_k\left(\frac{\partial F}{\partial p_k}\frac{\partial K}{\partial q_k}\right)\right]
\end{aligned}$$

Now, using the property $[u,\ vw]=[u,\ v]w+[u,\ w]v$, we have

$$\begin{aligned}
[F,\ [G,\ K]]+[G,\ [K,\ F]] &=\sum_k\left\{\left[F,\ \frac{\partial G}{\partial q_k}\right]\frac{\partial K}{\partial p_k}+\left[F,\ \frac{\partial K}{\partial p_k}\right]\frac{\partial G}{\partial q_k}-\left[F,\ \frac{\partial G}{\partial p_k}\right]\frac{\partial K}{\partial q_k}-\left[F,\ \frac{\partial K}{\partial q_k}\right]\frac{\partial G}{\partial q_k}\right.\\
&\quad\left.-\left[G,\ \frac{\partial F}{\partial q_k}\right]\frac{\partial K}{\partial p_k}-\left[G,\ \frac{\partial K}{\partial p_k}\right]\frac{\partial F}{\partial q_k}+\left[G,\ \frac{\partial F}{\partial p_k}\right]\frac{\partial K}{\partial q_k}+\left[G,\ \frac{\partial K}{\partial q_k}\right]\frac{\partial F}{\partial p_k}\right\}\\
&=\sum_k\left\{\left[F,\ \frac{\partial G}{\partial q_k}\right]-\left[G,\ \frac{\partial F}{\partial q_k}\right]\right\}\frac{\partial K}{\partial p_k}+\sum_k\left\{-\left[F,\ \frac{\partial G}{\partial p_k}\right]+\left[G,\ \frac{\partial F}{\partial p_k}\right]\right\}\frac{\partial K}{\partial q_k}\\
&\quad+\sum_k\left\{\left[F,\ \frac{\partial K}{\partial p_k}\right]\frac{\partial G}{\partial q_k}-\left[F,\ \frac{\partial K}{\partial q_k}\right]\frac{\partial G}{\partial q_k}-\left[G,\ \frac{\partial K}{\partial p_k}\right]\frac{\partial F}{\partial q_k}+\left[G,\ \frac{\partial K}{\partial q_k}\right]\frac{\partial F}{\partial p_k}\right\}\\
&=\sum_k\left\{\left[F,\ \frac{\partial G}{\partial q_k}\right]-\left[G,\ \frac{\partial F}{\partial q_k}\right]\right\}\frac{\partial K}{\partial p_k}+\sum_k\left\{-\left[F,\ \frac{\partial G}{\partial p_k}\right]+\left[G,\ \frac{\partial F}{\partial p_k}\right]\right\}\frac{\partial K}{\partial q_k}+I
\end{aligned}$$

$$\text{where,}\quad I=\sum_k\left\{\left[F,\ \frac{\partial K}{\partial p_k}\right]\frac{\partial G}{\partial q_k}-\left[F,\ \frac{\partial K}{\partial q_k}\right]\frac{\partial G}{\partial q_k}-\left[G,\ \frac{\partial K}{\partial p_k}\right]\frac{\partial F}{\partial q_k}+\left[G,\ \frac{\partial K}{\partial q_k}\right]\frac{\partial F}{\partial p_k}\right\}$$

$$\therefore\quad [F,\ [G,\ K]]+[G,\ [K,\ F]]=\sum_k\left\{\left(\frac{\partial}{\partial q_k}[F,\ G]\right)\frac{\partial K}{\partial p_k}-\left(\frac{\partial}{\partial p_k}[F,\ G]\right)\frac{\partial K}{\partial q_k}\right\}+I$$

$$=[[F,\ G],\ K]+I=-[K,\ [F,\ G]]+I$$

Now,

$$
\begin{aligned}
I =& \sum_k \left\{ \left[F,\ \frac{\partial K}{\partial p_k}\right] \frac{\partial G}{\partial q_k} - \left[F,\ \frac{\partial K}{\partial q_k}\right] \frac{\partial G}{\partial q_k} - \left[G,\ \frac{\partial K}{\partial p_k}\right] \frac{\partial F}{\partial q_k} + \left[G,\ \frac{\partial K}{\partial q_k}\right] \frac{\partial F}{\partial p_k} \right\} \\
=& \sum_k \left\{ \frac{\partial G}{\partial q_k} \frac{\partial F}{\partial q_k} \frac{\partial^2 K}{\partial p_k^2} - \frac{\partial F}{\partial p_k} \frac{\partial^2 K}{\partial q_k \partial p_k} \frac{\partial G}{\partial q_k} - \frac{\partial G}{\partial p_k} \frac{\partial F}{\partial q_k} \frac{\partial^2 K}{\partial p_k \partial q_k} + \frac{\partial G}{\partial p_k} \frac{\partial F}{\partial p_k} \frac{\partial^2 K}{\partial q_k^2} \right. \\
& \left. - \frac{\partial G}{\partial q_k} \frac{\partial^2 K}{\partial p_k^2} \frac{\partial F}{\partial q_k} + \frac{\partial G}{\partial p_k} \frac{\partial^2 K}{\partial q_k \partial p_k} \frac{\partial F}{\partial q_k} + \frac{\partial G}{\partial q_k} \frac{\partial^2 K}{\partial p_k \partial q_k} \frac{\partial F}{\partial p_k} - \frac{\partial G}{\partial p_k} \frac{\partial^2 K}{\partial q_k^2} \frac{\partial F}{\partial p_k} \right\} \\
=& 0
\end{aligned}
$$

Hence, $$[F,\ [G,\ K]] + [G,\ [K,\ F]] = -[K,\ [F,\ G]]$$

$$\implies \boxed{[F,\ [G,\ K]] + [G,\ [K,\ F]] + [K,\ [F,\ G]] = 0}$$

which proves *Jacobi's identity*

9.2.3 Poisson Bracket and Canonical Transformation

The importance of the Poisson Bracket essentially lies on its property of the covariance or the form invariance under canonical transformation. It states that

The Poisson bracket of dynamical variables defined on a set of canonically conjugate variables is covariant under the canonical transformation.

Proof: Let $F(q,\ p)$ and $G(q,\ p)$ be two dynamical variables at any given moment t and corresponding to a system described by Hamilton's canonical equations corresponding to n degrees of freedom and with the canonically conjugate pairs $(q,\ p)$. Then the Poisson Bracket associated with the system is defined by

$$[F, G]_{q,p} = \sum_i \left(\frac{\partial F}{\partial q_i} \frac{\partial G}{\partial p_i} - \frac{\partial G}{\partial p_i} \frac{\partial F}{\partial q_i} \right)$$

Now we impose canonical transformation of the set $(q,\ p)$ to a new set say $(Q,\ P)$ such that

$$
\begin{aligned}
q_k &= q_k(Q_i, P_i) \\
p_k &= p_k(Q_i, P_i)
\end{aligned}
$$

which will modify the form the dynamical variables in the Poisson Bracket as

$$
\begin{aligned}
[F,\ G]_{(q,p)} &= \sum_{ij} \left[\left(\frac{\partial F}{\partial Q_j} \frac{\partial Q_j}{\partial q_i} + \frac{\partial F}{\partial P_j} \frac{\partial P_j}{\partial q_i} \right) \frac{\partial G}{\partial p_i} - \left(\frac{\partial F}{\partial Q_j} \frac{\partial Q_j}{\partial p_i} + \frac{\partial F}{\partial P_j} \frac{\partial P_j}{\partial p_i} \right) \frac{\partial G}{\partial q_i} \right] \\
&= \sum_j \left[\frac{\partial F}{\partial Q_j} \sum_i \left(\frac{\partial Q_j}{\partial q_i} \frac{\partial G}{\partial p_i} - \frac{\partial Q_j}{\partial p_i} \frac{\partial G}{\partial q_i} \right) + \frac{\partial F}{\partial P_j} \sum_i \left(\frac{\partial P_j}{\partial q_i} \frac{\partial G}{\partial p_i} - \frac{\partial P_j}{\partial p_j} \frac{\partial G}{\partial q_i} \right) \right]
\end{aligned}
$$

$$[F,\ G]_{(q,p)} = \sum_j \left[\frac{\partial F}{\partial Q_j}[Q_j,\ G] + \frac{\partial F}{\partial P_j}[P_j,\ G]\right] \tag{9.2.2}$$

Putting $G = Q_k$ and $F = G$ in (9.2.2) we get

$$\begin{aligned}[G,\ Q_k]_{(q,p)} &= \sum_j \left[\frac{\partial G}{\partial Q_j}[Q_j,\ Q_k] + \frac{\partial G}{\partial P_j}[P_j,\ Q_k]\right] \\ &= \sum_j \frac{\partial G}{\partial P_j}(-\delta_{jk}) \\ &= -\frac{\partial G}{\partial P_k}\end{aligned}$$

Again putting $G = P_k$ and $F = G$ in (9.2.2)

$$\begin{aligned}[G,\ P_k]_{(q,p)} &= \sum_j \left[\frac{\partial G}{\partial Q_j}[Q_j,\ P_k]\right] \\ &= \sum_j \frac{\partial G}{\partial Q_j}(\delta_{jk}) \\ &= \frac{\partial G}{\partial Q_k}\end{aligned}$$

From (9.2.2)

$$\begin{aligned}[F,\ G]_{(q,p)} &= \sum_j \left(\frac{\partial F}{\partial Q_j}\frac{\partial G}{\partial P_j} - \frac{\partial F}{\partial P_j}\frac{\partial G}{\partial Q_j}\right) \\ &= [F,\ G]_{(Q,P)}\end{aligned}$$

Therefore the Poisson Bracket is conserved under canonical transformation.

