Course Materials for III B. Sc

Physics

Subject Name:***Theoretical Physics***

Prepared by

Dr. S. Sathiskumar
Assistant Professor
Post Graduate Department of Physics
Srinivasan College of Arts and Science
Perambalur – 621 212

Unit - I Fundamental Principles and Lagrangian Formulation

defined on the basis of Newton's law of gravitation. The mass of a body defined on the basis of gravitational properties is called the **gravitational mass**. Naturally a question arises: Is the inertial mass of a body equal to its gravitational mass? Recently it was established that these masses are equal to within a few parts in 10^{12}. This equivalence of the inertial and gravitational masses of a body is the **principle of equivalence** postulated by Einstein in general relativity.

1.4 MECHANICS OF A PARTICLE

In this section, we shall discuss mainly the conservation laws for a particle in motion in Newtonian formalism.

Conservation of Linear Momentum From Newton's first law, we have already indicated the law of conservation of momentum of a single particle in Eq. (1.7). It also follows from Newton's second law of motion which states that

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

If no external force is acting on the particle

$$\frac{d\mathbf{p}}{dt} = 0 \text{ or } \mathbf{p} = \text{constant in time} \qquad (1.13)$$

If the total force acting on a particle is zero, then the linear momentum **p** *is conserved.*

Angular Momentum and Torque Angular momentum and torque are two important quantities in rotational motion. A force causes linear acceleration whereas a torque causes angular acceleration. The *angular momentum* **of a particle about a point O (say origin), denoted by L, is defined as L = r p (1.14) where r is the radius vector of the particle. The** *torque* **(N) or moment of a force about O is**

defined as $\mathbf{N} = \mathbf{r} \times \mathbf{F} = \mathbf{r} \times \frac{d\mathbf{p}}{dt}$ (1.15)

which is perpendicular to the plane containing the vectors **r** and **F** points in the direction of the advance of a right hand screw from **r** to **F**. Since

$$\frac{d\mathbf{r}}{dt} \times \mathbf{p} = m\frac{d\mathbf{r}}{dt} \times \frac{d\mathbf{r}}{dt} = 0$$

from Eq. (1.15), we have

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

$$\mathbf{N} = \frac{d\mathbf{L}}{dt} \qquad (1.16)$$

which is the analogue of Newton's second law in rotational motion.

Conservation of Angular Momentum The angular momentum conservation comes automatically from Eq. (1.16).
If the torque N acting on the particle is zero, then

$$\frac{d\mathbf{L}}{dt} = 0 \text{ or } \mathbf{L} = \text{constant} \qquad (1.17)$$

If the torque **N** *acting on a particle is zero, the angular momentum* **L** *is a constant.* Planets moving around the sun and satellites around the earth are some of the very common examples.

Work Done by a Force Work done by an external force in moving a particle from position 1 to position 2 is given by

$$W_{12} = \int_1^2 \mathbf{F} \cdot d\mathbf{r} = \int_1^2 m\frac{d\mathbf{v}}{dt} d\mathbf{r}$$

Assuming the mass of the particle constant

$$W_{12} = m\int_1^2 \frac{d\mathbf{v}}{dt}\frac{d\mathbf{r}}{dt} dt = m\int_1^2 \frac{d\mathbf{v}}{dt}\mathbf{v}\, dt$$

$$= m\int_1^2 \mathbf{v}\,d\mathbf{v} = \frac{m}{2}\int_1^2 d(\mathbf{v}^2) = \frac{m}{2}(\mathbf{v}_2^2 - \mathbf{v}_1^2)$$

$$= T_2 - T_1 \qquad (1.18)$$

where T_2 and T_1 are the kinetic energies of the particle in positions 2 and 1 respectively. If $T_2 > T_1$, $W_{12} > 0$, work is done by the force on the particle and as a result the kinetic energy of the particle is increased. If $T_1 > T_2$, $W_{12} < 0$, work is done by the particle against the force and as a result the kinetic energy of the particle is decreased.

Conservative Force If the force acting on a system is such that the work done along a closed path is zero, then the force is said to be conservative. That is, for a conservative force F

$$\oint \mathbf{F}\cdot d\mathbf{r} = 0 \qquad (1.19)$$

If the closed curve encloses the surface **S**, by Stokes theorem, we have

$$\oint \mathbf{F}\cdot d\mathbf{r} = \int_s (\nabla \times \mathbf{F})\cdot d\mathbf{S} = 0 \qquad (1.20)$$

Since the surface is arbitrary, this is possible only if

$$\nabla \times \mathbf{F} = Curl\ \mathbf{F} = 0 \qquad (1.21)$$

which is the necessary and suficient condition for a force to be conservative. The curl of a vector is zero if it can be expressed as the gradient of a scalar function of position. Hence, we can write

$$\mathbf{F} = -\nabla V(\mathbf{r}) \qquad (1.22)$$

The scalar function $V(\mathbf{r})$ in Eq. (1.22) is called the **potential energy** of the particle at the point or simply the **potential** at the point. In terms of V, the components of the force are

$$F_x = \frac{\partial V}{\partial x} \qquad F_y = \frac{\partial V}{\partial y} \qquad F_z = \frac{\partial V}{\partial z} \qquad (1.23)$$

Conservation of Energy The work done by a force F in moving a particle of mass *m* from position 1 to position 2 is given by Eq.

(1.18). Now consider the work done W_{12} by taking F to be a conservative force derivable from a potential V. Then W_{12} takes the form

$$W_{12} = \int_1^2 \mathbf{F} \cdot d\mathbf{r} = -\int_1^2 \nabla V \cdot d\mathbf{r}$$

$$= -\int_1^2 \left(\frac{\partial V}{\partial x} dx + \frac{\partial V}{\partial y} dy + \frac{\partial V}{\partial z} dz \right) = -\int_1^2 dV$$

$$= V_1 - V_2 \qquad (1.24)$$

Combining Eqs. (1.18) and (1.24), we have $T_1 + V_1 = T_2 + V_2$ which gives the energy conservation theorem.

If the force acting on a particle is conservative, then the total energy of the particle, $T + V$, is a constant.

Equation (1.22) is satisfied even if we replace V by $V + C$, where C is a constant. Then $F = -\nabla V = -\nabla(V + C)$ (1.25)

Hence, the potential introduced through Eq. (1.22) is not unique and therefore an absolute value of the potential has no meaning. It may be noted that the kinetic energy also has no absolute value since we use an inertial frame of reference for measuring the velocity and hence the kinetic energy. For measuring the absolute kinetic energy we required a reference frame which is absolutely at rest. It is not possible to find such a reference frame and therefore the kinetic energy we measure is only relative.

1.5 MOTION UNDER A CONSTANT FORCE

When the applied force **F** on a particle is constant in time and hence there is a constant acceleration, we write Eq. (1.10) in the form

$$\frac{d^2\mathbf{r}}{dt^2} = \frac{d\mathbf{v}}{dt} = \frac{\mathbf{F}}{m} = \mathbf{a} = \text{constant} \qquad (1.26)$$

Direct integration of Eq. (1.26) is possible if the initial conditions are known.

2

System of Particles

The mechanics of a system of particles can be studied by using a straightforward application of Newton's laws.This application of Newton's laws considers the forces acting between particles in addition to the externally applied forces. One can easily extend the considerations of the mechanics of a single particle to a system of particles also.

2.1 CENTRE OF MASS

The mass of a point particle is concentrated at a particular point. When we consider the motion of a system of n particles, there is a point in it which behaves as if the entire mass of the system is concentrated at that point. This point is called the **centre of mass** of the system. The centre of mass C of a system of particles (see Fig. 2.1) whose radius vector is $\mathbf{R}$ is related to the masses m_i and radius vectors $\mathbf{r}_i$ of all n particles of the system by the equation

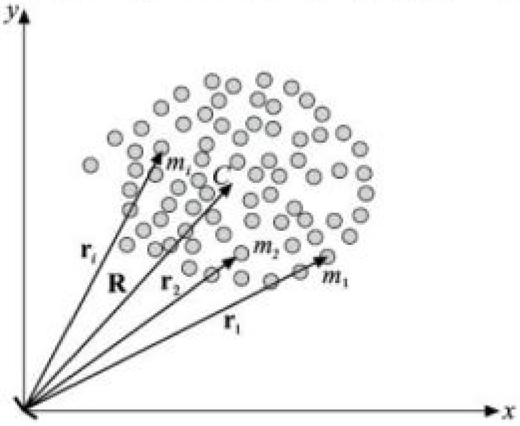


Fig. 2.1 Centre of mass of a system of *n* particles.

$$\mathbf{R} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2 + \ldots + m_i\mathbf{r}_i + \ldots}{m_1 + m_2 + \ldots + m_i + \ldots} = \frac{\sum_i m_i\mathbf{r}_i}{M} \qquad (2.1)$$

where M is the total mass of the system. For a continuous body, the co-ordinates of the centre of mass are

$$X = \frac{1}{M}\int_v \rho x\, dV \quad Y = \frac{1}{M}\int_v \rho y\, dV \quad Z = \frac{1}{M}\int_v \rho z\, dV$$

For a homogeneous body, the density ρ is constant and V is the volume of the body

$$X = \frac{1}{V}\int_v x\, dV \quad Y = \frac{1}{V}\int_v y\, dV \quad Z = \frac{1}{V}\int_v z\, dV \qquad (2.1a)$$

A frame of reference with the centre of mass as the origin is called the **centre of mass frame of reference**. In this frame of reference, obviously, the position vector of the centre of mass **R** is equal to zero. Consequently, the linear momentum **P** of the system ($d\mathbf{R}/dt$) is also zero. It is the practice to deal with all scattering problems in nuclear physics in this frame of reference.

2.2 CONSERVATION OF LINEAR MOMENTUM

Consider a system of n particles of masses $m_1, m_2, m_3, \ldots m_n$. Let their position vectors at time t be $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \ldots \mathbf{r}_n$. The force acting on the ith particle $\mathbf{F}_i$ has two parts: (i) a force applied on the system from outside or external force (ii) an internal force which is a force among the particles of the system. Newton's second law for the ith particle of the system can be written as

$$\frac{d\mathbf{p}_i}{dt} = \mathbf{F}_i = \mathbf{F}_{ie} + \sum_{j=1}^{n} \mathbf{F}_{ij} \quad j \neq i \qquad (2.2)$$

where $\mathbf{F}_{ie}$ is the external force on the ith particle and $\mathbf{F}_{ij}$ is the internal force on the ith particle due to the jth one. Since $\mathbf{F}_{ii} = 0$, j π i in the summation. Summing over all particles of the system, Eq. (2.2) takes the form

$$\sum_{i=1}^{n} \frac{d\mathbf{p}_i}{dt} = \sum_{i=1}^{n} \mathbf{F}_{ie} + \sum_{i=1}^{n} \sum_{j=1}^{n} \mathbf{F}_{ij} \quad j \neq i \qquad (2.3)$$

Assuming that Newton's third law is valid for the internal force $\mathbf{F}_{ij} = -\mathbf{F}_{ji}$ (2.4) Use of this condition reduces the second term on the right of Eq. (2.3) to zero. The first term $\sum_i \mathbf{F}_{ie} = \mathbf{F}_e$, the total external force acting on the system. The sum $\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3 + \ldots + \mathbf{p}_n = \mathbf{P}$
is the total linear momentum of the system. Now Eq. (2.3) reduces to

$$\frac{d\mathbf{P}}{dt} = \mathbf{F}_e \qquad (2.5)$$

which provides the law of conservation of linear momentum of a system of particles: *If the external force acting on a system of particles is zero, then the total linear momentum of the system is conserved.*
When external force acting on a system is zero, it is called a ***closed system***. For a closed system, linear momentum is conserved.

Another interesting result is the relation connecting the total linear momentum and the velocity of the centre of mass. With the definition of centre of mass in Eq. (2.1)

$$\sum_i \frac{d\mathbf{p}_i}{dt} = \frac{d^2}{dt^2} \sum_i m_i \mathbf{r}_i = M \frac{d^2\mathbf{R}}{dt^2}$$

and Eq. (2.3) takes the form

$$M \frac{d^2\mathbf{R}}{dt^2} = \frac{d\mathbf{P}}{dt} = \mathbf{F}_e \qquad (2.6)$$

That is, the centre of mass moves as if the total external force were acting on the entire mass of the system concentrated at the centre of mass.

2.3 ANGULAR MOMENTUM

We now derive the angular momentum **L** of a system of particles which is defined as $\mathbf{L} = \sum_i \mathbf{r_i} \times \mathbf{p}_i$ (2.7)

Figure 2.2 illustrates the position vector of the centre of mass of the system and that of the *i*th particle.

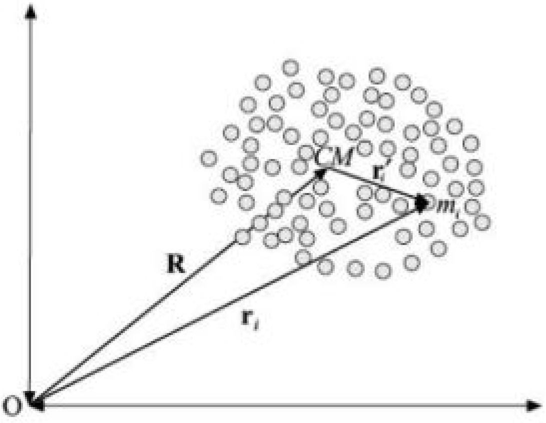


Fig. 2.2 Position of centre of mass and *i*th particle.

From Fig. 2.2 we have

$$\mathbf{r}_i = \mathbf{R} + \mathbf{r}_i{}' \quad \text{or} \quad \dot{\mathbf{r}}_i = \dot{\mathbf{R}} + \dot{\mathbf{r}}_i{}' \tag{2.8}$$

Consequently,

$$\mathbf{p}_i = m_i \dot{r}_i = m_i(\dot{\mathbf{R}} + \dot{\mathbf{r}}_\mathbf{i}{}') \tag{2.9}$$

Substituting these values in Eq. (2.7)

$$L = \sum_i (R + r_i') \times m_i \; (\dot{\mathbf{R}} + \dot{r}_i')$$

$$= \sum_i m_i \; (\mathbf{R} \times \dot{\mathbf{R}}) + \left(\mathbf{R} \times \sum_i m_i \; \dot{r}'_i \right) + \left(\sum_i m_i \mathbf{r}_i{}' \times \dot{\mathbf{R}} \right) + \sum_i (m_i \mathbf{r}_i{}' \times \dot{r}'_i) \tag{2.10}$$

The quantity $\sum_i m_i \, r_i{}'$ vanishes as it defines the radius vector of the centre of mass in the co-ordinate system in which the origin is the centre of mass. The quantity

$$\sum_i m_i \dot{\mathbf{r}}_i' = \frac{d}{dt}\left(\sum_i m_i \mathbf{r}_i'\right) = 0$$

Hence, the total angular momentum

$$\mathbf{L} = \sum_i m_i\,(\mathbf{R} \times \dot{\mathbf{R}}) + \sum_i m_i r_i' \times \dot{r}_i$$

$$= \mathbf{R} \times M\mathbf{V}_{CM} + \sum_i (\mathbf{r}_i' \times \mathbf{p}_i') \qquad (2.11)$$

where $\mathbf{V}_{CM}$ is the velocity of the centre of mass with respect to the origin O. The meaning of the equation is that the total angular momentum about a point O is equal to the sum of the angular momentum of the system concentrated at the centre of mass and the angular momentum of the system of particles about the centre of mass.

2.4 CONSERVATION OF ANGULAR MOMENTUM

We now consider the angular momentum of a system of n particles which is

$$\mathbf{L} = \sum_{i=1}^{n} \mathbf{r}_i \times \mathbf{p}_i$$

defined as

$$\frac{d\mathbf{L}}{dt} = \sum_i \frac{d\mathbf{r}_i}{dt} \times \mathbf{p}_i + \sum_i \mathbf{r}_i \times \frac{d\mathbf{p}_i}{dt} \qquad (2.12)$$

The first term on the right is zero since the vector product of a vector with itself is zero. Substituting for $(d\mathbf{p}_i \,/\, dt)$ from Eq. (2.2)

$$\frac{d\mathbf{L}}{dt} = \sum_i \left[\mathbf{r}_i \times \left(\mathbf{F}_{ie} + \sum_j \mathbf{F}_{ij}\right)\right]$$

$$= \sum_i (\mathbf{r}_i \times \mathbf{F}_{ie}) + \sum_i \sum_j (\mathbf{r}_i \times \mathbf{F}_{ij}) \quad j \neq i \qquad (2.13)$$

The second term on the right contains pairs of terms like

$$\mathbf{r}_i \times \mathbf{F}_{ij} + \mathbf{r}_j \times \mathbf{F}_{ji}$$

Since $\mathbf{F}_{ij} = -\mathbf{F}_{ji}$,this pair reduces to

$$\mathbf{r}_i \times \mathbf{F}_{ij} - \mathbf{r}_j \times \mathbf{F}_{ij} = (\mathbf{r}_i - \mathbf{r}_j) \times \mathbf{F}_{ij} = \mathbf{r}_{ij} \times \mathbf{F}_{ij} \qquad (2.14)$$

which is zero if the internal forces are central, that is, the internal forces are along the line joining the two particles. Hence, the second term on the right of Eq. (2.13) vanishes. Since $\mathbf{r}_i$ $\mathbf{F}_{ie}$ is the torque due to the external force on the *i*th particle, Eq. (2.13) reduces to

$$\frac{d\mathbf{L}}{dt} = \sum_{\mathrm{i}} \mathbf{N}_{ie} = \mathbf{N}_e \qquad (2.15)$$

where N_{e} is the total external torque acting on the system. Eq. (2.15) leads to the conservation law: *If the total torque due to external forces on a system of particles is zero, then the total angular momentum is a constant of motion.*

2.5 KINETIC ENERGY FOR A SYSTEM OF PARTICLES

For a system of particles the kinetic energy of the system

$$T = \frac{1}{2}\sum_{i=1}^{n} m_i \mathbf{v}_i^2 \qquad (2.16)$$

The position of the centre of mass of the system and that of the *i*th particle is shown in Fig. 2.2. From the figure, we have

$$\mathbf{r}_i = \mathbf{R} + \mathbf{r}_i' \quad \text{or} \quad \mathbf{v}_i = \mathbf{V}_{CM} + \mathbf{v}_i'$$

With this value of $\mathbf{v}_i$, Eq. (2.16) takes the form

$$T = \frac{1}{2}\sum_i m_i\,(\mathbf{V}_{CM} + \mathbf{v}_i').(\mathbf{V}_{CM} + \mathbf{v}_i')$$

The term $\sum m_i \mathbf{r}_i'$ vanishes as it defines the radius vector of the centre of mass in the co-ordinate system in which the origin is the centre of mass. Hence,

$$T = \frac{1}{2} M\, \mathbf{V}_{CM}^2 + \frac{1}{2}\sum_i m_i\, \mathbf{v}_i'^2 \qquad (2.18)$$

Thus, like angular momentum, the kinetic energy also consists of two parts:

(i) the kinetic energy obtained if all the mass were concentrated at the centre of mass, and (ii) the kinetic energy of motion about the centre of mass.

2.6 ENERGY CONSERVATION OF A SYSTEM OF PARTICLES

The energy conservation law of a single particle system can easily be extended to a system of particles. The force acting on the ith particle is given by Eq. (2.2). As in the case of a single particle, the work done by all forces in moving the system from an initial position 1 to a final position 2 is given by

$$W_{12} = \sum_i \int_1^2 \mathbf{F}_i \,.\, d\mathbf{r}_i = \sum_i \int_1^2 \mathbf{F}_{ie} \,.\, d\mathbf{r}_i + \sum_i \sum_j \int_1^2 \mathbf{F}_{ij} \,.\, d\mathbf{r}_i \quad i \neq j \qquad (2.19)$$

Again, reducing the integral $\sum_i \int_1^2 \mathbf{F}_i \,.\, d\mathbf{r}_i$ using equation of motion, we have

$$W_{12} = \sum_i \int_1^2 \mathbf{F}_i \,.\, d\mathbf{r}_i = \sum_i m_i \int_1^2 \frac{d\mathbf{v}_i}{dt} . d\mathbf{r}_i = \sum_i m_i \int_1^2 \frac{d\mathbf{v}_i}{dt} . \mathbf{v}_i dt$$

$$= \sum_i m_i \int_1^2 \mathbf{v}_i \, d\mathbf{v}_i = \sum_i \frac{1}{2} m_i [\mathbf{v}_i^2]_1^2$$

$$= \mathrm{T}_2 - \mathrm{T}_1 \qquad (2.20)$$

where T is the total kinetic energy of the system.

Next we consider the right hand side of Eq. (2.19). If both $\mathbf{F}_{ie}$ and $\mathbf{F}_{ij}$ are conservative, they are derivable from potential functions $\mathbf{F}_{ie} = -\nabla_i V_{ie}(\mathbf{r}_i)$ and $\mathbf{F}_{ij} = -\nabla V_{ij}(\mathbf{r}_i, \mathbf{r}_j)$ (2.21) where the subscript i on the del operator indicates that the derivative is with respect to the co-ordinates of the ith particle. The first term on the right side of Eq. (2.19) now takes the form

$$\sum_i \int_1^2 \mathbf{F}_{ie} . d\mathbf{r}_i = -\sum_i \int_1^2 \nabla_i V_{ie} (\mathbf{r}_i) . d\mathbf{r}_i$$

$$= -\sum_i \int_1^2 dV_{ie} = -\left[\sum_i V_{ie}\right]_1^2 \tag{2.22}$$

In order to satisfy Newton's third law, $\mathbf{F}_{ij} = -\mathbf{F}_{ji}$. Hence,

$$\sum_i \sum_j \mathbf{F}_{ij} . d\mathbf{r}_i = \sum_i \sum_j \mathbf{F}_{ji} . d\mathbf{r}_j = -\sum_i \sum_j \mathbf{F}_{ij} . d\mathbf{r}_j \quad i \neq j \tag{2.23}$$

Consequently,

$$\sum_i \sum_j \mathbf{F}_{ij} . d\mathbf{r}_i = \frac{1}{2} \sum_i \sum_j \mathbf{F}_{ij} . (d\mathbf{r}_i - d\mathbf{r}_j) \quad i \neq j$$

$$= \frac{1}{2} \sum_i \sum_j \mathbf{F}_{ij} . d\mathbf{r}_{ij} \qquad d\mathbf{r}_{ij} = d\mathbf{r}_i - d\mathbf{r}_j \quad i \neq j \tag{2.24}$$

where the factor ½ is introduced to avoid each member of a pair being included twice, first in the *i* summation and then in the *j* summation. Substituting this value

$$\sum_i \sum_j \int_1^2 \mathbf{F}_{ij} . d\mathbf{r}_i = \frac{1}{2} \sum_i \sum_j \int_1^2 \mathbf{F}_{ij} . d\mathbf{r}_{ij} \quad i \neq j$$

$$= -\frac{1}{2} \sum_i \sum_j \int_1^2 \nabla_{ij} V_{ij} . d\mathbf{r}_{ij} \quad i \neq j$$

$$= -\frac{1}{2} \sum_i \sum_j \int_1^2 d\mathbf{V}_{ij}$$

$$= -\frac{1}{2} \left[\sum_i \sum_j V_{ij}\right]_1^2 \quad i \neq j \tag{2.25}$$

Here ∇_{ij} stands for the gradient with respect to r_{ij}.

Equation (2.19) can now be written as

3

Lagrangian Formulation

In the previous chapters we were able to demonstrate the effectiveness of Newton's laws of motion in solving variety of problems. However, if the system is subject to external constraints, solving the equations of motion may be difficult, and sometimes it may be difficult even to formulate them. The forces of constraints are usually very complex or unknown, which makes the formalism more difficult. To circumvent these difficulties, two different methods, Lagrange's and Hamilton's formulations, have been developed. These techniques use an energy approach and are constructed in such a way that the Newtonian formalism follows from it. Before going over to these procedures, we try to understand certain terms such as constraints, generalized coordinates, *etc*. In this chapter a discussion on the Lagrangian formalism is given.

3.1 CONSTRAINTS

A motion that cannot proceed arbitrarily in any manner is called a **constrained motion**. The conditions which restricts the motion of the system are called **constraints**. For example, gas molecules within a container are constrained by the walls of the vessel to move only inside the container. A particle placed on the surface of a solid sphere is restricted by the constraint, so that it can only move on the surface or in the region exterior to the sphere. There are two main types of constraints, holonomic and non-holonomic.

Holonomic Constraints

In holonomic constraints, the conditions of constraint are expressible as equations connecting the coordinates and time, having the form $f(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \ldots \mathbf{r}_n, t) = 0$ (3.1) We give below a few typical examples of holonomic constraint: (i) In a rigid body, the distance between any two particles of the body remains constant during motion. This is expressible as $|\mathbf{r}_i - \mathbf{r}_j|^2 = c_{ij}^2$ (3.2) where c_{ij} is the distance between the particles i and j at $\mathbf{r}_i$ and $\mathbf{r}_j$.

(ii) The sliding of a bead on a circular wire of radius a in the xy-plane is another example. The equation of constraint is $x^2 + y^2 = a^2$ (3.3) which can also be expressed in the differential form as $x\,dx + y\,dy = 0$ (3.3a) Equations (3.2) and (3.3) are of the same form as Eq. (3.1). The differential equation denoted by Eq. (3.3a) can be integrated to obtain Eq. (3.3). Holonomic constraints are also known as **integrable constraints**. The term *integrable* is used here since Eq. (3.1) is equivalent to the differential equation $\sum_i \frac{\partial f}{\partial r_i} d\mathbf{r}_i = 0$ (3.4) Equation (3.4) can be readily integrated to Eq. (3.1).

Non-holonomic Constraints

Non-holonomic constraints are those which are not expressible in the form of Eq. (3.1). The coordinates in this case are restricted either by inequalities or by non-integrable differentials.

(i) The constraint involved in the example of a particle placed on the surface of a sphere is non-holonomic, which may be expressed as the inequality $r^2 - a^2 \geq 0$ (3.5) where a is the radius of the sphere.

(ii) Gas molecules in a spherical container of radius R. If $\mathbf{r}_i$ is the position vector of the ith molecule, $x_i^2 + y_i^2 + z_i^2 \leq R^2$ (3.6) Here, the centre of the sphere is the origin of the coordinate system.

In non-holonomic constraints, if the constraints are expressible as relations among the velocities of the particles of the system, that is, $f(x_1, x_2, \ldots, \dot{x}_1, \dot{x}_2, \ldots t) = 0$ (3.6a) and if these equations of non-holonomic constraints can be integrated to give relations among the coordinates, then the constraints become holonomic.

Scleronomous and Rheonomous Constraints Constraints are further classified as scleronomous and rheonomous.

A scleronomous constraint is one that is independent of time

whereas a rheonomous constraint contains time explicity. A pendulum with an inextensible string of length l_0 is described by the equation $x^2 + y^2 = l_0^2$ (3.7) As the constraint equation is independent of time, it is a scleronomous constraint. A pendulum with an extensible string is rheonomous, the condition of constraint is $x^2 + y^2 = l^2(t)$ (3.8) where $l(t)$ is the length of the string at time t.

Constraints introduce two types of difficulties in the solution of mechanical problems. The coordinates r_i are no longer independent as they are connected by the equations of constraint. In the case of holonomic constraints, this difficulty is solved by the introduction of generalized coordinates. The second difficulty is due to the fact that the forces of constraints cannot be specified explicitly. They are among the unknowns of the problem and must be obtained from the solution. This difficulty can be solved if the problem is formulated in the Lagrangian form, in which the forces of constraint do not appear.

In most of the systems of interest, the constraints involved are holonomic. Hence, we restrict ourselves mainly to holonomic systems.

3.2 GENERALIZED COORDINATES

Degrees of Freedom

The number of independent ways in which a mechanical system can move without violating any constraint is called the **number of degrees of freedom** of the system. It is the minimum possible number of coordinates required to describe the system completely. When a particle moves in space, it has three degrees of freedom. If it is constrained to move along a space curve it has only one degree of freedom whereas it has two degrees of freedom if it moves in a plane.

Generalized Coordinates

For a system of N particles, free from constraints, we require a total of $3N$ independent coordinates to describe its configuration completely. Let there are k

constraints of the type $f_s(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N, t) = 0 \quad s = 1, 2, 3, ..., k$ (3.9) acting on the system. Now the system has only $3N - k$ independent coordinates or degrees of freedom. These $3N - k$ independent coordinates represented by the variables $q_1, q_2, q_3, ..., q_{3N-k}$ are called the **generalized coordinates**. In terms of the new coordinates, the old coordinates $r_1, r_2, ..., r_N$ can be written as

$$\begin{aligned} \mathbf{r}_1 &= \mathbf{r}_1(q_1, q_2, ..., q_{3N-k}, t) \\ \mathbf{r}_2 &= \mathbf{r}_2(q_1, q_2, ..., q_{3N-k}, t) \qquad (3.10) \\ &\cdots\cdots\cdots\cdots \\ \mathbf{r}_N &= \mathbf{r}_N(q_1, q_2, ..., q_{3N-k}, t) \end{aligned}$$

These are the transformation equations from the set of r_i variables to q_l variables. In analogy with cartesian coordinates, time derivatives $\dot{q}_1, \dot{q}_2, \dot{q}_3 ...$ are defined as **generalized velocities**.

Generalized coordinates are not unique. They may or may not have dimensions of length. Depending on the problem, it may prove more convenient to select some of the coordinates with dimensions of energy, some others with dimensions of L^2, and yet some others could be combinations of angles and coordinates, and so on.