9.2.4 Equation of Motion in Poisson Bracket

From Newton's time there has been the attempt to write down equations of motion for any dynamical system and look for conserved quantities, if any. The Newton's equation of motion, though elegant, had some inconveniences as the forces of constraints which are required to be included in the equaiton of motion yet could not be evaluated apriori. Moreover, in Newton's formalism there was no direct way of counting the number of conserved quantities, let alone their estimation. The quantities are first suspected to be conserved quantities requiring the verification subsequently - whether they are conserved quantities or not. In lagrangian and Hamiltonian formalism, these difficulties are partially removed: the number conserved quantities could be counted and estimated by reading off the dependence of the Lagrangian or the Hamiltonian on the generalised coordinates. The number of cyclic coordinates directly gives the number of conserved quantities of the system. Moreover the Hamiltonian formalism could provide some amount of symmetry in the structure of the equations.

The equaitons of motion for a dynamical system can also be written in terms of Poisson Bracket. The Poisson Bracket formalism of classical dynamics is beset with the symmetry in the dynamical equations and is capable of not only ascertaining the conserved quantities but also evaluating the conserved quantities in an exhaustive manner.

Following is the discussion and the deduction of the necessary equations and relations towards the description of a dynamical system and the associated conserved quantities.

We first suppose that for a given dynamical system, the Hamiltonian $H(q_i,\ p_i, t)$ is known, where q's are the generalised coordiinates and p's the canonical momenta.

Let us proceed to evaluate the Poisson Bracket of one of the generalised coordinates with the Hamiltonian H and also generalized momentum with the Hamiltonian H:

$$[q_k,\ H]_{(q,\ p)} = \sum_i \left(\frac{\partial q_k}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial q_k}{\partial p_i}\frac{\partial H}{\partial q_i} \right) \tag{9.2.3}$$

$$[p_k,\ H]_{(q,\ p)} = \sum_i \left(\frac{\partial p_k}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial p_k}{\partial p_i}\frac{\partial H}{\partial q_i} \right) \tag{9.2.4}$$

The corresponding Hamilton's canonical equations of motion are

$$\dot{q_i} = \frac{\partial H}{\partial p_i}, \qquad \dot{p_i} = -\frac{\partial H}{\partial q_i} \tag{9.2.5}$$

With the help of the equations (9.2.3), (9.2.4) and (9.2.5) we can write the following

$$[q_k,\ H] = \sum_i \delta_{ik}\frac{\partial H}{\partial p_i} = \frac{\partial H}{\partial p_k} = \dot{q_i} \tag{9.2.6}$$

$$[p_k,\ H] = \sum_i (-\delta_{ik})\frac{\partial H}{\partial q_i} = -\frac{\partial H}{\partial q_k} = \dot{p_k} \tag{9.2.7}$$

Now let u be any dynamical quantity associated with the system such that $u \equiv u(q_i, p_i, t)$. Then we can express the equation of motion for u as

$$\begin{aligned}
\frac{du}{dt} &= \sum_i \left(\frac{\partial u}{\partial q_i}\frac{dq_i}{dt} + \frac{\partial u}{\partial p_i}\frac{dp_i}{dt} \right) + \frac{\partial u}{\partial t} \\
&= \sum_i \left(\frac{\partial u}{\partial q_i}\dot{q_i} + \frac{\partial u}{\partial p_i}\dot{p_i} \right) + \frac{\partial u}{\partial t} \\
&= \sum_i \left(\frac{\partial u}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial u}{\partial p_i}\frac{\partial H}{\partial q_i} \right) + \frac{\partial u}{\partial t} \\
\frac{du}{dt} &= [u, H] + \frac{\partial u}{\partial t}
\end{aligned}$$

If u is explicitly independent of time, then $\dfrac{\partial u}{\partial t} = 0$.

$$\therefore \quad \frac{du}{dt} = [u, H]$$

If the dynamical quantity u happens to be a conserved quantity, we must have

$$\frac{du}{dt} = 0, \qquad \text{so that} \quad [u,\ H] = 0$$

i.e., If the poisson bracket of a dynamical quantity with the hamiltonian vanishes, the quantity must be a constant of motion.

Now we consider two case which leads to two important results:

Case(1) If $u = q_i \qquad \Longrightarrow \qquad \dot{q}_i = [q_i, H]$

Case(2) If $u = p_i \qquad \Longrightarrow \qquad \dot{p}_i = [p_i, H]$

Conserved quantities

An interesting result follows in regards of the conserved quantities of a dynamical system when we use the poisson bracket and in particular, the Jacobi identity. If the system has at least two conserved quantities known apriori, it is possible to excavate all the remaining conserved quantities thought the poisson bracket and the Jacobi identity. To proceed, let u and v be two given conserved quantities associated with a system with hamiltonian H. Then it follows that

$$[u,\ H] = 0 \qquad \text{and} \qquad [v,\ H] = 0 \tag{9.2.8}$$

Using Jacobi identity,

$$\big[u,\ [v,\ w]\big] + \big[v,\ [w,\ u]\big] + \big[w,\ [u,\ v]\big] = 0$$

Let, $w = H$. Then Jacobi identity yields

$$\big[u,\ [v,\ H]\big] + \big[v,\ [H,\ u]\big] + \big[H,\ [u,\ v]\big] = 0 \tag{9.2.9}$$

Since, we have $[u,\ H] = 0$ and $[v,\ H] = 0$, the first and the second term of (9.2.9) becomes zero and we get

$$\big[H,\ [u,\ v]\big] = 0 \qquad \Longrightarrow \qquad \big[[u,\ v],\ H\big] = 0$$

i.e., the poisson bracket of $[u,\ v]$ with the hamiltonian vanishes, implying that $[u,\ v] = z$ is a constant of motion.

Note!

Once the hamiltonian for a system is known along with two conserved quantities, the use of Jacobi identity yields the third constant of motion. Repeating this process with the choice of a pair of conserved quantities out of the three available now, can similarly yield another constant of motion. This process can be made to go on till the exhaustive set of constants of motion associated with the dynamical system are excavated. Any further repetition of the process thereafter will merely yield some member of the set of the exhaustive constants of motion.

Example 9.2.3 *A harmonic oscillator of mass m and the spring constant k executes motion in one dimension such that the hamiltonian is given in terms of a single generalised coordinate q and single conjugate momentum p, by*

$$H = \frac{p^2}{2m} + \frac{kq^2}{2}.$$

Show, with the help of Poisson Bracket that the transformation defined for the system by

$$P = \frac{\alpha q^2}{2}\left(1 + \frac{p^2}{\alpha^2 q^2}\right)$$

$$Q = tan^{-1}\frac{\alpha q}{p}, \qquad where, \qquad \alpha = \sqrt{km}$$

is Canonical.

Solution: As the system is defined by the canonical pair $(q,\ p)$, the Poisson Bracket expression yields

$$[q,\ p] = 1, \qquad [q,\ q] = 0 = [p,\ p]$$

The transformation of the variables to $(Q,\ P)$ will be a canonical transformation for the harmonic oscillator if

$$[Q,\ P]_{(q,\,p)} = 1, \qquad \text{with} \qquad [P,\ P] = [Q,\ Q] = 0.$$

We find,

$$\frac{\partial Q}{\partial p} = -\frac{\alpha q}{p^2 + \alpha^2 q^2} \quad ; \quad \frac{\partial Q}{\partial q} = \frac{\alpha p}{p^2 + \alpha^2 q^2} \quad ; \quad \frac{\partial P}{\partial p} = \frac{p}{\alpha} \quad ; \quad \frac{\partial P}{\partial q} = \alpha q.$$

$$\begin{aligned}[Q, P]_{(q,p)} &= \frac{\partial Q}{\partial q}\frac{\partial P}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial P}{\partial q} \\ &= \frac{\alpha p}{p^2 + \alpha^2 q^2}.\frac{p}{\alpha} + \frac{\alpha q}{p^2 + \alpha^2 q^2}.\alpha q = \frac{p^2 + \alpha^2 q^2}{p^2 + \alpha^2 q^2} = 1\end{aligned}$$

Hence the transformation is canonical.

Example 9.2.4 *Using Poisson Brackets, show that the transformation* $Q = \frac{1}{p}$; $\quad P = qp^2$ *is canonical.*

Solution: For the old pair of conjugate variables $(q,\ p)$, we have $[q,\ p] = 1$.
Now we need to show that, $[Q,\ P]_{(q,p)} = 1$

$$\begin{aligned}[Q,\ P] &= \left[\frac{1}{p},\ qp^2\right] \\ &= \left[\frac{1}{p},\ q\right]p^2 + \left[\frac{1}{p},\ p^2\right]q \\ &= \left(\frac{\partial(1/p)}{\partial q}\frac{\partial q}{\partial p} - \frac{\partial(1/p)}{\partial p}\frac{\partial q}{\partial q}\right)p^2 \qquad \because\ [f(p),\ g(p)] = 0 \\ &= 0 - p^2(-\frac{1}{p^2}.1) \\ &= 1\end{aligned}$$

Thus the transformation is Canonical.