Configuration Space

We have seen that the configuration of a system can be specified completely by the values of $n = 3N - k$ independent generalized coordinates $q_1, q_2, ..., q_n$. It is convenient to think of the n q's as the coordinates of a point in an n-dimensional space. This n-dimensional space is called the **configuration space** with each dimension represented by a coordinate. As the generalized coordinates are not necessarily position coordinates, configuration space is not necessarily connected to the physical 3-dimensional space and the path of motion also does not necessarily resemble the path in space of actual particle.

3.3 PRINCIPLE OF VIRTUAL WORK

A **virtual displacement**, denoted by $d\mathbf{r}_i$, refers to an imagined, infinitesimal, instantaneous displacement of the coordinate that is consistent with the constraints. It is different from an actual displacement $d\mathbf{r}_i$ of the system occurring in a time interval *dt*. It is called *virtual* as the displacement is instantaneous. As there is no actual motion of the system, the work done by the forces of constraint in such a virtual displacement is zero.

Consider a scleronomic system of *N* particles in equilibrium. Let $\mathbf{F}_i$ be the force acting on the *i*th particle. The force $\mathbf{F}_i$ is a vector addition of the externally applied force $\mathbf{F}_i^e$ and the forces of constraints $\mathbf{f}_i$. Then $\mathbf{F}_i = \mathbf{F}_i^e + \mathbf{f}_i$ (3.11)

If $d\mathbf{r}_i$ is a virtual displacement of the *i*th particle, the virtual work done dW_i on the *i*th particle is given by $\delta W_i = \mathbf{F}_i . \delta\mathbf{r}_i$ (3.12) If the system is in equilibrium, the total force on each particle must be zero:

$\mathbf{F}_i = 0$ for all *i*. Therefore, the dot product $\mathbf{F}_i . \delta\mathbf{r}_i$ is also zero. That is, $\delta W_i = (\mathbf{F}_i^e + \mathbf{f}_i).\delta\mathbf{r}_i = 0 \quad i = 1, 2, \ldots, N$ (3.13) The total virtual work done on the system dW is the sum of the above vanishing products:

$$\delta W = \sum_{i=1}^{N} \delta W_i = \sum_{i=1}^{N} (\mathbf{F}_i^e + \mathbf{f}_i).\delta\mathbf{r}_i = 0$$

$$= \sum_{i=1}^{N} F_i^e .\delta\mathbf{r}_i + \sum_{i=1}^{N} \mathbf{f}_i .\delta\mathbf{r}_i = 0 \qquad (3.14)$$

Under a virtual displacement, the work done by the forces of constraints is zero. This is valid for rigid bodies and most of the constraints that commonly occur. Therefore, Eq. (3.14) reduces to

$$\delta W = \sum_{i=1}^{N} \boldsymbol{F}_i^e .\delta \boldsymbol{r}_i = 0 \qquad (3.15)$$

which is the **principle of virtual work** and is stated as : *In an N-particle system, the total work done by the external forces when virtual displacements are made is called virtual work and the total virtual work done is zero.*

The coefficients $\delta\mathbf{r}_i$ in Eq. (3.15) can no longer be set equal to zero as they are not independent. It should also be noted that the principle of virtual work deals only with statics.

3.4 D’ALEMBERT’S PRINCIPLE

The principle of virtual work deals only with statics and the general motion of the system is not relevant here. A principle that involves the general motion of the system was suggested by D' Alembert.

Consider the motion of an N-particle system. Let the force acting on the *i*th particle be $\mathbf{F}_i$. By Newton's law $\mathbf{F}_i = \mathbf{p}_i$ or $\mathbf{F}_i - \dot{\mathbf{p}}_i = 0$ (3.16) This means that the *i*th particle in the system will be in equilibrium under a force equal to the actual force plus a "reversed effective force", $-\dot{\mathbf{p}}_i$, as named by D'Alembert. Then dynamics reduces to statics. To this equivalent static problem, give a virtual displacement $\delta\mathbf{r}_i$ which leads to

$$\sum_{i=1}^{N} (\mathbf{F}_i - \dot{\mathbf{p}}_i).\delta\mathbf{r}_i = 0 \qquad (3.17)$$

$$\sum_{i=1}^{N} (\mathbf{F}_i^e + \mathbf{f}_i - \dot{\mathbf{p}}_i).\delta\mathbf{r}_i = 0 \qquad (3.18)$$

Restricting to situations where the virtual work done by forces of constraints is zero $\sum_{i=1}^{N} (\mathbf{F}_i^e - \dot{\mathbf{p}}_i).\boldsymbol{\delta}\mathbf{r}_i = 0$ (3.19) which is **D'Alembert's principle**.

3.5 LAGRANGE'S EQUATIONS

Lagrange used D'Alembert's principle as the starting point to derive the equations of motion, now known as Lagrange's equations. Dropping the superscript *e* in Eq. (3.19) $\sum_{i=1}^{N} (F_i - \dot{\mathbf{p}}_i).\boldsymbol{\delta}\mathbf{r}_i = 0$ (3.20) The virtual displacements $\boldsymbol{\delta}\mathbf{r}_i$ in Eq. (3.20) are not independent. Lagrange changed Eq. (3.20) into an equation involving virtual displacment of the generalized coordinates which are independent.

Consider a system with N particles at $\mathbf{r}_1$, $\mathbf{r}_2$,..., $\mathbf{r}_N$ having k equations of holonomic constraints. The system will have $n = 3N - k$ generalized coordinates $q_1, q_2, \ldots, q_n$. The transformation equations from the r variables to the q variables are given by Eq. (3.10).

$\mathbf{r}_i = \mathbf{r}_i (q_1, q_2, \ldots, q_n, t)$ (3.21) Since virtual displacement does not involve time, from Eq. (3.21)

$$\delta \mathrm{r}_i = \frac{\partial \mathbf{r}_i}{\partial q_1}\delta q_1 + \frac{\partial \mathbf{r}_i}{\partial q_2}\delta q_2 + \ldots + \frac{\partial \mathbf{r}_i}{\partial q_n}\delta q_n = \sum_j \frac{\partial \mathbf{r}_i}{\partial q_j}\delta q_j \qquad (3.22)$$

Here dq_j's are the virtual displacements of generalized coordinates. From Eq. (3.21) we also have $\dot{\mathbf{r}}_i = \frac{d\mathbf{r}_i}{dt} = \frac{\partial \mathbf{r}_i}{\partial q_1}\dot{q}_1 + \frac{\partial \mathbf{r}_i}{\partial q_2}\dot{q}_2 + \ldots + \frac{\partial \mathbf{r}_i}{\partial q_n}\dot{q}_n + \frac{\partial \mathbf{r}_i}{\partial t}$

$$= \sum_j \frac{\partial \mathbf{r}_i}{\partial q_j}\dot{q}_j + \frac{\partial \mathbf{r}_i}{\partial t} \qquad (3.23) \qquad \frac{\partial \dot{\mathbf{r}}_i}{\partial \dot{q}_j} = \frac{\partial \mathbf{r}_i}{\partial q_j} \qquad (3.24)$$

The form of D'Alembert's principle, Eq. (3.20), can be changed easily by substituting $\boldsymbol{\delta r_i}$ from Eq. (3.22). The first term of Eq. (3.20) is

$$\sum_i \mathbf{F}_i . \delta \mathbf{r}_i = \sum_i \mathbf{F}_i . \left(\sum_j \frac{\partial \mathbf{r}_i}{\partial q_j}\delta q_j \right) = \sum_j \left(\sum_i \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \right) \delta q_j$$

$$= \sum_{j=1}^{n} Q_j \delta q_j \qquad (3.25)$$

where,

$$Q_j = \sum_i \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} \qquad (3.26)$$

The quantity Q_j is the jth component of the generalized force **Q**. The generalized force components need not have the dimension of force as the q's need not have the dimension of length. However, Q_jdq_j must have the dimension of work.

We now write the inertial force term of Eq. (3.20)

$$\sum_i \dot{\mathbf{p}}_i \, .\delta \mathbf{r}_i = \sum_i m_i \ddot{\mathbf{r}}_i \, .\delta \mathbf{r}_i = \sum_i m_i \ddot{\mathbf{r}}_i \left(\sum_j \frac{\partial \mathbf{r}_i}{\partial q_j} \delta q_j \right)$$

$$= \sum_j \left[\sum_i \frac{d}{dt} \left(m_i \dot{\mathbf{r}}_i \frac{\partial \mathbf{r}_\mathbf{i}}{\partial q_j} \right) - \sum_i m_i \dot{\mathbf{r}}_i \frac{d}{dt} \frac{\partial \mathbf{r}_i}{\partial q_j} \right] \delta q_j \tag{3.27}$$

Using Eq. (3.24)

$$\sum_i \frac{d}{dt} \left(m_i \dot{\mathbf{r}}_i \frac{\partial \mathbf{r}_i}{\partial q_j} \right) = \sum_i \frac{d}{dt} \left(m_i \dot{\mathbf{r}}_i \frac{\partial \dot{\mathbf{r}}_\mathbf{i}}{\partial \dot{q}_j} \right) = \frac{d}{dt} \left(\sum_i m_i \mathbf{v}_i \frac{\partial \mathbf{v}_i}{\partial \dot{q}_j} \right)$$

$$= \frac{d}{dt} \frac{\partial}{\partial \dot{q}_j} \left(\sum_i \frac{1}{2} m_i \mathbf{v}_i^2 \right) = \frac{d}{dt} \frac{\partial T}{\partial \dot{q}_j} \tag{3.28}$$

where T is the total kinetic energy of the system. Changing the order of differentiation in the second term of Eq. (3.27)

$$\sum_i m_i \dot{\mathbf{r}}_i \frac{d}{dt} \frac{\partial \mathbf{r}_i}{\partial q_j} = \sum_i m_i \dot{\mathbf{r}}_i \frac{\partial \dot{\mathbf{r}}_i}{\partial q_j} = \frac{\partial}{\partial q_j} \left(\sum_i \frac{1}{2} m_i \mathbf{v}_i^2 \right) = \frac{\partial T}{\partial q_j} \tag{3.29}$$

Use of Eqs. (3.28) and (3.29) reduces Eq. (3.27) to

$$\sum_i \dot{\mathbf{p}}_i \, .\delta \mathbf{r}_i = \sum_{j=1}^{n} \left(\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} \right) \tag{3.30}$$

With Eqs. (3.26) and (3.30), Eq. (3.20) becomes

$$\sum_{j=1}^{n} \left(\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} - Q_j \right) \delta q_j = 0 \tag{3.31}$$

The *dq's* are independent and therefore each of the coefficients must separately vanish. From which it follows that

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} - Q_j = 0 \quad j = 1, 2, \ldots, n \tag{3.32}$$

Equation (3.32) can be simplified further if the external forces $\mathbf{F}_i$ are conservative: $\mathbf{F}_i = -\nabla_i V$ where $V = V(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$. Then

$$Q_j = -\sum_i \nabla_i V \cdot \frac{\partial \mathbf{r}_i}{\partial q_j} = -\sum_i \frac{\partial V}{\partial \mathbf{r}_i} \frac{\partial \mathbf{r}_i}{\partial q_j}$$

$$= -\frac{\partial V}{\partial q_j} \qquad (3.33)$$

since

$$dV = \sum_i \frac{\partial V}{\partial \mathbf{r}_i} d\mathbf{r}_i = \sum_i \frac{\partial V}{\partial \mathbf{r}_i} \left(\sum_j \frac{\partial \mathbf{r}_i}{\partial q_j} dq_j \right)$$

$$\frac{\partial V}{\partial q_j} = \sum_i \frac{\partial V}{\partial \mathbf{r}_i} \frac{\partial \mathbf{r}_i}{\partial q_j} \qquad (3.34)$$

Equation (3.32) becomes

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} + \frac{\partial V}{\partial q_j} = 0 \quad j = 1, 2, \ldots, n \qquad (3.35)$$

If the potential V is a function of position only, $\left(\frac{\partial V}{\partial \dot{q}_j} \right) = 0$. We can now include this term in Eq. (3.35). Then

$$\frac{d}{dt}\frac{\partial}{\partial \dot{q}_j}(T - V) - \frac{\partial}{\partial q_j}(T - V) = 0 \quad j = 1, 2, \ldots, n \qquad (3.36)$$

We now introduce a new function L defined by $L(q, \dot{q}, t) = T(q, \dot{q}, t) - V(q)$

(3.37) where q stands for q_1, q_2, q_3,..., q_n and $\dot{q}$ stands for $\dot{q}_1, \dot{q}_2, \dot{q}_3, \ldots, \dot{q}_n$. This function L is called the **Lagrangian function** of the system. In terms of L, Eq. (3.36) becomes $\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j} \right) - \frac{\partial L}{\partial q_j} = 0 \quad j = 1, 2, 3, \ldots, n$

(3.38) These n equations, one for each independent generalized coordinate, are known as **Lagrange's equations**. These constitute a set of n second order differential equations for n unknown functions $q_j(t)$ and the general solution contains
2n integration constants.

3.7 GENERALIZED MOMENTUM

Consider the motion of a particle of mass m moving along x-axis. Its linear momentum $\mathbf{p}$ is $m\dot{x}$ and kinetic energy $T = (\tfrac{1}{2})\ m\ \dot{x}^2$. Differentiating T with respect to $\dot{x}$ we have $\frac{\partial T}{\partial \dot{x}} = m\ddot{x} = \mathbf{p}$ (3.50) If the potential V is not a function of the velocity $\dot{x}$, since $L = T - V$

$\mathbf{p} = \frac{\partial T}{\partial \dot{x}} = \frac{\partial L}{\partial \dot{x}}$ (3.51) Let us use this concept to define generalized momentum. For a system described by a set of generalized coordinates q_1, q_2,..., q_n, we define **generalized momentum** $\mathbf{p}_i$ corresponding to generalized coordinate q_i as $\mathbf{p}_i = \frac{\partial L}{\partial \dot{q}_i}$ (3.52) Sometimes it is also known as **conjugate momentum** (conjugate to co-ordinate q_i).

In general, generalized momentum is a function of the q's, $\dot{q}$'s and t . As the Lagrangian is utmost quadratic in the $\dot{q}$'s, $\mathbf{p}_i$ is a linear function of the $\dot{q}$'s. The generalized momentum $\mathbf{p}_i$ need not always have the dimension of linear momentum. However, the product of any generalized momentum and the associated coordinate must always have the dimension of angular momentum. For a conservative system, the use of the expression for generalized momentum, Eq. (3.52), reduces Lagrange's equations of motion to $\dot{\mathbf{p}}_j = \frac{\partial L}{\partial q_j} \quad j = 1, 2, \ldots, n$

(3.53)

3.8 FIRST INTEGRALS OF MOTION AND CYCLIC

COORDINATES

Lagrange's equations of motion for a system having n degrees of freedom will have n differential equations that are second order in time. As the solution of each equation requires two integration constants, a total of $2n$ constants have to be evaluated from the initial values of n-generalized coordinates and n-generalized velocities. In general, it is either very difficult to solve the problem completely or very tedious. However, very often a great deal of information about the system is possible from the first integrals of equations of motion. The *first integrals of motion* are functions of the generalized coordinates q's and generalized velocities $\dot{q}$'s of the form $f(q_1, q_2, ..., q_n, \dot{q}_1, \dot{q}_2, ..., \dot{q}_n, t) = \alpha_i$ (constant) (3.54) These first integrals are of interest because they give good deal of information about the system. The conservation laws of energy, momentum and angular momentum that we deduced in Newtonian formalism are of this type. In the process, the relation between conservation laws and the symmetry properties of the system is revealed.

Cyclic Coordinates

Coordinates that do not appear explicitly in the Lagrangian of a system (although it may contain the corresponding generalized velocities) are said to be **cyclic or ignorable**. If q_i is a cyclic coordinate $L = L(q_1, q_2, ... q_{i-1}, q_{i+1}, ..., q_n, \dot{q}_1, \dot{q}_2, ..., \dot{q}_n, t)$ (3.55) In such a case $(\partial L/\partial q_i) = 0$ and Lagrange's equation reduces to

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = 0 \quad \text{or} \quad \frac{\partial L}{\partial \dot{q}_i} = \text{constant } \alpha_i$$

which means that

$$\frac{\partial L}{\partial \dot{q}_i} = \mathbf{p}_i = \text{constant } \alpha_i$$

(3.56) Equation (3.56) constitutes a first integral for the equations of motion. We may state this result as a general conservation theorem: *The generalized momentum conjugate to a cyclic coordinate is conserved during the motion.*

3.9 CONSERVATION LAWS AND SYMMETRY PROPERTIES

The title suggests the possibility of a relationship between the conservation laws and symmetries. In this section, we shall investigate the connection between the two in detail. A closed system is one that does not interact with other systems.

Homogeneity of Space and Conservation of Linear Momentum

Homogeneity in space means that the mechanical properties of a closed system remain unchanged by any parallel displacement of the entire system in space. That means that the Lagrangian is unchanged ($dL = 0$) if the system is displaced by an infinitesimal amount $\delta \mathbf{r} : \mathbf{r}_i \rightarrow \mathbf{r}_i + \delta \mathbf{r}_i$. The change in L due to infinitesimal displacement $d\mathbf{r}$, the velocities remaining fixed, is given by

$$\delta L = \sum_i \frac{\partial L}{\partial \mathbf{r}_i} \delta \mathbf{r}_i \qquad (3.57)$$

The second term in this equation vanished as velocities remained constant $(\delta \dot{\mathbf{r}}_i = 0)$. Since each of the $\delta \mathbf{r}_i$ in Eq. (3.57) is an arbitrary independent displacement, the coefficient of each term is zero separately. Hence,

$$\frac{\partial L}{\partial \mathbf{r}_i} = 0 \qquad (3.58)$$

With this condition, Lagrange's equation reduces to

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{r}}_i} = 0 \quad \text{or} \quad \frac{\partial L}{\partial \dot{\mathbf{r}}_i} = \text{constant}$$

$$\mathbf{p}_i = \text{constant} \qquad (3.59)$$

As the $\mathbf{p}_i$'s are additive, the total linear momentum $\mathbf{p}$ of a closed system is a constant. Thus, the homogeneity of space implies that the linear momentum $\mathbf{p}$ is a constant of motion.

It can also be proved that if the Lagrangian of a system (not necessarily closed) is invariant with respect to translation in a certain direction, then the linear momentum of the system in that direction is constant in time.

Isotropy of Space and Conservation of Angular Momentum

Space is isotropic, which means the mechanics (i.e., the Lagrangian) of a closed system is unaffected by an infinitesimal rotation of the system in space, i.e.,
$dL = 0$. Consider a cartesian frame of reference with O as the origin. Let $\mathbf{r}_i$ be the radius vector of the ith particle located at P. Let the system as a whole undergoes an infinitesimal rotation df. The displacement is denoted by the vector df and its direction is that of the axis of rotation. Due to this rotation, the position vector

of the ith particle is shifted from P to P and the radius vector $\mathbf{r}_i$ to $\mathbf{r}_i + d\mathbf{r}_i$

(see Fig. 3.1)

$$|\delta\mathbf{r}_i| = \mathbf{r}_i \sin\theta_i \,\delta\phi$$

$$\delta\mathbf{r}_i = \delta\boldsymbol{\phi} \times \mathbf{r}_i \qquad (3.60)$$

When the system is rotated, the position vectors of all particles change their directions in this way. The corresponding change in the velocity vector is given by $\delta\dot{\mathbf{r}}_i = \delta\mathbf{f} \times \dot{\mathbf{r}}_i$ (3.61)

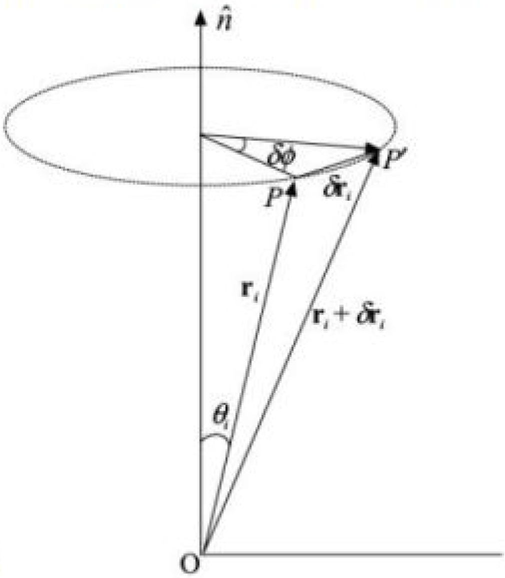


Fig. 3.1 Change of a position vector under rotation of the system.

The condition that $dL = 0$ leads to

$$\delta L = \sum_i \left(\frac{\partial L}{\partial \mathrm{r}_i} \delta \mathbf{r}_i + \frac{\partial L}{\partial \dot{\mathbf{r}}_i} \delta \dot{\mathbf{r}}_i \right) = 0 \tag{3.62}$$

Equations (3.38) and (3.52) give

$$\frac{\partial L}{\partial \mathbf{r}_i} = \frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{r}}_i} = \dot{\mathbf{p}}_i \quad \text{and} \quad \frac{\partial L}{\partial \dot{\mathbf{r}}_i} = \mathbf{p}_i$$

Equations (3.62) now becomes

$$\sum_i \left(\dot{\mathbf{p}}_i \, \delta \mathbf{r}_i + \mathbf{p}_i \, \delta \dot{\mathbf{r}}_i \right) = 0 \tag{3.63}$$

Substituting $\delta \mathbf{r}_i$ from Eq.(3.60) and $\delta \dot{\mathbf{r}}_i$ from Eq. (3.61)

$$\sum_i [\dot{\mathbf{p}}_i \, (\delta \boldsymbol{\phi} \times \mathbf{r}_i) + \mathbf{p}_i \, (\delta \boldsymbol{\phi} \times \dot{\mathbf{r}}_i)] = 0$$

Using the vector relation $\mathbf{A}.(\mathbf{B} \times \mathbf{C}) = \mathbf{B}.(\mathbf{C} \times \mathbf{A}) = \mathbf{C}.(\mathbf{A} \times \mathbf{B})$, the above relation can be written as

$$\sum_i [\delta \boldsymbol{\phi}.(\mathbf{r}_i \times \dot{\mathbf{p}}_i) + \delta \boldsymbol{\phi}.(\dot{\mathbf{r}}_i \times \mathbf{p}_i)] = 0$$

$$\delta \boldsymbol{\phi}. \sum_i \frac{d}{dt} (\mathbf{r}_i \times \mathbf{p}_i) = 0$$

$$\delta \boldsymbol{\phi} \cdot \frac{d\mathbf{L}}{dt} = 0 \tag{3.64}$$

where **L** is the total angular momentum of the system. Since *df* is arbitrary

$$\frac{d\mathbf{L}}{dt} = 0$$

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} = \text{constant} \tag{3.65}$$

Thus, the rotational invariance of the Lagrangian of a closed sytem is equivalent to the conservation of total angular momentum.

Homogeneity of Time and Conservation of Energy Homogeneity in time implies that the Lagrangian of a closed system does not depend explicitly on the time t. That is, $(\partial L / \partial t) = 0$. The total time

derivative of the Lagrangian is

$$\frac{dL}{dt} = \sum_i \frac{\partial L}{\partial q_i}\dot{q}_i + \sum_i \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i + \frac{\partial L}{\partial t} \tag{3.66}$$

Use of the condition that $(\partial L/\partial t) = 0$ gives

$$\frac{dL}{dt} = \sum_i \frac{\partial L}{\partial q_i}\dot{q}_i + \sum_i \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i \tag{3.67}$$

Replacing $(\partial L/\partial q_i)$ using Lagrange's equation, we have

$$\frac{dL}{dt} = \sum_i \left(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\right)\dot{q}_i + \sum_i \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i = \sum_i \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\dot{q}_i\right) = \sum_i \frac{d}{dt}(p_i\dot{q}_i)$$

$$\frac{d}{dt}\left(\sum_i p_i\dot{q}_i - L\right) = 0 \tag{3.68}$$

That is, the quantity in parenthesis must be constant in time. Denoting the constant by H called the **Hamiltonian** of the system $\sum_i p_i\dot{q}_i - L = H$ (constant) (3.69) It can be shown that H is the total energy of the system if (i) the potential energy V is velocity-independent and (ii) the transformation equations connecting the rectangular and generalized coordinates do not depend on time explicitly. When condition (ii) is satisfied, the kinetic energy T is a homogeneous quadratic function of the generalized velocities and by Euler's theorem, Eq. (3.48) $\sum_i \frac{\partial T}{\partial \dot{q}_i}\dot{q}_i = 2T$ (3.70) Now, Eq. (3.69) can be written as

$$H = \sum_i \frac{\partial L}{\partial \dot{q}_i}\dot{q}_i - L = \sum_i \frac{\partial (T - V)}{\partial \dot{q}_i}\dot{q}_i - L$$

$$= \sum_i \frac{\partial T}{\partial \dot{q}_i}\dot{q}_i - L = 2T - (T - V)$$

$$H = T + V = E \text{ (total energy)} \tag{3.71}$$

Substituting these values in the expression for generalized force, we have

$$Q_\rho = F_x \frac{\partial x}{\partial \rho} + F_y \frac{\partial y}{\partial \rho} + F_z \frac{\partial z}{\partial \rho}$$

$$= F_x \cos\phi + F_y \sin\phi = F_\rho$$

$$Q_\phi = -F_x \rho \sin\phi + F_y \rho \cos\phi = \rho F_\phi$$

$$Q_z = F_z$$

where F_r, F_f and F_z are the components of the force along the increasing directions of r, f and z.

Example 3.5 Find Lagrange's equation of motion of the bob of a simple pendulum.

Solution: Let us select the angle q made by the string with the vertical axis as the generalized coordinate as shown in Fig. 3.3. Since l is a constant, kinetic energy of the bob $T = \frac{1}{2} m\left(l\dot{\theta}\right)^2 = \frac{1}{2} m l^2 \dot{\theta}^2$

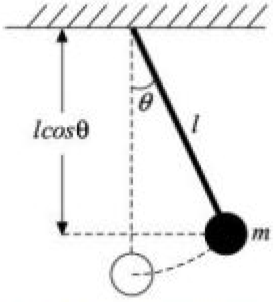


Fig. 3.3 Simple pendulum.

Taking the mean position of the bob as the reference point

The potential energy of the bob $V = mg\left(l - l\cos\theta\right)$

The Lagrangian $L = \frac{1}{2}ml^2\dot{\theta}^2 - mgl\,(1 - \cos\theta)$

$$\frac{\partial L}{\partial \dot{\theta}} = ml^2\dot{\theta} \quad \text{and} \quad \frac{\partial L}{\partial \theta} = -\,mgl\sin\theta$$

Substituting these quantities in Lagrange's equation

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

$$ml^2\ddot{\theta} + mgl\sin\theta = 0$$

$$\ddot{\theta} + \left(g/l\right)\sin\theta = 0$$

which is Lagrange's equation of motion of the bob of a simple pendulum.

Example 3.6 Obtain the equations of motion for the motion of a particle of mass *m* in a potential $V(x, y, z)$ in spherical polar coordinates.

Solution: In spherical polar coordinates the elementary lengths are *dr*, *rdq*, *r* sin*q* *df* and velocities are $\dot{r}, r\dot{\theta}, r\sin\theta\dot{\phi}$

Kinetic energy $T = \frac{1}{2}m\left(\dot{r}^2 + r^2\,\dot{\theta}^2 + r^2\sin^2\theta\,\dot{\phi}^2\right)$

Lagrangian $L = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\phi}^2\right) - V\left(r, \theta, \phi\right)$

Lagrange's equations of motion: $\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = 0 \qquad j = 1, 2, 3, \ldots$

Identifying *r*, *q* and *f* as as the generalized coordinates, the equations of motion are *r* coordinate:

$$\frac{d(m\dot{r})}{dt} - mr\left(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2\right) + \frac{\partial V}{\partial r} = 0$$

θ co-ordinate:

$$\frac{d}{dt}\left(mr^2\dot{\theta}\right) - mr^2 \sin\theta\cos\theta\,\dot{\phi}^2 + \frac{\partial V}{\partial \theta} = 0$$

ϕ co-ordinate:

$$\frac{d}{dt}\left(mr^2 \sin^2\theta\dot{\phi}\right) + \frac{\partial V}{\partial \phi} = 0$$

Note: If the force is central, $\quad \dfrac{\partial V}{\partial \theta} = \dfrac{\partial V}{\partial \phi} = 0$

Example 3.7 Masses m and $2m$ are connected by a light inextensible string which passes over a pulley of mass $2m$ and radius a. Write the Lagrangian and find the acceleration of the system.

Solution: The system has only one degree of freedom, and x (see Fig. 3.4) is taken as the generalized coordinate. The length of the string be l and the centre of the pulley is taken as zero for potential energy.

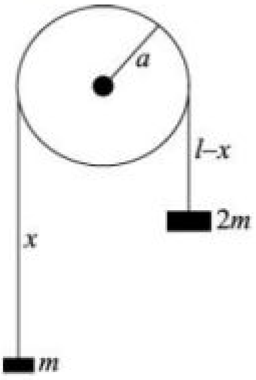


Fig. 3.4 A pulley with a string carrying masses m and $2m$ at its end.