9.3 Summary

A transformation from (q,p) to (Q,P) which preserves the canonical form of the equations of motion is known as canonical transformation, provided that the conditions apply to all Hamiltonian systems. Here we have discussed a specific procedure for transforming one set into another set of canonical variables which may be more convenient. If a problem has been formulated in the form of Hamilton's canonical equations, the canonical transformations can be used to cast these equations into a more easily soluble form *i.e.*, to make integration of the equations of motion simpler.

Canonical transformations are the transformation of phase space. They are characterised by the property that they leave the form of Hamilton's equations of motion invariant.

Further, it has been established that a canonical tranformation of canonically conjugate variables does preserve the form of the Poisson Brackets. This means, Poisson brackets can be used to check if a given transformation of the set of variables are canonical or not. Thus,

The fundamental Poisson Brackets provide the most convenient way to decide whether a given transformation is Canonical.

The Canonical invariance of Poisson Brackets implies that the equations expressed in terms of Poisson brackets are invariant under Canonical transformation. Therefore we can develop a structure of classical mechanics paralleling the Hamiltonian formulation, expressed solely in terms of Poisson brackets. Which is especially useful for transition from classical mechanics to quantum mechanics.

Self Study Questions:

1. What is a generating function? Deduce the expression for the transformed generalised coordinates and generalised momenta for four different types of generating functions.

2. Deduce the conditions for a trnasformation to be canonical.

3. Find the values of α and β for which the transformation

$$Q = q^{\alpha} \cos \beta p, \qquad P = q^{\alpha} \sin \beta p$$

   is canonical. Find the gnerating function F_3 correponding to the system.

4. Verify the following properties of the Poisson brackets:

   (i) $\dfrac{dg}{dt} = [g, H] + \dfrac{\partial g}{\partial t}$ (ii) $\dot{q}_j = [q_j, H], \quad \dot{p}_j = [p_j, H]$

   (iii) $[t, H] = 1,$ (iv) $[p_k, p_j] = 0, \quad [q_k, q_j] = 0$

   (v) $[p_k, p_j] = \delta_{kj}$

5. Find the Poisson bracket of

$$\alpha p^2 + 2\beta pq + \gamma q^2$$

with the Hamiltonian H given by

$$H = ap^2 + bq^2 + cp + dq + e$$

where α, β, γ, a, b, c, d, e are all constants.

6. By direct calculation, show that the Poisson brackets are invariant under canonical transformation.

7. Prove that the Poisson bracket of two constant of motion is itself a constant of motion, even when the constants depend upon time explicitly.

UNIT 10

Hamilton-Jacobi Theory

Preparatory inputs to this unit

1. Hamilton's Canonical equations of motion
2. Canonical Transformation and Poisson Bracket
3. Basics of Ordinary and Partial differential equations.
4. Central force problem and Keplar's laws.

10.1 Hamilton-Jacobi Theory

The basic programme of the Hamilton-Jacobi theory is to extend the ideas of the Canonical transformations further so that all the new position and momentum coordinates become constants. The development of the theory is further motivated by the idea to seek a canonical trnasformation from the corrdinates and momenta $(q,\ p)$ at time t to a new set chosen as the initial values of the coordinates and momenta $(q_0,\ p_0)$ at time $t = 0$, which are in fact constant quantities. Essentially this is to seek for evolution equations in the canonical variables at different moments of time which are to be obtained as a sequence of infinitesimal canonical transformations, also known as the infinitesimal contact transformation from the initial set $(q_0,\ p_0)$. Such a programme invariably demands a trnasformation of the type

$$q = q(q_0, p_0, t),$$
$$p = p(q_0, p_0, t)$$

To grasp the ideas involving such a transformation, we start with discussion of the infinitesimal contact transformation.

10.1.1 Infinitesimal contact transformation

The transformations in which the new set of coordinates $(Q_k,\ P_k)$ differ from the old set $(q_k, ; p_k)$ by infinitesimals or by very small amounts, are called *infinitesimal contact transformations.* In other words the relation between the new coordinates and the old coordinates is given as

$$Q_k = q_k + \delta q_k$$
$$P_k = p_k + \delta p_k$$

For any *identity transformation* we have $Q_k = q_k$ and $P_k = p_k$. It is first assumed that the generating function for infinitesimal contact transformation can be constructed from that for the identity transformations with an infinitesimal change in the canonical variables. For simplicity let us consider the example we have used earlier in the case of identity transformation where $F = \sum_k q_k P_k$ and hence $F \equiv F_2(q_k,\ P_k, t)$. The generating function giving an infinitesimal change in the canonical variables can be written as

$$F_2 = \sum_k q_k P_k + \epsilon G(q_k, P_k)$$

where ϵ is an *infinitesima* parameter with $G(q_k, P_k)$ as an arbitrary function. From this generating function, we can solve for the required transformation relations which we obtain as

$$p_k = \frac{\partial F_2}{\partial q_k} = P_k + \epsilon \frac{\partial G}{\partial q_k}$$

$$Q_k = \frac{\partial F_2}{\partial P_k} = q_k + \epsilon \frac{\partial G}{\partial P_k}$$

and

$$H' = H.$$

Hence,

$$\begin{aligned} Q_k - q_k &= \delta q_k = \epsilon \frac{\partial G}{\partial P_k} \\ P_k - p_k &= \delta p_k = -\epsilon \frac{\partial G}{\partial q_k} \end{aligned}$$

Since the difference $(P_k - p_k)$ is infinitesimal, we can replace P_k by p_k in the derivative in $G(q_k, P_k)$ to write as $G(q_k, p_k)$. So the above equations becomes

$$\delta q_k = \epsilon \frac{\partial G}{\partial p_k} \tag{10.1.1a}$$

and
$$\delta p_k = -\epsilon \frac{\partial G}{\partial q_k} \tag{10.1.1b}$$

Thus in case of *infinitesimal transformations*, the transformation relations are transformed to the function G instead of the original generating function F. Thus G becomes the new generating function which generates the **infinitesimal** contact transformation.

To illustrate the point, we consider an infinitesimal canonical transformation with the hamiltonian $H(q,\ p)$ as the generating function G and an infinitesimal interval of time dt as the infinitesimal parameter ϵ in the above formalism. The corresponding δ-changes in the canonical variables can then be calculated as

$$\delta q_i = dt \frac{\partial H}{\partial p_i} = dt\ \dot{q}_i = dq_i$$

and
$$\delta p_i = -dt \frac{\partial H}{\partial q_i} = dt\ \dot{p}_i = dp_i$$

These equations suggest that the infinitesimal canonical transformation causes to change the canonical variables, *i.e.*, the generalised coordinates and canonically conjugate momenta, defined at time t to the new values that are differed by infinitesimal amount defined at time $t + dt$. What is meant by the statement is that the motion of a system during an interval of time dt can be described by the infinitesimal cononical transformation generated by the Hamiltonian H. Further, the fact that a successive canonical transformations equivanlently produce a single canonical transformation, the canonical variables of a system at time t can be thought to have evolved from those at the initial time t_0 as a sequence of successive infinitesimal transformations with the corresponding hamiltonian at every time interval as the associated generator.

10.1.2 Infinitesimal Transformation and Poisson Bracket

Let us consider a infinitesimal canonical transformation to canonical variables $(q,\ p)$ of a dynamical system, the transformation generated with the generator $G(q,\ p)$. Any dynamical variable, say $u(q,\ p,\ t)$ associated with the system will then also change accordingly, such that

$$\delta u = u(q_i + \delta q_i,\ p_i + \delta p_i,\ t + \delta t) - u(q_i,\ p_i,\ t) = \sum_i \left(\frac{\partial u}{\partial q_i} \delta q_i + \frac{\partial u}{\partial p_i} \delta p_i \right) + \frac{\partial u}{\partial t}$$

The values of the changes in canonical variables given in equations (10.1.1a) and (10.1.1b) are substituted here, which yields

$$\delta u = \epsilon \sum_i \left(\frac{\partial u}{\partial q_i}\frac{\partial G}{\partial p_i} - \frac{\partial u}{\partial p_i}\frac{\partial G}{\partial q_i} \right) + \frac{\partial u}{\partial t}$$
$$= \epsilon[u,\ G] + \frac{\partial u}{\partial t}$$

For $u(q,\ p,\ t) = H(q,\ p,\ t)$, the hamiltonian, we can find the change in the hamiltonian as

$$\delta H = \epsilon[H,\ G] + \frac{\partial H}{\partial t}$$

If the Hamiltonian is explicitly independent of time, the change in the Hamiltonian out of the infinitesimal transformation then reduces to

$$\delta H = \epsilon[H,\ G]$$

From this expression it is clear that if the generator $G(q,\ p)$ is a constant of motion, then it poisson bracket with the hamiltonian $[H,\ G]$ vanishes, giving $\delta H = 0$. This result may be summarised in the form of the statement that *a constant serves to generate an infinitesimal canonical transformation without affecting the Hamiltonian.*

Translational symmetry and conservation of momentum

Suppose that coordinate q_i is cyclic. Then, the Hamiltonian will be independent of q_i. It will moreover be invariant under an infinitesimal canonical transformation which involves a displacement in q_i alone. Then, the transformation equation is given by

$$\delta q_i = \epsilon \delta_{ij}$$

$$\delta p_i = 0$$

where ϵ signifies an infinitesimal displacement in q_i The second statement follows from the fact that canonical momentum corresponding to a cyclic coordinate is a conserved quantity. The generating function corresponding to infinitesimal transformation and satisfying the equations is given as $G = p_i$. Therefore, we can conclude that the generating function G must be constant of the motion, since infinitesimal canonical transformation renders the Hamiltonian invariant, *viz.*, $[H,\ G] = 0$.