K.E. of the system $$T = \frac{1}{2} m\dot{x}^2 + m\dot{x}^2 + \frac{1}{2} I\omega^2$$

$$= \frac{3}{2} m\dot{x}^2 + \frac{1}{2} I\left(\frac{\dot{x}}{a}\right)^2$$

P.E. of the system $V = -mgx - 2mg\,(l - x)$

Lagrangian $$L = \frac{3}{2} m\dot{x}^2 + \frac{I}{2a^2}\dot{x}^2 - mgx + 2mgl$$

$$\frac{\partial L}{\partial \dot{x}} = \left(3m + \frac{I}{a^2}\right)\dot{x} \qquad \frac{\partial L}{\partial x} = -mg$$

Substitution of these derivatives in Lagrange's equation gives the equation of motion:

$$\left(3m + \frac{I}{a^2}\right)\ddot{x} + mg = 0$$

Acceleration $$\ddot{x} = -\frac{mg}{(3m + I/a^2)} = -\frac{g}{4}$$

since moment of inertia of the disc $= \frac{1}{2} \times 2ma^2 = ma^2$. Minus sign indicates mass m moves upwards with the acceleration $g/4$.

Example 3.8 A simple pendulum has a bob of mass m with a mass m_1 at the moving support (pendulum with moving support) which moves on a horizontal line in the vertical plane in which the pendulum oscillates. Find the Lagrangian and Lagrange's equation of motion.

Solution: This pendulum (see Fig. 3.5) has two degrees of freedom, and x and q can be taken as the generalized coordinates. Taking the point of support as the

Unit - II Hamilton’s Formulation

4

Variational Principle

In Chapter 3, Lagrange's equations of motion were derived from D'Alembert's principle which is a differential principle. In this chapter the basic laws of mechanics are obtained from an integral principle known as *Hamilton's variational principle*. In this procedure, Lagrange's equations of motion are obtained from a statement about the value of the time integral of the Lagrangian between times t_1 and t_2. In D'Alembert's principle, we considered the instantaneous state of the system and virtual displacements from the instantaneous state. However, in the following variational procedure, infinitesimal virtual variations of the entire motion from the actual one is considered.

4.1 HAMILTON'S PRINCIPLE

Hamilton's principle is a variational formulation of the laws of motion in configuration space. It is considered more fundamental than Newton's equations as it can be applied to a variety of physical phenomena.

The configuration of a system at any time is defined by the values of the n generalized co-ordinates $q_1, q_2, q_3, \ldots, q_n$. This corresponds to a particular point in the n-dimensional configuration space in which the q_i's are components along the n co-ordinate axes. Hamilton's principle states: *For a conservative holonomic system, the motion of the system from its position at time t_1 to its position at time t_2 follows a path for which the line integral*

$$I = \int_{t_1}^{t_2} L(q_1, q_2, \ldots, q_n, \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n, t)\,dt = \int_{t_1}^{t_2} L(q, \dot{q}, t)\,dt \qquad (4.1)$$

has a stationary value.

That is, out of all possible paths by which the system point could travel from its position at time t_1 to its position at time t_2 in the configuration space consistent with the constraints, the path followed by the system is that for which the value of the above integral is stationary. Mathematically, the principle can be stated as:

$$\delta I = \delta \int_{t_1}^{t_2} L(q,\dot{q},t)\,dt = 0 \qquad (4.2)$$

where $q_i(t)$ and hence $\dot{q}_i(t)$ is to be varied such that $\delta q_i(t_1) = \delta q_i(t_2) = 0$. The time integral of the Lagrangian L, Eq.(4.1), is called the **action integral** or simply **action**. The *d*-variation considered here refers to the variation in a quantity at the same instant of time (see section 3.3) while the *d*-variation as usual refers to a variation in quantity along a path at different instants of time (see Fig.4.1). The two paths are infinitely close but arbitrary.

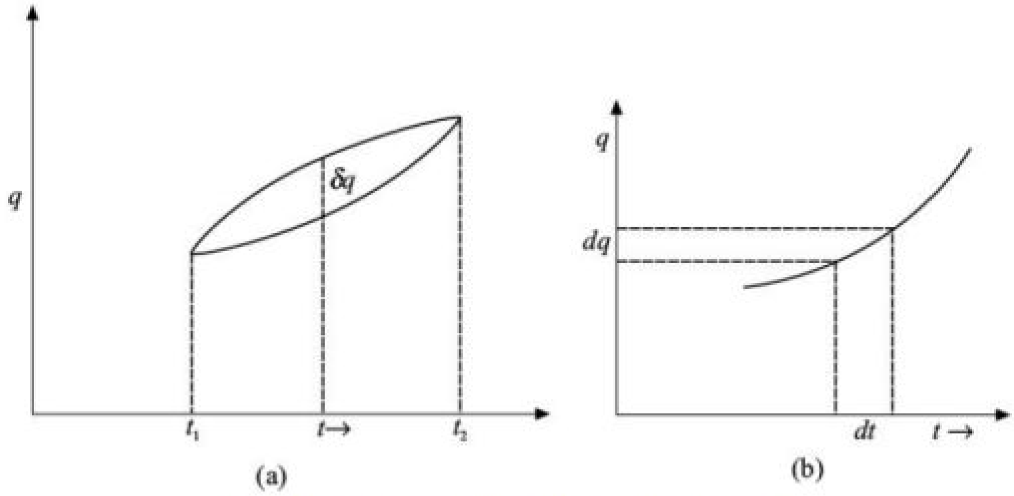


Fig. 4.1 (a) *d*-variation in motion; (b) *d*-variation in motion.

4.2 DEDUCTION OF HAMILTON'S PRINCIPLE

Hamilton's principle can be easily deduced from D'Alembert's principle given by Eq. (3.19). Consider a system of N particles of masses m_i, $i = 1, 2, \ldots, N$, located at points r_i and acted upon by external forces $\mathbf{F}_i$. According to D'Alembert's principle

$$\sum_{i=1}^{N}\left(\mathbf{F}_i - m_i\,\ddot{\mathbf{r}}_i\right).\delta\mathbf{r}_i = 0 \tag{4.3}$$

The term $\sum_i \mathbf{F}_i\,.\delta\mathbf{r}_i$ is the virtual work δW done by the external applied forces $\mathbf{F}_i$, its value is given by Eq. (3.25):

$$\delta W = \sum_{i=1}^{N}\mathbf{F}_i\,.\delta\mathbf{r}_i = \sum_{j=1}^{n} Q_j\,\delta q_j \tag{4.4}$$

where Q_j is the generalized force, defined by Eq. (3.26), and q's are the generalized co-ordinates of the system. The second term in Eq. (4.3) can be written as

$$\sum_i m_i\,\ddot{\mathbf{r}}_i\,.\delta\mathbf{r}_i = \frac{d}{dt}\left(\sum_i m_i\dot{\mathbf{r}}_i\,.\delta\mathbf{r}_i\right) - \sum_i m_i\dot{\mathbf{r}}_i\,\frac{d}{dt}(\delta\mathbf{r}_i)$$

Since virtual displacement is at the same instant of time, the order of δ and d-variations can be interchanged.

$$\begin{aligned}\sum_i m_i\,\ddot{\mathbf{r}}_i\,.\delta\mathbf{r}_i &= \frac{d}{dt}\left(\sum_i m_i\,\dot{\mathbf{r}}_i\,.\delta\mathbf{r}_i\right) - \sum_i m_i\dot{\mathbf{r}}_i\,.\delta\dot{\mathbf{r}}_i \\ &= \frac{d}{dt}\left(\sum_i m_i\dot{\mathbf{r}}_i\,.\delta\mathbf{r}_i\right) - \delta\left(\frac{1}{2}\sum_i m_i\dot{\mathbf{r}}_i\,.\dot{\mathbf{r}}_i\right)\end{aligned} \tag{4.5}$$

The second term on the right hand side of Eq. (4.5) is the d-variation of kinetic energy T. Now Eq. (4.5) takes the form

$$\sum_i m_i \ddot{\mathbf{r}}_i . \delta \mathbf{r}_i = \frac{d}{dt} \sum_i m_i \dot{\mathbf{r}}_i . \delta \mathbf{r}_i - \delta T \tag{4.6}$$

Combining Eqs. (4.3), (4.4) and (4.6)

$$\frac{d}{dt} \sum_i m_i \dot{\mathbf{r}}_i . \delta \mathbf{r}_i = \delta T + \delta W \tag{4.7}$$

Integrating with respect to time between the limits t_1 and t_2, we get

$$\int_{t_1}^{t_2} (\delta T + \delta W)\, dt = \left[\sum_i m_i \dot{\mathbf{r}}_i . \delta \mathbf{r}_i \right]_{t_1}^{t_2} \tag{4.8}$$

The right hand side is zero as $\delta \mathrm{r}_i (t_1) = \delta \mathrm{r}_i (t_2) = 0$. Replacing δW by Eq. (4.4), Eq. (4.8) reduces to

$$\int_{t_1}^{t_2} \left(\delta T + \sum_{j=1}^{n} Q_j\, \delta q_j \right) dt = 0 \tag{4.9}$$

Eq. (4.9) is sometimes referred to as *the integral form of D Alembert's principle* or *the generalized version of Hamilton's principle*. The integral form is more advantageous since it is independent of the choice of co-ordinates with which we describe the system. If the external forces are conservative, $\mathbf{F}_i = -\nabla_i V$ and by Eq. (3.33), $Q_j = (-\partial V / \partial q_j)$. Consequently,

$$\sum_j Q_j\, \delta q_j = -\sum_j \frac{\partial V}{\partial q_j} \delta q_j = -\delta V \tag{4.10}$$

and Eq. (4.9) becomes

$$\int_{t_1}^{t_2} (\delta T - \delta V) dt = 0$$

$$\int_{t_1}^{t_2} \delta (T - V)\, dt = \int_{t_1}^{t_2} \delta L\, dt = 0 \tag{4.11}$$

For a holonomic system the *d*-variation and integration can be interchanged.

Then $\delta\int_{t_1}^{t_2} L(q,\dot{q},t)\,dt = 0$ (4.12)

which is Hamilton's principle.

4.3 LAGRANGE'S EQUATION FROM HAMILTON'S PRINCIPLE

The action integral must have a stationary value for the actual path. Let us label each possible path in the configuration space by an infinitesimal parameter, say a. That is, the set of paths may be labelled by (q, a) with $q\,(t, 0)$ representing the correct path. In terms of the parameter a, each path may be written as $q_i\,(t, a) = q_i\,(t, 0) + a\,h_i\,(t)$, $i = 1, 2, 3,\ldots, n$ (4.13) where h_i(t) is a completely arbitrary well-behaved function of time with the condition $h\,(t_1) = h(t_2) = 0$. From Eq. (4.13)

$$\delta q_i = \eta_i(t)\delta\alpha \qquad (4.14)$$

As the q_i's and $\dot{q}_i$'s are only functions of t and a, for a given $h_i(t)$, the action integral I is a function of a only:

$$I\,(\alpha) = \int_{t_1}^{t_2} L\left\{q_i\,(t,\alpha),\,\dot{q}_i\,(t,\alpha),\,t\right\}dt \qquad (4.15)$$

Expanding the integrand L by Taylor series

$$I\,(\alpha) = \int_{t_1}^{t_2} \sum_i \left[L\left\{q_i\,(t,0),\,\dot{q}_i\,(t,0),t\right\} + \frac{\partial L}{\partial q_i}(\alpha\eta_i) + \frac{\partial L}{\partial \dot{q}_i}(\alpha\dot{\eta}_i)\right]dt \quad (4.16)$$

where the higher order terms in the expansion are left out, which is reasonable as a 0 . Since the integration limits t_1 and t_2 are not dependent, differentiating with respect to a under the integral sign

$$\frac{\partial I}{\partial \alpha} = \int_{t_1}^{t_2} \sum_i \left[\frac{\partial L}{\partial q_i} \eta_i + \frac{\partial L}{\partial \dot{q}_i} \dot{\eta}_i \right] dt \qquad (4.17)$$

Integrating the second term on the right hand side by parts

$$\int_{t_1}^{t_2} \frac{\partial L}{\partial \dot{q}_i} \dot{\eta}_i dt = \left[\frac{\partial L}{\partial \dot{q}_i} \eta_i \right]_{t_1}^{t_2} - \int_{t_1}^{t_2} \eta_i(t) \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i} \right) dt \qquad (4.18)$$

The integrated term vanishes since $\eta_i\ (t_1) = \eta_i(t_2) = 0$. Substituting Eq. (4.18) in Eq. (4.17) we get

$$\frac{\partial I}{\partial \alpha} = \int_{t_1}^{t_2} \sum_i \left[\frac{\partial L}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i} \right) \right] \eta_i(t)\, dt \qquad (4.19)$$

Multiplying Eq. (4.19) by $\delta\alpha$

$$\frac{\partial I}{\partial \alpha} \delta\alpha = \int_{t_1}^{t_2} \sum_i \left[\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right] \delta\alpha\, \eta_i(t)\, dt$$

Using Eq. (4.14) and remembering that the left side is simply δI.

$$\delta I = \int_{t_1}^{t_2} \sum_i \left[\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right] \delta q_i(t)\, dt \qquad (4.20)$$

For I to be stationary, $dI = 0$. Since q_i's are independent, the variations dq_i's are arbitrary and the necessary condition for the right side of Eq. (4.20) to be zero is that the coefficients of dq_i's vanish separately. Hence,

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \qquad i = 1, 2, \ldots, n \qquad (4.21)$$

which is Lagrange's equation, given by Eq. (3.38).

The above result is a special case of the more general Euler-Lagrange differential equation which determines the path $y = y(x)$ such that the line

6

Hamiltonian Mechanics

In Lagrangian formalism, generalized coordinates (q_i's) and generalized velocities ($\dot{q}_i$'s) are used as independent coordinates to formulate dynamical problems which result in second order linear differential equations. In Hamilton's formalism, generalized coordinates and generalized momenta (p_i's) are used as basic variables to formulate problems. The formulation is mainly based on the Hamiltonian function of the system which is a function of q_i's and p_i's of the system. The resulting first order linear differential equations are easier to handle mathematically. Hamilton's formalism also serves as the basis for further developments such as Hamilton – Jacobi theory and quantum mechanics. Throughout this chapter, we shall assume that the systems are holonomic and the forces are derivable from a position-dependent potential.

6.1 THE HAMILTONIAN OF A SYSTEM

The Hamiltonian H of a system, defined by Eq. (3.69), is

$$H = \sum_i p_i \dot{q}_i - L(q, \dot{q}, t) \tag{6.1}$$

where, as before, q stands for $q_1, q_2, \ldots, q_n$. Using the relation

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \tag{6.2}$$

it is possible to express $\dot{q}_i$ in terms of p_i. When this is done, we can write

$$H = H(p, q, t) \qquad q = q_1, q_2, \ldots, q_n \qquad p = p_1, p_2, \ldots, p_n \tag{6.3}$$

That is, H is expressed as a function of the generalized coordinates, generalized momenta and time. In Lagrangian formalism, the configuration space is spanned by the n generalized coordinates. Here, the q's and p's are treated in the same way and the involved space is called the **phase space**. It is a space of $2n$ variables $q_1, q_2,..., q_n, p_1, p_2,..., p_n$. Every point in the space represents both the position and momenta of all particles in the system.

As already pointed out, in general, H need not represent the total energy of the system. However, if the transformation equations connecting the cartesian and generalized coordinates do not depend on time explicitly, H is equal to the total energy of the system.

6.2 HAMILTON'S EQUATIONS OF MOTION

Hamilton's equations of motion can be derived in the following different ways: (i) From the Hamiltonian of the system (ii) From the variational principle.

In this section we shall derive them from the Hamiltonian of a system given by Eq. (6.1). Differentiating Eq. (6.1), we have

$$dH = \sum_i \left[p_i d\dot{q}_i + \dot{q}_i \, dp_i - \frac{\partial L}{\partial q_i} dq_i - \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i \right] - \frac{\partial L}{\partial t} dt \tag{6.4}$$

Since $p_i = \left(\frac{\partial L}{\partial \dot{q}_i} \right)$, the first and fourth terms on the right side of Eq. (6.4) together vanish. Hence,

$$dH = \sum_i \left(\dot{q}_i dp_i - \frac{\partial L}{\partial q_i} dq_i \right) - \frac{\partial L}{\partial t} dt \tag{6.5}$$

Taking the differential of H in Eq. (6.3)

$$dH = \sum_i \left(\frac{\partial H}{\partial q_i} dq_i + \frac{\partial H}{\partial p_i} dp_i \right) + \frac{\partial H}{\partial t} dt \tag{6.6}$$

Comparing Eqs. (6.5) and (6.6) and using Eq. (3.53), we get

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \; i = 1, 2, \ldots, n \tag{6.7}$$

$$\dot{p}_i = -\frac{\partial H}{\partial q_i}, i = 1, 2, \ldots, n \tag{6.8}$$

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \tag{6.9}$$

Equations (6.7) and (6.8) are **Hamilton's equations of motion**. They are also called the **canonical equations of motion**. They constitute a set of $2n$ first order differential equations replacing the n second order differential equations of Lagrange.

Hamilton's equations are applicable to holonomic conservative systems. If part of the forces acting on the system is not conservative, Lagrange's equations take the form

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = Q_i \tag{6.10}$$

where Q_i represents the forces not arising from a potential and L contains the conservative forces. Replacing $(\partial L/\partial \dot{q}_i)$ by p_i in Eq. (6.10), we have

$$\dot{p}_i = \frac{\partial L}{\partial q_i} + Q_i \tag{6.11}$$

In such cases, Hamilton's equations are

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \tag{6.12}$$

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} + Q_i \tag{6.13}$$

6.3 HAMILTON'S EQUATIONS FROM VARIATIONAL PRINCIPLE

Hamilton's variational principle stated in Eq. (4.2) is

$$\delta I = \delta \int_{t_1}^{t_2} L(q, \dot{q}, t)\, dt = 0 \tag{6.14}$$

Here $q_i(t)$ and hence $\dot{q}_i(t)$ is to be varied such that

$$\delta q_i\,(t_1) = \delta q_i\,(t_2) = 0 \tag{6.15}$$

which refers to paths in configuration space. In Hamilton's formalism, the integral I has to be evaluated over the trajectory of the system point in phase space, and the varied paths must be in the neighbourhood of this phase space trajectory. Therefore, to make the principle applicable to phase space trajectories, we have to express the integrand of the integral I as a function of the independent coordinates p and q and their time derivatives. This can be achieved only by replacing L in Eq. (6.14) using Eq. (6.1). We then get

$$\delta I = \delta \int_{t_1}^{t_2} \left[\sum_i p_i \dot{q}_i - H(p, q, t) \right] dt = 0 \qquad (6.16)$$

where $q(t)$ is varied subject to $\delta q_i(t_1) = \delta q_i(t_2) = 0$ and $p_i(t)$ is varied without any end-point restriction. Since the original variational principle is modified to suit phase space, it is known as **modified Hamilton's principle**. Carrying out the variations in Eq. (6.16) we have

$$\int_{t_1}^{t_2} \sum_i \left(p_i \delta \dot{q}_i + \dot{q}_i \delta p_i - \frac{\partial H}{\partial q_i} \delta q_i - \frac{\partial H}{\partial p_i} \delta p_i \right) dt = 0 \qquad (6.17)$$

We now integrate the first term in Eq. (6.17) by parts

$$\int_{t_1}^{t_2} p_i \delta \dot{q}_i \, dt = \int_{t_1}^{t_2} p_i \delta \frac{d}{dt} q_i \, dt = \int_{t_1}^{t_2} p_i \frac{d}{dt} \delta q_i \, dt$$

$$= [p_i \delta q_i]_{t_1}^{t_2} - \int_{t_1}^{t_2} \dot{p}_i \delta q_i \, dt$$

The integrated term vanishes at the end-points t_1 and t_2 and therefore

$$\int_{t_1}^{t_2} p_i \delta \dot{q}_i \, dt = - \int_{t_1}^{t_2} \dot{p}_i \delta q_i \, dt \qquad (6.18)$$

Substituting Eq. (6.18) in Eq. (6.17), we have

$$\int_{t_1}^{t_2} \sum_i \left[-\left(\dot{p}_i + \frac{\partial H}{\partial q_i} \right) \delta q_i + \left(\dot{q}_i - \frac{\partial H}{\partial p_i} \right) \delta p_i \right] dt = 0 \qquad (6.19)$$

Since the modified Hamilton's principle is a variational principle in phase space, the dq's and dp's are arbitrary and therefore the coefficients of dq_i and dp_i in Eq. (6.19) must vanish separately. Hence,

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i} \qquad i = 1, 2, \ldots, n \qquad (6.20)$$

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} = 0 \qquad (6.26)$$

$$p_i = \text{constant} = b \qquad (6.27)$$

That is, the momentum conjugate to a generalized coordinate which is cyclic is conserved.

Now, if we have a system in which the coordinates $q_1, q_2, \ldots, q_i$ are cyclic, then the Lagrangian of the system is of the form

$$L = L(q_{i+1}, \ldots, q_n, \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n, t) \qquad (6.28)$$

However, the Hamiltonian will be of the form

$$H = H(q_{i+1}, \ldots, q_n, b_1, b_2, \ldots, b_i, p_{i+1}, \ldots, p_n, t) \qquad (6.29)$$

When i cyclic coordinates are present in a system, in Lagrangian formalism the problem is still one of n degrees of freedom, whereas in Hamilton's formalism it is one of $(n - i)$ degrees of freedom. This is because even if q_j is absent in the Lagrangian, we have the equation

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} = 0 \qquad (6.30)$$

6.5 CANONICAL TRANSFORMATIONS

The transformation of one set of coordinates q_i to another set Q_i by transformation equations of the type $Q_i = Q_i(q_1, q_2, \ldots, q_n, t)$ (6.31) is called **point transformation** or **contact transformation**. What we have been doing in earlier chapters are transformations of this type. In Hamilton's formalism, the momenta are also independent variables on the same level as the generalized coordinates. Therefore, it is appropriate to have a more general type of transformation that involves both generalized coordinates and momenta. Considerable advantages will be there if the equations of motion are simpler in the new set of variables (Q, P) than in the original set (q, p). If all the coordinates are made cyclic by a transformation, the solutions will be much simpler. When there is a transformation from the original set (q, p) to the new set (Q, P), a corresponding change in the Hamiltonian $H(q, p, t)$ to a new Hamiltonian $K(Q, P, t)$ is expected. The transformation equations for the (q, p) to (Q, P) set are $Q_i = Q_i(q, p, t)$ and $P_i = P_i(q, p, t)$ (6.32) The (q, p) set obeys Hamilton's canonical equations

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \quad \text{and} \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} \tag{6.33}$$

We are interested only in those transformations that are governed by Hamilton's canonical equations

$$\dot{Q}_i = \frac{\partial K}{\partial P_i} \qquad \text{and} \qquad \dot{P}_i = -\frac{\partial K}{\partial Q_i} \tag{6.34}$$

Such transformations are called **canonical transformations.**

The original set of variables satisfy the modified Hamilton's principle, given in Eq. (6.16):

$$\delta \int_{t_1}^{t_2} \left[\sum_i p_i \dot{q}_i - H(q, p, t) \right] dt = 0 \tag{6.35}$$

The new set of variables (Q, P) should also hold the modified Hamilton's principle:

$$\delta \int_{t_1}^{t_2} \left[\sum_i P_i \dot{Q}_i - K(Q, P, t) \right] dt = 0 \tag{6.36}$$

The simultaneous validity of Eqs. (6.35) and (6.36) means that their integrands must be either equal or connected by a relation of the type

$$\sum_i p_i \dot{q}_i - H(q, p, t) = \alpha \left[\sum_i P_i \dot{Q}_i - K(Q, P, t) \right] + \frac{dF}{dt} \tag{6.37}$$

Here a is a constant independent of coordinates, momenta and time. This a is related to a simple type of scale transformation and therefore it is always possible to set $a = 1$. F is a function of the coordinates, momenta and time. The total time derivative of F in Eq. (6.37) will not contribute to the modified Hamilton's principle since

$$\delta\int_{t_1}^{t_2}\frac{dF}{dt}\,dt=\delta\int_{t_1}^{t_2}dF=\delta[\,F(t_2)-F(t_1)]=\delta F(t_2)-\delta F(t_1)=0$$

In the above, we have used the condition that the variation at the end-points is zero. This leads to

$$\sum_i p_i\dot{q}_i-H(q,p,t)=\sum_i P_i\dot{Q}_i-K(Q,P,t)+\frac{dF}{dt} \tag{6.38}$$

The left hand side of Eq. (6.38) is a function of original coordinates and momenta, and the first two terms on the right hand side depends only on the (*Q*, *P*) set. Hence, in general, *F* must be a function of the original and new variables in order for a transformation to be effected. They are 4*n* in all. Of these 4*n*, only 2*n* are independent as the 4*n* variables are connected by the 2*n* equations of constraints given by Eq. (6.32). Hence, the function *F* can be written in 4 forms: F_1 (*q*, *Q*, *t*), F_2 (*q*, *P*, *t*), F_3 (*p*, *Q*, *t*) and F_4 (*p*, *P*, *t*). The problem in question will dictate which form is to be selected. Next we consider these 4 types in detail.

Type 1 – $\mathbf{F_1}$ (q, Q, t): When the function *F* is of this form, Eq. (6.38) can be written as

$$\sum_i p_i\dot{q}_i-\sum_i P_i\dot{Q}_i+(K-H)=\sum_i\left(\frac{\partial F_1}{\partial q_i}\dot{q}_i+\frac{\partial F_1}{\partial Q_i}\dot{Q}_i\right)+\frac{\partial F_1}{\partial t} \tag{6.39}$$

Multiplying by *dt* and rearranging

$$\sum_i\left(p_i-\frac{\partial F_1}{\partial q_i}\right)dq_i-\sum_i\left(P_i-\frac{\partial F_1}{\partial Q_i}\right)dQ_i+\left(K-H-\frac{\partial F_1}{\partial t}\right)dt=0 \tag{6.40}$$

Since the original and new coordinates are separately independent, Eq. (6.40) is valid only if the coefficients of dq_i and dQ_i separately vanish. Therefore, from Eq. (6.40) we have

$$p_i=\frac{\partial F_1}{\partial q_i}(q,Q,t) \tag{6.41}$$

$$P_i = -\frac{\partial F_1}{\partial Q_i}(q, Q, t) \tag{6.42}$$

$$K = H + \frac{\partial F_1}{\partial t} \tag{6.43}$$

From Eq. (6.41) we can compute Q in terms of q and p if the arbitrary function F_1 is known: $Q = Q(q, p, t)$ (6.44) Using this value of Q in Eq. (6.42) we can compute P in terms of q and p: $P = P(q, p, t)$ (6.45) Eqs. (6.44) and (6.45) are the desired transformations from the original (q, p) to the new (Q, P) set. Eq. (6.43) gives the relation connecting the original and new Hamiltonians. Thus, we can express (Q, P) in terms of (q, p) only if the arbitrary function F_1 is known. Hence, F_1 is called the **generating function** of the transformation. If the generating function F_1 does not contain time explicitly, then $K = H$.

Type 2 – F_2 (q, P, t): Addition of the term $-\frac{d}{dt}\sum_i P_i Q_i$ to the right hand side of Eq. (6.38) will not affect the value since F_2 is arbitrary and

$$\delta\int_{t_1}^{t_2}\frac{d}{dt}\sum_i P_iQ_i\,dt=\delta\int_{t_1}^{t_2}\sum_i d\left(P_iQ_i\right)=\delta\left[\sum_i P_iQ_i\right]_{t_1}^{t_2}=0$$

It follows immediately from Eq. (6.38).

$$\sum_i p_i\dot{q}_i-H=\sum_i P_i\dot{Q}_i-K+\sum_i\frac{\partial F_2}{\partial q_i}\dot{q}_i+\sum_i\frac{\partial F_2}{\partial P_i}\dot{P}_i$$
$$+\frac{\partial F_2}{\partial t}-\sum_i P_i\dot{Q}_i-\sum_i Q_i\dot{P}_i \tag{6.46}$$

As two terms on the right hand side together vanish, multiplying by dt and rearranging

$$\sum_i\left(p_i-\frac{\partial F_2}{\partial q_i}\right)dq_i+\sum_i\left(Q_i-\frac{\partial F_2}{\partial P_i}\right)dP_i+\left(K-H-\frac{\partial F_2}{\partial t}\right)dt=0 \tag{6.47}$$

Since q_i's and P_i's are independent

$$p_i=\frac{\partial F_2}{\partial q_i}(q,P,t) \tag{6.48}$$

$$Q_i=\frac{\partial F_2}{\partial P_i}(q,P,t) \tag{6.49}$$

$$K=H+\frac{\partial F_2}{\partial t} \tag{6.50}$$

From Eq. (6.48)

$$P_i=P_i\,(q,p,t) \tag{6.51}$$

Using this result in Eq. (6.49)

$$Q_i=Q_i\,(q,p,t) \tag{6.52}$$

Eqs. (6.51) and (6.52) are the required transformation equations.

Type 3 – $F_3(p, Q, t)$: Proceeding on similar lines and adding the term $\frac{d}{dt}\sum_i p_i q_i$ to the right hand side of Eq. (6.38) and simplifying

$$P_i = -\frac{\partial F_3}{\partial Q_i}(p, Q, t) \tag{6.53}$$

$$q_i = -\frac{\partial F_3}{\partial p_i}(p, Q, t) \tag{6.54}$$

$$K = H + \frac{\partial F_3}{\partial t} \tag{6.55}$$

Type 4 – $F_4(p, P, t)$: Adding

$$-\frac{d}{dt}\sum_i P_i Q_i + \frac{d}{dt}\sum_i p_i q_i$$

to the right hand side of Eq. (6.38) and simplifying

$$q_i = -\frac{\partial F_4}{\partial p_i}(p, P, t) \tag{6.56}$$

$$Q_i = \frac{\partial F_4}{\partial P_i}(p, P, t) \tag{6.57}$$

$$K = H + \frac{\partial F_4}{\partial t} \tag{6.58}$$

In all these transformations, t is unchanged and therefore it may be regarded as an independent parameter. However, in relativistic formalism this cannot be so as space and time are treated on an equal footing. Sometimes a suitable generating function does not conform to one of the 4 types discussed above. Different combinations of the 4 types may be needed in such cases. If the generating function does not contain time explicitly, $K = H$ and Eq. (6.38) reduces to $\sum_i (p_i dq_i - P_i dQ_i) = dF$ (6.59)

Then the condition for a transformation to be canonical is that $\sum_i (p_i dq_i - P_i dQ_i)$ must be a perfect differential.