Rotational Symmetry and Conservation of Angular Momentum

Let us consider that the infinitesimal canonical transformation of canonically conjugate variables produces an infinitesimal rotation of $d\theta$ in the system. If We consider the cartesian coordinate system to describe the system. The rotation is assumed to be about the z axis such that the new coordinates and momenta upon the $d\theta$-rotation are given by

$$\begin{aligned} X_i &= x_i - y_i d\theta\ ; \qquad & P_{ix} &= p_{ix} - p_{iy} d\theta \\ Y_i &= y_i + x_i d\theta\ ; & P_{iy} &= p_{iy} + p_{ix} d\theta \\ Z_i &= z_i\ ; & P_{iz} &= p_{iz} \end{aligned}$$

Thus the infinitesimal changes in x and y of the coordinates will be written as

$$\delta x_i = -y_i d\theta, \qquad \delta y_i = x_i d\theta, \qquad \delta z_i = 0$$

Similarly the equations involving changes in the momentum components are

$$\delta p_{ix} = -p_{iy} d\theta, \qquad \delta p_{iy} = p_{ix} d\theta, \qquad \delta p_{iz} = 0.$$

Then the generating function G which will yield the above equation with the use of equations (10.1.1a), (10.1.1b) is

$$G = \sum_i (x_i p_{iy} - y_i p_{ix})$$

Here the role of the infinitesimal parameterϵ is played by the infinitesimal rotation $d\theta$. Further, in view of the fact that the rotation is about the z-axis, the expression for G can also be written as

$$G = L_z = \vec{L}.\hat{e}$$

where $\vec{L}$ and $\hat{e}$ are the angular momentum vector and the unit vector along the direction of the infinitesimal rotation vector.

The generating function G can be used for computing the individual values of δx_i, δy_i, δp_{ix}, δp_{iy}. through the equations (10.1.1a) and (10.1.1b).

By using the properties of Poisson Bracket we can verify the relation

$$[L_x, L_y] = L_z$$

If L_x and L_y happens to be the constants of motion, then L_z is also destined to be a constant of motion. Thus, if any two components of the angular momentum are constants of motion, the total angular momentum is a conserved quantity.

10.2 Hamilton-Jacobi Equations

It has already been discussed that Canonical transformations can be handy to transform the Hamiltonian's canonical equations of motion into simpler forms. The modus operandi towards achieving the simpler form consists of two methods normally used to solve the mechanical problems. If the Hamiltonian is conserved, then the solution of mechanical problem is determined by transferring it to new canonical coordinates. Because in this case, all generalized coordinates become cyclic and hence the integration of new coordinates become easy.

An alternative approach is to find a canonical transformation from the old coordinates q and old momenta p at time t to a new set of constants quantities, which may be $2n$ initial values at $t = 0$. With such a transformation, the equations of transformation, between old and new co ordinates and momenta, are exactly the solution of the mechanical problems as follows:

$$q = (q_0, p_0, t), \qquad p = p(p_0, q_0, t).$$

The advantage of this type of transformation is that in the process of finding the transformation equations we arrive at the solution also. Here, the coordinates and momenta are expressed as functions of their initial values q_0, p_0 and time. The concept of this process was first suggested by Jacobi.

10.2.1 Deduction of Hamilton-Jacobi Equation

The basic purpose is to reduce the canonical variables of a system to constants via canonical transformation with the objective that the transformed quantities are constants. To this effect, we seek a canonical transformation from the old set of coordinates $(q_j,\ p_j,\ t)$, to a new set $(Q_j,\ P_j,\ t)$, such that

$$\dot{Q}_j = \frac{\partial K}{\partial P_j}, \qquad \dot{P}_j = -\frac{\partial K}{\partial Q_j}$$
$$\text{and} \qquad K = H + \frac{\partial F}{\partial t} \tag{10.2.1}$$

where K, H and F are transformed Hamiltonian, old Hamiltonian and associated generating function respectively.

Our objective is to choose the canonical transformation such that the new coordinate and new momenta both are the constants of the motion. For this purpose, we invoke some special conditions to the problem: first we choose the generating function F to be second type such that $F \equiv F_2(q_j,\ P_j,\ t)$, and further demand that the new Hamiltonian K reduces to a constant, zero in particular. The Hamilton's equations of motion under these conditions will become

$$\dot{Q}_j = \frac{\partial K}{\partial P_j} = 0, \qquad \dot{P}_j = -\frac{\partial K}{\partial Q_j} = 0$$

i.e., the new generalised coordinates and momenta are constants of motion. The new generalised momenta $P_1, P_2, \ldots, P_n$ are usually denoted by the constants $\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n$, *i.e.*, $P_j = \alpha_j$. Thus equation (10.2.1) becomes

$$K(Q_1,\ Q_2,\ \ldots,\ Q_n;\ P_1,\ P_2,\ \ldots,\ P_n,\ t) = H(q_1,\ q_2,\ \ldots,\ q_n;\ p_1,\ p_2,\ \ldots,\ p_n,\ t) + \frac{\partial F_2}{\partial t} = 0$$

which implies

$$H(q_1, q_2, q_3, \ldots, q_n;\ p_1, p_2, \ldots, p_n, t) + \frac{\partial F_2}{\partial t} = 0 \tag{10.2.2}$$

The generating function F_2 is a function of $q_j,\ P_j,\ t$ and it should satisfy the equations

$$p_j = \frac{\partial F_2}{\partial q_j}, \qquad Q_j = \frac{\partial F_2}{\partial P_j} \tag{10.2.3}$$

Substituting these values to (10.2.2), reduces the latter to

$$H\left(q_1, q_2, \ldots, q_n, \frac{\partial F_2}{\partial q_1}, \frac{\partial F_2}{\partial q_2}, \ldots, \frac{\partial F_2}{\partial q_n}, t\right) + \frac{\partial F_2}{\partial t} = 0 \tag{10.2.4}$$

The equation (10.2.4) is a partial differential equation in $(n+1)$ variables while equation (10.2.2) is of $(2n+1)$ variables. This shows that the above substitutions have reduced the number of variables by n. The equation (10.2.4) is called *Hamilton-Jacobi equation.*

The solution of this equation is called *Hamilton's principal function*, denoted by S. Thus the Hamilton-Jacobi equation can be rewritten as

$$H\left(q_1, q_2, \ldots, q_n, \frac{\partial S}{\partial q_1}, \frac{\partial S}{\partial q_2}, \ldots, \frac{\partial S}{\partial q_n}, t\right) + \frac{\partial S}{\partial t} = 0$$

This equation can be written with generic notation as

$$\frac{\partial S}{\partial t} + H\left(q_j, \frac{\partial S}{\partial q_j}\right) = 0$$

or simply as
$$H + \frac{\partial S}{\partial t} = 0 \tag{10.2.5}$$

We can now see that by the design of the equation, the new generalised coordinates Q_j's are also constants, generally denoted by $\beta_1, \beta_2, \ldots, \beta_n$. If $S(q_j, \alpha_j, t)$ is known, then one can determine Q_j from (10.2.3), *i.e.*,

$$Q_j = \beta_j = \frac{\partial S}{\partial \alpha_j}, \qquad j = 1, 2, \ldots, n. \tag{10.2.6}$$

These n relations can be inverted to find the old generalised coordinates q_j's as the functions of constants α_j and β_j as

$$q_j = q_j(\alpha, \beta, t). \tag{10.2.7}$$

Further, the expression for p_j in (10.2.3) can be used to write the old set of canonical momenta as $p_j = \frac{\partial S}{\partial q_j}$, expressed in terms of the constants α, β and t as

$$p_j = p_j(\alpha, \beta, t) \tag{10.2.8}$$

Equations (10.2.7) and (10.2.8) are the solutions of the equations of motion of the system in terms of the original set of canonical variables q_j and p_j, *i.e.*, solution to the original problem follows.