6.6 POISSON BRACKETS

6.11 THE PRINCIPLE OF LEAST ACTION

The principle of least action is another variational principle associated with the Hamiltonian formulation. It involves the type of *D*-variation discussed in Section 6.10. To prove the principle of least action, consider the action integral

$$I = \int_{t_1}^{t_2} L dt \tag{6.98}$$

The Δ-variation of I is written using Eq. (6.97):

$$\begin{aligned} \Delta I &= \Delta \int_{t_1}^{t_2} L\, dt = \int_{t_1}^{t_2} \delta L\, dt + \Delta t \int_{t_1}^{t_2} \frac{d}{dt}\,(L\, dt) \\ &= \int_{t_1}^{t_2} \delta L\, dt + \Delta t \int_{t_1}^{t_2} dL \\ &= \int_{t_1}^{t_2} \sum_i \left(\frac{\partial L}{\partial \dot{q}_i}\, \delta \dot{q}_i + \frac{\partial L}{\partial q_i}\, \delta q_i \right) dt + \left[L \Delta t \right]_{t_1}^{t_2} \end{aligned} \tag{6.99}$$

According to Lagrange's equation

$$\frac{\partial L}{\partial q_i} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i} \right) \tag{6.100}$$

An interchange in the order of differentiation gives

$$\delta \dot{q}_i = \frac{d}{dt}\, \delta q_i \tag{6.101}$$

Using Eqs. (6.100) and (6.101), the part in the parenthesis of the first term in Eq. (6.99) is

$$\frac{\partial L}{\partial \dot{q}_i}\,\delta \dot{q}_i + \frac{\partial L}{\partial q_i}\,\delta q_i = \frac{\partial L}{\partial \dot{q}_i}\,\frac{d}{dt}(\delta q_i) + \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right)\delta q_i$$

$$= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\,\delta q_i\right) = \frac{d}{dt}(p_i\,\delta q_i) \tag{6.102}$$

Applying Eq. (6.97) for the co-ordinate q_i

$$\Delta q_i = \delta q_i + \Delta t\,\frac{d}{dt}\,q_i$$

$$\delta q_i = \Delta q_i - \Delta t \dot{q}_i$$

Multiplying by p_i

$$p_i \delta q_i = p_i \Delta q_i - p_i \Delta t\, \dot{q}_i \tag{6.103}$$

Substituting this value of $p_i\,\delta q_i$ in Eq. (6.102), we have

$$\frac{\partial L}{\partial \dot{q}_i}\,\delta \dot{q}_i + \frac{\partial L}{\partial q_i}\,\delta q_i = \frac{d}{dt}(p_i \Delta q_i) - \frac{d}{dt}(p_i\,\Delta t\,\dot{q}_i) \tag{6.104}$$

Combining Eqs. (6.99) and (6.104)

$$\Delta I = \Delta \int_{t_1}^{t_2} L\,dt = \int_{t_1}^{t_2} \sum_i \left[\frac{d}{dt}(p_i \Delta q_i) - \frac{d}{dt}(p_i \Delta t \dot{q}_i)\right] dt + \left[L\,\Delta t\right]_{t_1}^{t_2}$$

$$= \sum_i \left[p_i\;\Delta q_i\right]_{t_1}^{t_2} - \left[p_i \dot{q}_i \Delta t\right]_{t_1}^{t_2} + \left[L \Delta t\right]_{t_1}^{t_2} \tag{6.105}$$

As $\Delta q_i = 0$ at the end-points, the first term on the right of Eq. (6.105) is zero. Hence,

$$\Delta \int_{t_1}^{t_2} L dt = \left[\left(L - \sum_i p_i \dot{q}_i \right) \Delta t \right]_{t_1}^{t_2} \tag{6.106}$$

Since $H = \sum_i p_i \dot{q}_i - L$, Eq. (6.106) reduces to

$$\Delta \int_{t_1}^{t_2} L\, dt = -\left[H\, \Delta t \right]_{t_1}^{t_2} \tag{6.107}$$

Restricting to systems for which $H = E$ $\left(\text{conservative and} \dfrac{\delta H}{\delta t} = 0 \right)$ or $\Delta H = 0$, we have

$$\Delta \int_{t_1}^{t_2} H\, dt = \int_{t_1}^{t_2} \Delta H\, dt + \int_{t_1}^{t_1} H \Delta\, (dt) = \int_{t_1}^{t_2} H d(\Delta t)$$

$$= \left[H\, \Delta t \right]_{t_1}^{t_2} \tag{6.108}$$

Combining Eqs. (6.107) and (6.108)

$$\Delta \int_{t_1}^{t_2} L\, dt = -\Delta \int_{t_1}^{t_2} H\, dt \quad \text{or} \quad \Delta \int_{t_1}^{t_2} (L + H)\, dt = 0$$

$$\Delta \int_{t_1}^{t_2} \sum_i \dot{q}_i p_i\, dt = 0 \tag{6.109}$$

which is the **principle of least action.**

Different Forms of Least Action Principle The principle of least action can be expressed in different forms. If the transformation equations do not depend on time explicitly, then the kinetic energy is a quadratic function of the generalized velocities. In such a case from Eq. (3.47) we have

$$\sum_i p_i \dot{q}_i = \sum_i \frac{\partial T}{\partial \dot{q}_i} \dot{q}_i = 2T \tag{6.110}$$

Equation (6.109) now reduces to

$$\Delta \int_{t_1}^{t_2} T \, dt = 0 \tag{6.111}$$

This is another form of the principle of least action. Further, if there is no external force on the system, T and H are conserved and the principle of least action takes the form

$$\Delta \int_{t_1}^{t_2} dt = 0 \quad \text{or} \quad \Delta(t_2 - t_1) = 0 \tag{6.112a}$$

$$t_2 - t_1 = \text{an extremum} \tag{6.112b}$$

That is, of all paths possible between two points that are consistent with the conservation of energy, the system moves along the path for which the time of transit is the least. In this form, the principle is similar to **Fermat's principle** in geometrical optics, which states that a light ray travels between two points along such a path that the time taken is the least.

Again, when the transformation equations do not involve time, the kinetic energy is given by Eq. (3.45):

$$T = \sum_j \sum_k a_{jk}\, \dot{q}_j\, \dot{q}_k \tag{6.113}$$

A configuration space for which the a_{ij} coefficients form the metric tensor can be constructed. The element of path length $d\rho$ in this space is defined by

$$(d\rho)^2 = \sum_j \sum_k a_{jk}\, dq_j\, dq_k \tag{6.114}$$

$$\left(\frac{d\rho}{dt}\right)^2 = \sum_j \sum_k a_{jk}\, \dot{q}_j\, \dot{q}_k \tag{6.115}$$

From Eqs. (6.113) and (6.115)

$$T = \left(\frac{d\rho}{dt}\right)^2 \quad \text{or} \quad dt = \frac{d\rho}{\sqrt{T}} \tag{6.116}$$

Equation (6.116) helps us to change the variable in Eq. (6.111) and the principle of least action takes the form

Equation (6.116) helps us to change the variable in Eq. (6.111) and the principle of least action takes the form

$$\Delta \int_{t_1}^{t_2} T dt = \Delta \int_{t_1}^{t_2} \sqrt{T}\ d\rho = 0 \tag{6.117}$$

For conservative systems, $H = T + V$. Consequently, Eq. (6.117) becomes

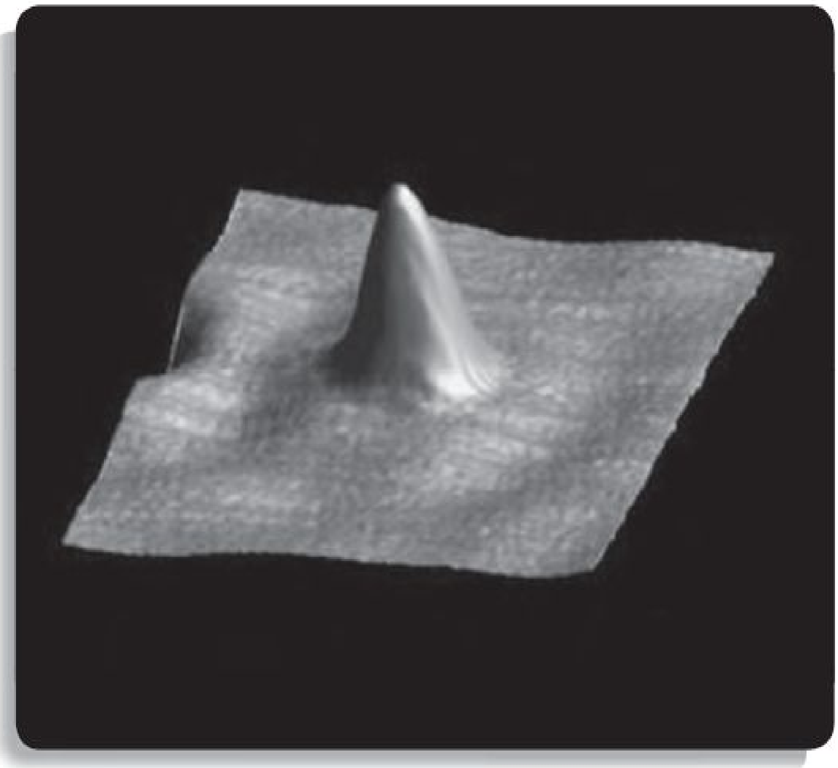
$$\Delta \int_{t_1}^{t_2} \sqrt{H - V(q)}\ d\rho = 0 \tag{6.118}$$

Eq. (6.118) is often referred to as **Jacobi's form of least action principle**. It now refers to the path of the system in a curvilinear configuration space characterized by the metric tensor with elements a_{jk}.

Unit - III Dual Nature of Matter

Chapter 4

THE WAVELIKE PROPERTIES OF PARTICLES



Just as we produce images from light waves that scatter from objects, we can also form images from "particle waves". The electron microscope produces images from electron waves that enable us to visualize objects on a scale that is much smaller than the wavelength of light. The ability to observe individual human cells and even sub-cellular objects such as chromosomes has revolutionized our understanding of biological processes. It is even possible to form images of a single atom, such as this cobalt atom on a gold surface. The ripples on the surface show electrons from gold atoms reacting to the presence of the intruder.

In classical physics, the laws describing the behavior of waves and particles are fundamentally different. Projectiles obey particle-type laws, such as Newtonian mechanics. Waves undergo interference and diffraction, which cannot be explained by the Newtonian mechanics associated with particles. The energy carried by a particle is confined to a small region of space; a wave, on the other hand, distributes its energy throughout space in its wavefronts. In describing the behavior of a particle we often want to specify its location, but this is not so easy to do for a wave. How would you describe the exact location of a sound wave or a water wave?

In contrast to this clear distinction found in classical physics, quantum physics requires that particles sometimes obey the rules that we have previously established for waves, and we shall use some of the language associated with waves to describe particles. The system of mechanics associated with quantum systems is sometimes called "wave mechanics" because it deals with the wavelike behavior of particles. In this chapter we discuss the experimental evidence in support of this wavelike behavior for particles such as electrons.

As you study this chapter, notice the frequent references to such terms as the *probability* of the outcome of a measurement, the *average* of many repetitions of a measurement, and the *statistical* behavior of a system. These terms are fundamental to quantum mechanics, and you cannot begin to understand quantum behavior until you feel comfortable with discarding such classical notions as fixed trajectories and certainty of outcome, while substituting the quantum mechanical notions of probability and statistically distributed outcomes.

4.1 DE BROGLIE'S HYPOTHESIS

Louis de Broglie (1892–1987, France). A member of an aristocratic family, his work contributed substantially to the early development of the quantum theory.

Progress in physics often can be characterized by long periods of experimental and theoretical drudgery punctuated occasionally by flashes of insight that cause profound changes in the way we view the universe. Frequently the more profound the insight and the bolder the initial step, the simpler it seems in historical perspective, and the more likely we are to sit back and wonder, "Why didn't I think of that?" Einstein's special theory of relativity is one example of such insight; the hypothesis of the Frenchman Louis de Broglie is another.*

In the previous chapter we discussed the double-slit experiment (which can be understood only if light behaves as a wave) and the photoelectric and Compton effects (which can be understood only if light behaves as a particle). Is this dual particle-wave nature a property only of light or of material objects as well? In a bold and daring hypothesis in his 1924 doctoral dissertation, de Broglie chose the latter alternative. Examining Eq. 3.20, $E = hf$, and Eq. 3.22, $p = h/\lambda$, we find some difficulty in applying the first equation in the case of particles, for we cannot be sure whether E should be the kinetic energy, total energy, or total relativistic energy (all, of course, are identical for light). No such difficulties arise from the second relationship. De Broglie suggested, lacking any experimental evidence in

---

*De Broglie's name should be pronounced "deh-BROY" or "deh-BROY-eh," but it is often said as "deh-BROH-lee."

support of his hypothesis, that associated with any material particle moving with momentum p there is a wave of wavelength λ, related to p according to

$$\lambda = \frac{h}{p} \tag{4.1}$$

where h is Planck's constant. The wavelength λ of a particle computed according to Eq. 4.1 is called its *de Broglie wavelength*.

Example 4.1

Compute the de Broglie wavelength of the following: (*a*) A 1000-kg automobile traveling at 100 m/s (about 200 mi/h). (*b*) A 10-g bullet traveling at 500 m/s. (*c*) A smoke particle of mass 10^{-9} g moving at 1 cm/s. (*d*) An electron with a kinetic energy of 1 eV. (*e*) An electron with a kinetic energy of 100 MeV.