10.2.2 Physical Significance of S

The Hamilton's principal function S corresponding to a given dynamical system is a function of the old generalised coordinates q's, the new generalised constant momenta α's and the time t, *i.e.*, $S \equiv S(q, \alpha, t)$. We calculate the total time derivative of S :

$$\begin{aligned}
\frac{dS}{dt} &= \sum_j \frac{\partial S}{\partial q_j}\dot{q}_j + \frac{\partial S}{\partial t} \\
&= \sum_j p_j \dot{q}_j + \frac{\partial S}{\partial t} \\
&= \sum_j p_j \dot{q}_j - H, \qquad \because \quad H + \frac{\partial S}{\partial t} = 0 \\
\frac{dS}{dt} &= L
\end{aligned}$$

Thus,

$$S = \int L dt + \text{Constant}$$

where $L = \sum_j p_j \dot{q}_j - H$ is the Lagrangian of the system. Thus we see that up to an additive constant, the Hamilton's principal function S turns out to be identical with the Hamilton's action, *i.e.*, the indefinite time integral of the Lagrangian.

10.2.3 Separation of variables in Hamilton-Jacobi Equation: Hamilton's Characteristic Function

We have already seen that for a conservative system, the hamiltonian is explicitly independent of time, and at the same time a constant of motion, designating the constant total energy of the system. For such a system the Hamilton-Jacobi equation takes a particular form and amenable to an additive separation of variables, paving the way for directly identifying the constants of motion, even without a complete solution.

With the explicit time independence in the hamiltonian, the Hamilton-Jacobi equation can be written in terms of the Hamilton's principal function S in the following form:

$$\frac{\partial S}{\partial t} + H\left(q_j; \frac{\partial S}{\partial q_j}\right) = 0$$

As we can see, the first term involves the explicit time dependence and the second term depends only on q_j the coordinates, time does not appear explicitly. For such a system the solution can be written as a sum of two terms- one explicitly depending on time while the other explicitly on q_j :

$$S(q_j, \alpha_j, t) = W(q_j; \alpha_j) - \alpha_1 t$$

with α_1 as a constant. Such that the explicit time independent part can be written as

$$H\left(q_j, \frac{\partial W}{\partial q_j}\right) = \alpha_1$$

identifying the constant α_1 to be the total energy E of the system.. Thus the Hamilton-Jacobi equation assumes a simpler form without involving the time in it. The role of the Hamilton's principal function S is now played by an explicit time independent quantity $W(q_j; \alpha_j)$, and this is called *Hamilton's characteristic function.*

It is very interesting to note that the canonical transformation generated by the function W alone is such that the new Hamiltonian will be cyclic in all the new coordinates. This can be seen from the following: Let us consider a canonical transformation in which the new momenta are all constants of the motion, α_j and consider α_1, in particular, to be the associated constant Hamiltonian. This gives

$$H(q_j, p_j) = \alpha_1 \ .$$

If the generating function of this canonical transformation be given by $W(q, P)$, the transformation equations must follow

$$p_j = \frac{\partial W}{\partial q_j}; \qquad Q_j = \frac{\partial W}{\partial P_j} = \frac{\partial W}{\partial \alpha_j}, \tag{10.2.9}$$

Also, $H(q_j; p_j) = \alpha_1$

Therefore, in view of the equation (10.2.9), we can write this as

$$H\left(q_j, \frac{\partial W}{\partial q_j}\right) = \alpha_1 \tag{10.2.10}$$

which is in the form of the Hamilton-Jacobi equation, with W as the generator. Now in such a transformation we can write the new Hamiltonian K as

$$\begin{aligned} K &= H + \frac{\partial W}{\partial t} \\ &= \alpha_1 \qquad \because \quad W \text{ is explicitly independent of time.} \end{aligned}$$

i.e., the new hamiltonian K is same as the old hamiltonian which is a constant.

We thus observe that the new hamiltonian $K = \alpha_1$, being a constant is again devoid of any generalised coordinates, *i.e.*, cyclic in all the coordinates.

10.3 Action-Angle variables

There are many occasion in which one needs to discuss and analyse periodic motion, wherein the key quantity which needs attention is the frequency of motion. In order to handle such a system there exists a very powerful method based on a variation of the Hamilton-Jacobi procedure discussed above. This technique consists in considering general constants J_i as functions of the integration constants α_i's appearing directly from the solution of the Hamilton-Jacobi equation. These J_i's obtained for a system of n degrees of freedom, essentially form a set of n independent functions of the α_i's. These functions are called the *action variables.*

To develop the ideas gradually, we first consider a conservative system with one degree of freedom, where the conserved Hamiltonian is explicitly independent of time so that it can be written as

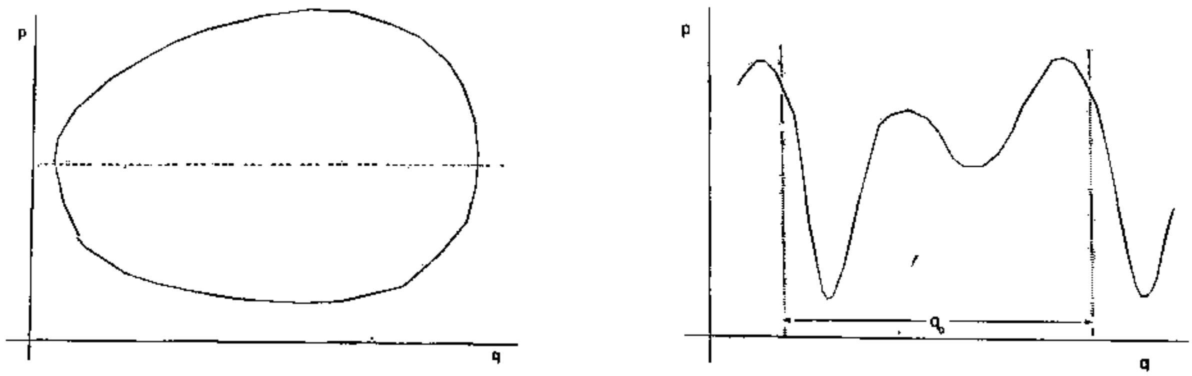
$$H(q,\ p) = \alpha_1.$$

This can be soloved for the momentum of the system as

$$p = p(q,\ \alpha_1), \tag{10.3.1}$$

This can be interpreted as the equation of the orbit traced out by the system in the two-dimensional phase space $(q,\ p)$ with constant value α_1 of the hamiltonian. The orbit so found in the phase space actually characterises the periodic motion of the system. Here we can see that there are two types of periodic motions: *closed orbit periodic motion* and *open path variation* with a periodicity of p in q.

1. The first type is characterised by the closed orbit in the phase space (10.1a). Here the system retraces a closed path periodically: both q and p becomes the periodic functions of the time with identical frequency. This type of situations arise in the system where the motion causes the kinetic energy function to vary periodically from zero to a certain maximum and back to zero again. This type of phenomena is known as *libration* in the astronomical jargon.



(a) Libration (b) Libration

Figure 10.1: Libration and Rotation

2. In the second type of periodic motion, the traced path maintains a periodicity of p in q in the phase space (10.1b). We can find such situtions in the case of rigid body rotations about certian axes, with the angle of rotation playing the role of the generalised coordinate q, where the increase in the value of q by 2π does not essentially change the state of the system. In this type of periodic motion it is not essential that the value of q be bounded between two fixed values: the value can increase indefinitely.

It is important to note that a single physical system may experience both type of periodic motions outlined as above. We can look at the case of simple pendulum where q serves to be the angle of deflection θ. With the length of the pendulum l and the zero reference of the potential energy fixed at the point of suspension, the energy of the system, a constant, is given by

$$E = \frac{p_\theta^2}{2ml} - mgl\cos\theta,$$

which can be solved for p_θ as a function of theta as

$$p_\theta = \pm\sqrt{2ml^2(E + mgl\cos\theta)}.$$

This solution traces out a path the system traverses in the two dimensional phase space of $(\theta,\ p_\theta)$.

A few cases may be considered from the solution:

Case I: $E < mgl$.
In this case the physical system can execute motion only if the absolute value of the angle, $|\theta|$ is less than a given bound, θ', where θ' is defined by

$$\cos\theta' = -\frac{E}{mgl}.$$

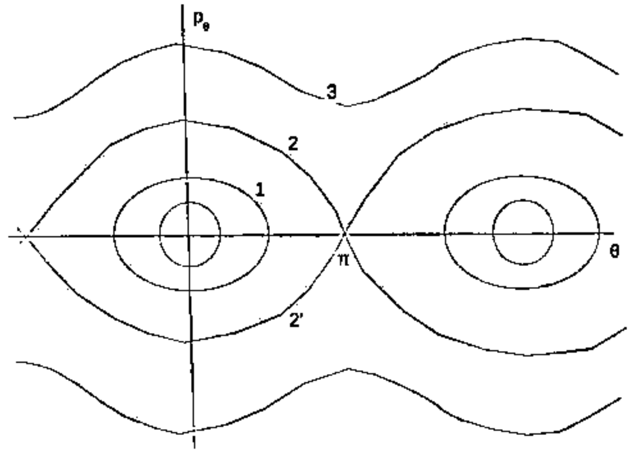


Figure 10.2: The Phase space orbits for the simple pendulum.

Under this circumstance, the pendulum executes its motion with θ-values between $-\theta'$ and θ', constituting the case of a periodic motion of the libration type. The system then traverses a trajectory in the phase space in the form of a curve type 1 as in the figure (10.2).

Case II: $E > mgl$.
In this case one finds a physical motion for all values of θ which can increase without limit to produce a rotation-type periodic motion. In the simple pendulum case, this case is chacterised by a higher value of energy so as to swing through the vertical position $\theta = \pi$ and thus continued rotation. Curve 3 in the figure (10.2) cooresponds to this type of motion.

Case III: The limiting case of $E = mgl$.
This case corresponds to an energy just sufficient for the pendulum to reach the limiting value of $\theta = \pi$ with zero values of both the angular momentum and kinetic energy. As illustrated by curve 2 or $2'$, in the figure (10.2), this is the configuration with ustable equilibrium and in principle retain this configuration for indefinite period of time. But any perturbation, be the slightest, in the system could the latter to shift to either type curve 2 or $2'$, in either way causing the system to shed off its energy.

The physics of the either type of the periodic motion, can be best dealt with the introduction of a new variable J which can replace the constant momentum after canonical transformation. This variable is known as the *action variable*, which is defined as

$$J = \oint p\, dq.$$

Depending on the type, the integration over the closed curve here is to be carried out over the complete period of libration or the rotation. Now we can see from equation (10.3.1) that J will be a function of α_1 such that we can write the hamiltonian as a function of J, *i.e.*,

$$\alpha_1 \equiv H = H(J)$$

and therefore the Hamilton's characteristic function is expressible as

$$W = W(q, J).$$

What we have understood from the above analysis is that essentially the momentum function of the problem has been generalised to J. The corresponding generalised coordinate, *i.e.*, the generalised coordinate conjugate to J is termed as the *angle variable* w, is accordingly defined by the transformation equation

$$w = \frac{\partial W}{\partial J}.$$

The corresponding equation of motion for w can then be written as

$$\dot{w} = \frac{\partial H(J)}{\partial J} = \mu(J),$$

where μ is a constant function, exclusively dependent on J. The solution of this equation can be found in a straightforward manner as

$$w = \mu t + \beta, \tag{10.3.2}$$

so that w is a linear function of time.

Physical interpretation of μ

In order to provide a physical interpretation of the function μ, we consider a system to undergo a complete cycle of libration or the rotation, in q so that the change in w is found as

$$\Delta w = \oint \frac{\partial w}{\partial q} dq$$

But as we have $w = \frac{\partial W}{\partial J}$, the above integral can be rewritten as

$$\Delta w = \oint \frac{\partial^2 W}{\partial q \partial J} dq.$$

In view of the constancy of J, the derivative with respect to J can be taken outside the integral so that

$$\Delta w = \frac{d}{dJ}\oint \frac{\partial W}{\partial q}dq = \frac{d}{dJ}\oint pdq = 1 \tag{10.3.3}$$

Equation (10.3.3) clears states that as q goes through a complete cycle of the libration or rotation w changes by unity. Now from equation ((10.3.2)) we can find $\Delta w = \mu\Delta t = \mu\tau$ for the time period of the cycle as τ. We can then write

$$\Delta w = \mu\Delta t = \mu\tau = 1, \qquad \Longrightarrow \qquad \mu = \frac{1}{\tau}$$

This means that the constant value of μ is the reciprocal of the time period for a complete cycle or the frequency of the periodic motion of q for the libration or rotation.

Thus the use of action-angle variables can provide us with the technique of obtaining the frequency of periodic motion without requiring to find the complete solution of the problem. For a periodic system with one degree of freedom, the frequency can directly be evaluated onces the hamiltonian is expressed as a function of J and finding the derivative of H with respect to J: the derivative is the frequency μ of the motion in q.

10.4 The Keplar problem of planetary orbits

The Keplar problem, which has been attracting the attention of scientists from centuries towards finding its solution from various angles. The same problem can also be discussed as an ideal problem for finding solution by use of Hamilton Jacobi equations. The problem, essentially consists in determining the orbit of a planet under the steady, inverse square gravitational force of attraction due to the sun. The problem thus falls in the category of the time independent central force; the associated hamiltonian being time independent and hence recognised as the total energy of the planet-sun system.

Let us consider a planet of reduced mass μ moving round the stationary sun in a given plane. Clearly the degree of freedom of the planet is 2. We consder the two dimentsional polar coordinate system $(r,\ \theta)$ to describe the motion of the system. If the radial and tangential components of the momentum are p_r and p_θ respectively, the hamiltonian of the system is given by

$$H = \frac{1}{2\mu}\left(p_r^2 + \frac{p_\theta^2}{r^2}\right) - \frac{k}{r}$$

As the system is conservative, the Hamiltonian equals the total energy $\alpha_1 = E$ of the system, *i.e.*,

$$H = \alpha_1 = E.$$

Since the Hamiltonian is explicitly independent of the time t, the solution of Hamilton Jacobi equation can be obtained through separation of variables, and the Hamilton's characteristic function W, the spatial part of the Hamilton's principal function S of the system, is involved in the solution as the generating function for the transformation. Let us assume the function S to be of the form

$$S(q_j, \alpha_j, t) = W(q_j, \alpha_j) - \alpha_1 t$$

Then,

$$p_i = \frac{\partial S}{\partial q_i} = \frac{\partial W}{\partial q_i}$$

which gives the radial and tangential components of the momentum as

$$p_r = \frac{\partial S}{\partial r} = \frac{\partial W}{\partial r}$$

and in view of the fact that in a central force motion , the agnular momentuam is a conserved quantity, we can write

$$p_\theta = \frac{\partial S}{\partial \theta} = \frac{\partial W}{\partial \theta} = \alpha_2 = \text{a constant}$$

Thus the expression

$$p_r^2 + \frac{p_\theta^2}{r^2} = 2\mu E + \frac{2\mu k}{r}$$

facilitates to write down the Hamilton-Jacobi equation for planetary orbits in plane polar coordinates as

$$\left(\frac{\partial W}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial W}{\partial \theta}\right)^2 = 2\mu E + \frac{2\mu k}{r} \tag{10.4.1}$$

Since this differential equation is cyclic in θ, the solution can be written as

$$W(r, \theta) = p_\theta \theta + W_1(r).$$

Substituting this to equation (10.4.1), we get

$$\left(\frac{dW_1}{dr}\right)^2 + \frac{p_\theta^2}{r^2} = 2\mu\left(E + \frac{k}{r}\right)$$

or,
$$W_1(r) = \int \left[2\mu\left(E + \frac{k}{r}\right) - \frac{p_\theta^2}{r^2}\right]^{\frac{1}{2}} dr,$$

Hence

$$S(r,\theta,p_\theta,E,t) = -Et + p_\theta\theta + \int \left[2\mu\left(E+\frac{k}{r}\right) - \frac{p_\theta^2}{r^2}\right]^{\frac{1}{2}} dr + A. \tag{10.4.2}$$

where, A is the constant of integration. Going further, we have

$$Q_1 = \frac{\partial W}{\partial \alpha_1} = t + \beta_1,$$

and $$\frac{\partial W}{\partial \alpha_i} = \beta_i \quad \text{when} \quad i \neq 1$$

$$\Longrightarrow \qquad \frac{\partial W}{\partial \alpha_2} = \beta_2$$

Using equation (10.4.2), β_2 can be expressed as

$$\beta_2 = \frac{\partial W}{\partial \alpha_2} = \theta - \int \frac{\alpha_2 dr}{r^2\sqrt{\left[2\mu\left(E+\frac{k}{r}\right) - \frac{\alpha_2^2}{r^2}\right]}} = \theta + I \quad \text{(say,)} \tag{10.4.3}$$

The integral I can be evaluated by changing the variable r as $u = \frac{1}{r}$, so that $du = -\frac{dr}{r^2}$. Thus

$$I = -\int \frac{\alpha_2 dr}{r^2\sqrt{\left[2\mu\left(E+\frac{k}{r}\right) - \frac{\alpha_2^2}{r^2}\right]}}$$

$$= \int \frac{\alpha_2 du}{\sqrt{\left[2\mu E + 2\mu k u - \alpha_2^2 u^2\right]}}$$

Now, in order to evaluate the integral, we consult the standard handbook results:

Given $R = a + bx + cx^2$

$$\int \frac{dx}{\sqrt{R}} = \frac{1}{\sqrt{c}} \ln(2\sqrt{cR} + 2cx + b), \quad c > 0.$$
$$= \frac{1}{\sqrt{c}} \sinh^{-1}\left(\frac{2cx+b}{\sqrt{\Delta}}\right); \quad c > 0, \quad \Delta > 0; \quad \Delta = 4ac - b^2$$
$$= \frac{\mathbf{-1}}{\sqrt{\mathbf{-c}}} \sin^{-1}\left(\frac{\mathbf{2cx+b}}{\sqrt{\mathbf{-\Delta}}}\right); \quad \mathbf{c < 0}, \quad \mathbf{\Delta < 0}$$
$$= \frac{1}{\sqrt{c}} \ln(2cx + b); \quad c > 0, \quad \Delta = 0$$

The result displayed in the boldface inside the box is the appropriate one to evaluate the integral, finally yielding,

$$I = \sin^{-1}\left(\frac{\mu k - \alpha_2^2 u}{\sqrt{2\mu E \alpha_2^2 + \mu^2 k^2}}\right)$$

so that from the integral expression for β_2 in (10.4.3), u can be written as

$$u = \frac{\mu k}{\alpha_2^2}\left[1 + \left(1 + \frac{2E\alpha_2^2}{\mu k^2}\right)^{\frac{1}{2}} \sin(\beta_2 - \theta)\right]$$

Rewriting $\beta_2' = \beta_2 + \frac{\pi}{2}$ and substituting $\frac{1}{r}$ for u we have

$$\frac{1}{r} = \frac{\mu k}{\alpha_2^2}\left[1 + \left(1 + \frac{2E\alpha_2^2}{\mu k^2}\right)^{\frac{1}{2}} \cos(\theta - \beta_2')\right]$$

which is in the form of the polar equation for a conic

$$\frac{1}{r} = \frac{1}{\epsilon l}\left(1 - \epsilon \cos(\theta - \beta_2')\right)$$

with the eccentricity

$$\epsilon = \sqrt{\left(1 + \frac{2E\alpha_2^2}{\mu k^2}\right)}, \qquad \epsilon l = \frac{\alpha_2^2}{\mu k}$$

Thus we have found that the orbit of the planet is described by a conic with a defined eccentricity, which in turn is decided by the value of the energy E, *i.e.*, which decides the nature of the conic, *viz.*, the trajectory of planetary motion is described by a parabola, a hyperbola or an ellipse:

If	$E < 0,$	$\epsilon < 1,$	the planet moves in an elliptic path.
	$E = 0,$	$\epsilon = 1,$	the path of the motion is parabolic.
	$E > 0,$	$\epsilon > 1,$	planetary motion is described by a hyperbolic path.

The Keplar problem actually involves an elliptical path signifying that the total energy of the system E is negative, which means that the system is bound. We can evaluate this value by finding the length of the semi major axis of the given elliptic orbit by considering the polar equation of the ellipse, *i.e.*,

$$\frac{1}{r} = \frac{1}{p}(1 + \epsilon \cos\theta),$$

which relates to the semi major axis a of the ellipse as

$$a = \frac{p}{1 - \epsilon^2}$$

Comparing this with the elliptical path of the planetary path as above, we have $p = \epsilon l$.

$$i.e., \qquad 2a = \frac{2\epsilon l}{1 - \epsilon^2}$$

Further with $\epsilon l = \dfrac{\alpha_2^2}{\mu k}$ and the expression of the eccentricity ϵ as found above, we find that the semi major axis a and the total energy E of the system are related by

$$E = -\frac{k}{2a}$$

The same result was also found while discussing the particle motion under central force earlier.

Example 10.4.1 ***The Harmonic Oscillator problem:*** *Work out the details of the simple harmonic oscillator by the Hamilton Jacobi theory.*

Solution: We consider the simple harmonic oscillator problem in one-dimension. The Hamiltonian of the system has been worked out in multiple occasions earlier, as

$$H = \frac{1}{2m}(p^2 + m^2\omega^2 q^2) \equiv E,$$

where $\omega = \sqrt{\frac{k}{m}}$; k being the force constant.

The Hamilton-Jocobi equation can be written for the principal function S by setting the momentum p equal to $\frac{\partial S}{\partial q}$ and then substituting in the Hamiltonian to the requirement that the new Hamiltonian vanishes.

$$\frac{1}{2m}\left[\left(\frac{\partial S}{\partial q}\right)^2 + m^2\omega^2 q^2\right] + \frac{\partial S}{\partial t} = 0 \tag{10.4.4}$$

Since the explicit dependence of S on t is involved only in the last term, the solution of the equation (10.4.4) can be obtained through the separation of variables and thus can be written in the form

$$S(q, \alpha, t) = W(q, \alpha) - \alpha t, \tag{10.4.5}$$

where α is the constant of integration. With the choice of the form of the solution, the time variable can be separated from (10.4.4), so that we have the time independent part of the equation as

$$\frac{1}{2m}\left[\left(\frac{\partial W}{\partial q}\right)^2 + m^2\omega^2 q^2\right] = \alpha \tag{10.4.6}$$

The integration constant α is thus to be identified with the total energy E. This can also be recognized directly from (10.4.5) and the relation

$$\frac{\partial S}{\partial t} + H = 0,$$

which then reduces to $H = \alpha$

Equation (10.4.6) can be integrated immediately to

$$W = \sqrt{2m\alpha}\int dq\sqrt{1 - \frac{m\omega^2 q^2}{2\alpha}},$$

$$\text{so that} \quad S = \sqrt{2m\alpha}\int dq\sqrt{\left(1 - \frac{m\omega^2 q^2}{2\alpha}\right)} - \alpha t \tag{10.4.7}$$

Now, we have

$$\beta = \frac{\partial S}{\partial \alpha} = \sqrt{\frac{m}{2\alpha}}\int \frac{dq}{\sqrt{\left(1 - \frac{m\omega^2 q^2}{2\alpha}\right)}} - t,$$

(where β are constants) which can be integrated easily as

$$t + \beta = \frac{1}{\omega} arc\sin q\sqrt{\frac{m\omega^2}{2\alpha}} \tag{10.4.8}$$

Equation (10.4.8) can be written as

$$q = \sqrt{\frac{2\alpha}{m\omega^2}}\sin\omega(t + \beta), \tag{10.4.9}$$

which is the familiar solution of a harmonic oscillator.

Formally, the solution for the momentum can be written as

$$p = \frac{\partial S}{\partial q} = \frac{\partial W}{\partial q} = \sqrt{2m\alpha - m\omega^2 q^2}$$

So, we have the form of the momentum p as

$$p = \sqrt{2m\alpha(1 - \sin^2 \omega(t+\beta))},$$

$$i.e., \quad p = \sqrt{2m\alpha}\cos\omega(t+\beta) \tag{10.4.10}$$

The constants α and β must be connected with the initial conditions q_0 and p_0 at time $t = 0$. By squaring (10.4.9) and (10.4.10) and adding we get

$$2m\alpha = p_0^2 + m^2\omega^2 q_0^2$$

The same result follows immediately from the identification of α as the conserved total energy E. Now, the phase constant β is related to q_0 and p_0 by

$$\tan\omega\beta = m\omega\frac{q_0}{p_0}$$

Thus, Hamilton's principal function is the generator of a canonical transformation to a new coordinate that measures the phase angle of the oscillation and to a new canonical momentum identified with the total energy.

If the solution for q is substituted in (10.4.7), Hamilton's principal function can be written as

$$S = 2\alpha\int \cos^2\omega(t+\beta)dt - \alpha t = 2\alpha\int\left(\cos^2\omega(t+\beta) - \frac{1}{2}\right)dt.$$

Now, the Lagrangian is

$$\begin{aligned} L &= \frac{1}{2m}(p^2 - m^2\omega^2 q^2) \\ &= \alpha\left(\cos^2\omega(t+\beta) - \sin^2\omega(t+\beta)\right) \\ &= 2\alpha\left(\cos^2\omega(t+\beta) - \frac{1}{2}\right), \end{aligned}$$

so that S is the time integral of the Lagrangian, in agreement with the general relation

$$S = \int L dt + Constant.$$

Example 10.4.2 ***Projectile motion:*** *Using the Hamilton-Jacobi theory, analyse the problem of the projectile motion of a particle under the action of the gravity of the earth.*

Solution: We choose the cartesian coordinates to describe the problem so that the potential energy for an object of mass m in projectile motion is given by $V = mgz$. Here g is the acceleration due to gravity. The form of the kinetic energy is

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

At $t_0 = 0$, the action is given by

$$S = \int_0^t \left[\frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) - mgz\right] dt$$

The conjugate momenta are then given by

$$p_x = m\dot{x}, \quad p_y = m\dot{y}, \quad p_z = m\dot{z}.$$

The Hamiltonian is

$$H = \frac{p_x^2 + p_y^2 + p_z^2}{2m} + mgz.$$

Since x and y are cyclic and $\dfrac{\partial H}{\partial t} = 0$, the corresponding momenta p_x and p_y are conserved, and the total energy $E = H$ is also conserved.

The Hamilton-Jacobi equation for the system can now be written as

$$\left.\begin{aligned} &\frac{1}{2m}\left[\left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2 + \left(\frac{\partial S}{\partial z}\right)^2\right] + mgz = -\frac{\partial S}{\partial t} \\ &\text{with separation of variables,} \\ &S = S_x(x) + S_y(y) + S_z(z) - Et \end{aligned}\right\} \qquad (10.4.11)$$

or,

$$\frac{1}{2m}\left[\left(\frac{\partial S_x}{\partial x}\right)^2 + \left(\frac{\partial S_y}{\partial y}\right)^2 + \left(\frac{\partial S_z}{\partial z}\right)^2\right] + mgz = E$$

or,

$$\left(\frac{\partial S_x}{\partial x}\right)^2 + \left(\frac{\partial S_y}{\partial y}\right)^2 + \left(\frac{\partial S_z}{\partial z}\right)^2 + 2m^2gz = 2mE,$$

which is possible only if

$$\left(\frac{\partial S_x}{\partial x}\right)^2 = \alpha^2, \qquad \left(\frac{\partial S_y}{\partial y}\right)^2 = \beta^2,$$

where α and β are constants.

Now, we have the z-part of the equation to be of the form

$$\left(\frac{\partial S_z}{\partial z}\right)^2 + 2m^2gz = 2mE - \alpha^2 - \beta^2$$

with

$$S_x = \alpha x + c_1 \qquad \text{and} \qquad S_y = \beta y + c_2.$$

Let $\quad \gamma^2 = 2mE - \alpha^2 - \beta^2,$

so that $\quad \left(\frac{\partial S_z}{\partial z}\right) = \sqrt{\gamma^2 - 2m^2 gz}$

$$S_z = -\frac{1}{2m^2 g}\int_{z_0}^{z} \sqrt{\xi}\,d\xi, \quad \text{where,} \quad \xi = \gamma^2 - 2m^2 gz$$

$$= \frac{-1}{3m^2 g}\left[\left(\gamma^2 - 2m^2 gz\right)^{3/2} - \left(\gamma^2 - 2m^2 gz_0\right)^{3/2}\right]$$

and Hamilton's principal function is

$$S = \alpha x + \beta y - \frac{1}{3m^2 g}\left(\gamma^2 - 2m^2 gz\right)^{3/2} - Et$$

Again

$$p_z = \sqrt{\gamma^2 - 2m^2 gz} = \sqrt{2m\left(E - \frac{p_x^2}{2m} - \frac{p_y^2}{2m} - mgz\right)},$$

and the final expression shows that $H = E$, as expected. The energy may be written as

$$2mE = \alpha^2 + \beta^2 + \gamma^2,$$

so that

$$p_z = \sqrt{\alpha^2 + \beta^2 + \gamma^2 - p_x^2 - p_y^2 - 2m^2 gz}$$

and

$$S = \alpha x + \beta y - \frac{1}{3m^2 g}\left(\gamma^2 - 2m^2 gz\right)^{3/2} - \frac{1}{2m}(\alpha^2 + \beta^2 + \gamma^2)t.$$

Taking constants of integration (α, β, γ) as the new momentum variables, we have

$$q_x = \frac{\partial S}{\partial \alpha} = x - \frac{\alpha}{m}t.$$
$$q_y = \frac{\partial S}{\partial \beta} = y - \frac{\beta}{m}t.$$
$$q_z = \frac{\partial S}{\partial \gamma} = -\frac{\gamma}{m^2 g}\left(\gamma^2 - 2m^2 gz\right)^{1/2} - \frac{1}{m}\gamma t.$$

Finally, we invert these relations to find x, y, z as function of the initial conditions and time:

$$x = q_x + \frac{\alpha}{m}t,$$
$$y = q_y + \frac{\beta}{m}t.$$

and

$$\left(q_z + \frac{\gamma}{m}t\right)^2 = \frac{\gamma^2}{m^4g^2}\left(\gamma^2 - 2m^2gz\right),$$
$$= \frac{\gamma^4}{m^4g^2} - \frac{2\gamma^2}{m^2g}z,$$

or,
$$\frac{2\gamma^2}{m^2g}z = \frac{\gamma^4}{m^4g^2} - \left(q_z + \frac{1}{m}\gamma t\right)^2,$$
$$= \frac{\gamma^4}{m^4g^2} - \left(q_z^2 + 2\frac{q_z}{m}\gamma t + \frac{1}{m^2}\gamma^2t^2\right).$$

Therefore,

$$z = \frac{m^2g}{2\gamma^2}\left[\frac{\gamma^4}{m^4g^2} - q_z^2 - \frac{2}{m}q_z\gamma t - \frac{1}{m^2}\gamma^2t^2\right],$$

or,

$$z = \left(\frac{\gamma^2}{2m^2g} - \frac{m^2gq_z^2}{2\gamma^2}\right) - \frac{mgq_zt}{\gamma} - \frac{g}{2}t^2,$$

i.e.,
$$z = A + Bt + Ct^2,$$

where
$$A = \left(\frac{\gamma^2}{2m^2g} - \frac{m^2gq_z^2}{2\gamma^2}\right), \qquad B = -\frac{mgq_z}{\gamma}, \qquad \text{and} \qquad C = -\frac{g}{2}.$$

At $t = 0$, we find $z \equiv z_0 = \left(\frac{\gamma^2}{2m^2g} - \frac{m^2gq_z^2}{2\gamma^2}\right)$. the initial value of the position of the projectile.

The initial velocity of the projectile can be found by differentiating z with respect to time and evaluating this at $t = 0$, *i.e.*,

$$\dot{z}|_{t=0} \equiv \dot{z}_0 = B = \frac{mgq_z}{\gamma}$$

Example 10.4.3 *Find the frequency of a one dimensional linear harmonic oscillator by use of the action-angle variables. The hamiltonian of the linear harmonic oscillator is given as*

$$H = \frac{1}{2m}(p^2 + m^2\omega^2q^2)), \qquad \textit{where} \qquad \omega = \sqrt{\frac{k}{m}}$$

Here m is the mass, k is the force constant, q is the generalised coordinate and p is the conjugate moomentum associated with the harmonic oscillator.

Solution: The given hamiltonian is

$$H = \frac{1}{2m}(p^2 + m^2\omega^2 q^2),$$

with

$$\omega = \sqrt{\frac{k}{m}}$$

The corresponding Hamilton-Jacobi equation is

$$\frac{1}{2m}\left[\left(\frac{\partial S}{\partial q}\right)^2 + m^2\omega^2 q^2\right] + \frac{\partial S}{\partial t} = 0$$

whose solution is sought in the form of

$$S(q, \alpha, t) = W(q, \alpha) - \alpha t,$$

with α as the constant of integration, giving the total energy, a constant quantity of the system. The form of the solution is chosen to be amenable for time separation such that the part independent of time is

$$\frac{1}{2m}\left[\left(\frac{\partial W}{\partial q}\right)^2 + m^2\omega^2 q^2\right] = \alpha$$

and the conjugate momentum $p \equiv \frac{\partial W}{\partial q} = \sqrt{2m\alpha - m^2\omega^2 q^2}$.

Now we define the quantity J as

$$J = \oint p dq$$

which, with the substitution of the values of $p \equiv \frac{\partial W}{\partial q}$ above yields

$$\begin{aligned} J &= \oint \left(\sqrt{2m\alpha - m^2\omega^2 q^2}\right) dq \\ &= m\omega \oint \left(\sqrt{\frac{2\alpha}{m\omega^2} - q^2}\right) dq \end{aligned}$$

Let

$$q = \left(\sqrt{\frac{2\alpha}{m\omega^2}}\right) \sin\theta,$$

so that

$$dq = \left(\sqrt{\frac{2\alpha}{m\omega^2}}\right) \cos\theta d\theta$$

Subsituting these to the equation above,

$$\begin{aligned} J &= \frac{2\alpha}{\omega} \int_0^{2\pi} \cos^2\theta d\theta \\ &= \frac{2\pi\alpha}{\omega} \end{aligned}$$

or, solving for α, we have

$$\alpha \equiv H = \frac{J\omega}{2\pi}$$

and therefore the frequency μ is given by

$$\mu = \frac{\partial H}{\partial J} = \frac{\omega}{2\pi} = \frac{1}{2\pi}\sqrt{\frac{k}{m}}$$

10.5 Summary

In this unit, we have discussed one of the important tools in classical mechanics namely the Hamilton Jacobi equation and the resulting time dependent and time independent functions- the principal function and the characteristic functions, along with some ideas of how these equations find applications in the simple harmonic oscillator problem, projectile motions and the keplar problem. The method of separation of variables have particularly been found to be suitable in finding the solution of Hamilton-Jacobi equations. In particular the following few points have been revealed during the discussion of the topics in the unit:

1. The Hamilton-Jacobi equation is a single, first order partial differential equation for function S of n generalized coordinates $q_1, q_2, \ldots, q_n$ along with the time t.

2. The generalized momenta donot appear, except as derivatives of S.

3. The Hamilton-Jacobi equation is an equivalent expression of an integral minimization promlem as Hamilton's principle. Because of this fact, Hamilton-Jacobi equation is useful in various branches of mathematics and physics, *viz.*, optimization problems in calculus of variations and dynamical systems, quantum chaos, and determining geodesics on a Riemannian manifold, an important variational problem in Riemannian geometry.

Self Study Questions:

1. Describe Hamilton's characteristic function. Use this to obtain the dynamics of a projectile, with the initial velocity $\vec{u} = u_1\hat{i} + u_2\hat{j}$ where $\frac{u_2}{u_1} = \tan\alpha$ and $\hat{j}$ is the vertical direction.

2. What is Hamilton's principal function? Use this to describe the dynamics of a paticles freely falling under the action of earth's gravity.

3. For a paticle moving in a potential field $V = \frac{\vec{k}\cdot\vec{r}}{r^3}$, ($\vec{k}$=constant), find all the constants of motion by use of the Hamilton-Jacobi method.

4. Use Hamilton-Jacobi theory to extablish that the orbit of a planet in the solar system is elliptic with the sun at of the foci.