Solution

(*a*) Using the classical relation between velocity and momentum,

$$\lambda = \frac{h}{p} = \frac{h}{mv} = \frac{6.6 \times 10^{-34}\ \mathrm{J \cdot s}}{(10^3\ \mathrm{kg})(100\ \mathrm{m/s})} = 6.6 \times 10^{-39}\ \mathrm{m}$$

(*b*) As in part (*a*),

$$\lambda = \frac{h}{mv} = \frac{6.6 \times 10^{-34}\ \mathrm{J \cdot s}}{(10^{-2}\ \mathrm{kg})(500\ \mathrm{m/s})} = 1.3 \times 10^{-34}\ \mathrm{m}$$

(*c*)

$$\lambda = \frac{h}{mv} = \frac{6.6 \times 10^{-34}\ \mathrm{J \cdot s}}{(10^{-12}\ \mathrm{kg})(10^{-2}\ \mathrm{m/s})} = 6.6 \times 10^{-20}\ \mathrm{m}$$

(*d*) The rest energy (mc^2) of an electron is 5.1×10^5 eV. Because the kinetic energy (1 eV) is much less than the rest energy, we can use nonrelativistic kinematics.

$$\begin{aligned} p &= \sqrt{2mK} \\ &= \sqrt{2(9.1 \times 10^{-31}\ \mathrm{kg})(1\ \mathrm{eV})(1.6 \times 10^{-19}\ \mathrm{J/eV})} \\ &= 5.4 \times 10^{-25}\ \mathrm{kg \cdot m/s} \end{aligned}$$

Then,

$$\begin{aligned} \lambda = \frac{h}{p} &= \frac{6.6 \times 10^{-34}\ \mathrm{J \cdot s}}{5.4 \times 10^{-25}\ \mathrm{kg \cdot m/s}} \\ &= 1.2 \times 10^{-9}\ \mathrm{m} = 1.2\ \mathrm{nm} \end{aligned}$$

We can also find this solution in the following way, using $p = \sqrt{2mK}$ and $hc = 1240\ \mathrm{eV \cdot nm}$.

$$\begin{aligned} cp &= c\sqrt{2mK} = \sqrt{2(mc^2)K} \\ &= \sqrt{2(5.1 \times 10^5\ \mathrm{eV})(1\ \mathrm{eV})} = 1.0 \times 10^3\ \mathrm{eV} \\ \lambda &= \frac{h}{p} = \frac{hc}{pc} = \frac{1240\ \mathrm{eV \cdot nm}}{1.0 \times 10^3\ \mathrm{eV}} = 1.2\ \mathrm{nm} \end{aligned}$$

This method may seem artificial at first, but with practice it becomes quite useful, especially because energies are usually given in electron-volts in atomic and nuclear physics.

(*e*) In this case, the kinetic energy is much greater than the rest energy, and so we are in the extreme relativistic realm, where $K \cong E \cong pc$, as in Eq. 2.40. The wavelength is

$$\lambda = \frac{hc}{pc} = \frac{1240\ \mathrm{MeV \cdot fm}}{100\ \mathrm{MeV}} = 12\ \mathrm{fm}$$

Note that the wavelengths computed in parts (*a*), (*b*), and (*c*) are far too small to be observed in the laboratory. Only in the last two cases, in which the wavelength is of the same order as atomic or nuclear sizes, do we have any chance of observing the wavelength. *Because of the smallness of h, only for particles of atomic or nuclear size will the wave behavior be observable.*

Two questions immediately follow. First, just what sort of wave is it that has this de Broglie wavelength? That is, what does the amplitude of the de Broglie wave measure? We'll discuss the answer to this question later in this chapter. For now, we assume that, associated with the particle as it moves, there is a de Broglie wave of wavelength λ, which shows itself *when a wave-type experiment (such as diffraction) is performed on it.* The outcome of the wave-type experiment depends on this wavelength. The de Broglie wavelength, which characterizes the wave-type behavior of particles, is central to the quantum theory.

The second question then occurs: Why was this wavelength not directly observed before de Broglie's time? As parts (*a*), (*b*), and (*c*) of Example 4.1 showed, for ordinary objects the de Broglie wavelength is very small. Suppose we tried to demonstrate the wave nature of these objects through a double-slit type of experiment. Recall from Eq. 3.16 that the spacing between adjacent fringes in a double-slit experiment is $\Delta y = \lambda D/d$. Putting in reasonable values for the slit separation d and slit-to-screen distance D, you will find that there is no achievable experimental configuration that can produce an observable separation of the fringes (see Problem 9). *There is no experiment that can be done to reveal the wave nature of macroscopic (laboratory-sized) objects.* Experimental verification of de Broglie's hypothesis comes only from experiments with objects on the atomic scale, which are discussed in the next section.

4.2 EXPERIMENTAL EVIDENCE FOR DE BROGLIE WAVES

The indications of wave behavior come mostly from interference and diffraction experiments. Double-slit interference, which was reviewed in Section 3.1, is perhaps the most familiar type of interference experiment, but the experimental difficulties of constructing double slits to do interference experiments with beams of atomic or subatomic particles were not solved until long after the time of de Broglie's hypothesis. We discuss these experiments later in this section. First we'll discuss diffraction experiments with electrons.

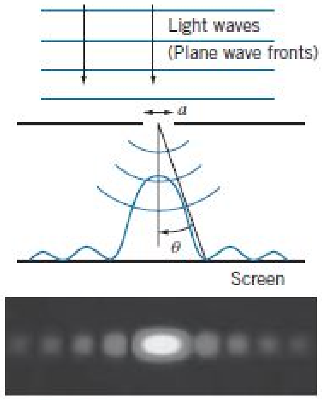


FIGURE 4.1 Light waves (represented as plane wave fronts) are incident on a narrow slit of width a. Diffraction causes the waves to spread after passing through the slit, and the intensity varies along the screen. The photograph shows the resulting intensity pattern.

Particle Diffraction Experiments

Diffraction of light waves is discussed in most introductory physics texts and is illustrated in Figure 4.1 for light diffracted by a single slit. For light of wavelength λ incident on a slit of width a, the diffraction minima are located at angles given by

$$a \sin\theta = n\lambda \qquad n = 1, 2, 3, \ldots \tag{4.2}$$

on either side of the central maximum. Note that most of the light intensity falls in the central maximum.

The experiments that first verified de Broglie's hypothesis involve *electron diffraction,* not through an artificially constructed single slit (as for the diffraction pattern in Figure 4.1) but instead through the atoms of a crystal. The outcomes of these experiments resemble those of the similar X-ray diffraction experiments illustrated in Section 3.1.

In an electron diffraction experiment, a beam of electrons is accelerated from rest through a potential difference ΔV, acquiring a nonrelativistic kinetic energy $K = e\,\Delta V$ and a momentum $p = \sqrt{2mK}$. Wave mechanics would describe the beam of electrons as a wave of wavelength $\lambda = h/p$. The beam strikes a crystal, and the scattered beam is photographed (Figure 4.2). The similarity between electron diffraction patterns (Figure 4.2) and X-ray diffraction patterns (Figure 3.7) strongly suggests that the electrons are behaving as waves.

The "rings" produced in X-ray diffraction of polycrystalline materials (Figure 3.8*b*) are also produced in electron diffraction, as shown in Figure 4.3, again providing strong evidence for the similarity in the wave behavior of electrons and X rays. Experiments of the type illustrated in Figure 4.3 were first done in 1927 by G. P. Thomson, who shared the 1937 Nobel Prize for this work. (Thomson's father, J. J. Thomson, received the 1906 Nobel Prize for his discovery of the electron and measurement of its charge-to-mass ratio. Thus it can be said that Thomson, the father, discovered the particle nature of the electron, while Thomson, the son, discovered its wave nature.)

An electron diffraction experiment gave the first experimental confirmation of the wave nature of electrons (and the quantitative confirmation of the de Broglie relationship $\lambda = h/p$) soon after de Broglie's original hypothesis. In 1926, at the Bell Telephone Laboratories, Clinton Davisson and Lester Germer were investigating the reflection of electron beams from the surface of nickel crystals. A schematic view of their apparatus is shown in Figure 4.4. A beam of electrons from a heated filament is accelerated through a potential difference ΔV. After passing through a small aperture, the beam strikes a single crystal of nickel. Electrons are scattered in all directions by the atoms of the crystal, some of them striking a detector, which can be moved to any angle ϕ relative to the incident beam and which measures the intensity of the electron beam scattered at that angle.

Figure 4.5 shows the results of one of the experiments of Davisson and Germer. When the accelerating voltage is set at 54 V, there is an intense reflection of the beam at the angle $\phi = 50°$. Let's see how these results give confirmation of the de Broglie wavelength.

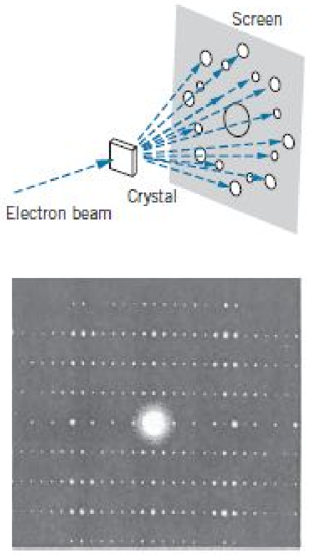


FIGURE 4.2 (Top) Electron diffraction apparatus. (Bottom) Electron diffraction pattern. Each bright dot is a region of constructive interference, as in the X-ray diffraction patterns of Figure 3.7. The target is a crystal of $Ti_2Nb_{10}O_{29}$.

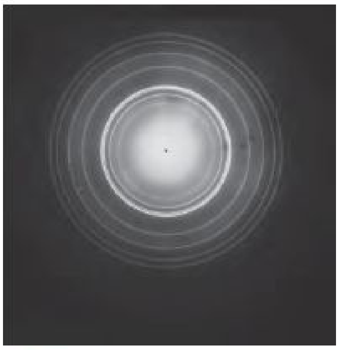


FIGURE 4.3 Electron diffraction of polycrystalline beryllium. Note the similarity between this pattern and the pattern for X-ray diffraction of a polycrystalline material (Figure 3.8*b*).

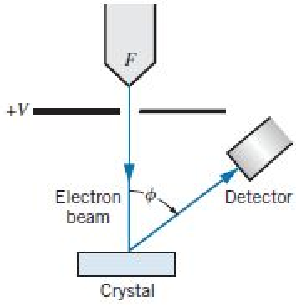


FIGURE 4.4 Apparatus used by Davisson and Germer to study electron diffraction. Electrons leave the filament F and are accelerated by the voltage V. The beam strikes a crystal and the scattered beam is detected at an angle ϕ relative to the incident beam. The detector can be moved in the range 0 to 90°.

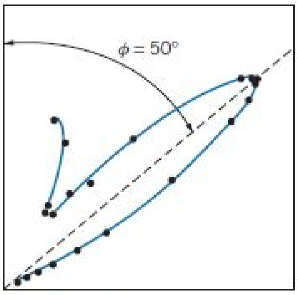


FIGURE 4.5 Results of Davisson and Germer. Each point on the plot represents the relative intensity when the detector in Figure 4.4 is located at the corresponding angle ϕ measured from the vertical axis. Constructive interference causes the intensity of the reflected beam to reach a maximum at $\phi = 50°$ for $V = 54$ V.

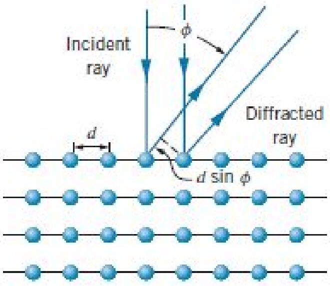


FIGURE 4.6 The crystal surface acts like a diffraction grating with spacing d.

Each of the atoms of the crystal can act as a scatterer, so the scattered *electron waves* can interfere, and we have a crystal diffraction grating for the electrons. Figure 4.6 shows a simplified representation of the nickel crystal used in the Davisson-Germer experiment. Because the electrons were of low energy, they did not penetrate very far into the crystal, and it is sufficient to consider the diffraction to take place in the plane of atoms on the surface. The situation is entirely similar to using a reflection-type diffraction grating for light; the spacing d between the rows of atoms on the crystal is analogous to the spacing between the slits in the optical grating. The maxima for a diffraction grating occur at angles ϕ such that the path difference between adjacent rays $d \sin \phi$ is equal to a whole number of wavelengths:

$$d \sin \phi = n\lambda \qquad n = 1, 2, 3, \ldots \tag{4.3}$$

where n is the order number of the maximum.

From independent data, it is known that the spacing between the rows of atoms in a nickel crystal is $d = 0.215\,\text{nm}$. The peak at $\phi = 50^\circ$ must be a first-order peak ($n = 1$), because no peaks were observed at smaller angles. If this is indeed an interference maximum, the corresponding wavelength is, from Eq. 4.3,

$$\lambda = d \sin \phi = (0.215\,\text{nm})(\sin 50^\circ) = 0.165\,\text{nm}$$

FIGURE 4.7 Diffraction of neutrons by a sodium chloride crystal.

We can compare this value with that expected on the basis of the de Broglie theory. An electron accelerated through a potential difference of 54 V has a kinetic energy of 54 eV and therefore a momentum of

$$p = \sqrt{2mK} = \frac{1}{c}\sqrt{2mc^2K} = \frac{1}{c}\sqrt{2(511{,}000\,\text{eV})(54\,\text{eV})} = \frac{1}{c}(7430\,\text{eV})$$

The de Broglie wavelength is $\lambda = h/p = hc/pc$. Using $hc = 1240\,\text{eV} \cdot \text{nm}$,

$$\lambda = \frac{hc}{pc} = \frac{1240\,\text{eV} \cdot \text{nm}}{7430\,\text{eV}} = 0.167\,\text{nm}$$

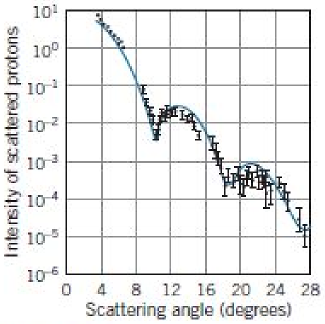


FIGURE 4.8 Diffraction of 1-GeV protons by oxygen nuclei. The pattern of maxima and minima is similar to that of single-slit diffraction of light waves. [*Source:* H. Palevsky et al., *Physical Review Letters* **18**, 1200 (1967).]

This is in excellent agreement with the value found from the diffraction maximum, and provides strong evidence in favor of the de Broglie theory. For this experimental work, Davisson shared the 1937 Nobel Prize with G. P. Thomson.

The wave nature of particles is not exclusive to electrons; *any* particle with momentum p has de Broglie wavelength h/p. Neutrons are produced in nuclear reactors with kinetic energies corresponding to wavelengths of roughly 0.1 nm; these also should be suitable for diffraction by crystals. Figure 4.7 shows that diffraction of neutrons by a salt crystal produces the same characteristic patterns as the diffraction of electrons or X rays. Clifford Shull shared the 1994 Nobel Prize for the development of the neutron diffraction technique.

To study the nuclei of atoms, much smaller wavelengths are needed, of the order of 10^{-15} m. Figure 4.8 shows the diffraction pattern produced by the scattering of 1-GeV kinetic energy protons by oxygen nuclei. Maxima and minima of the diffracted intensity appear in a pattern similar to the single-slit diffraction shown in Figure 4.1. (The intensity at the minima does not fall to zero because nuclei do not have a sharp boundary. The determination of nuclear sizes from such diffraction patterns is discussed in Chapter 12.)

Example 4.2

Protons of kinetic energy 1.00 GeV were diffracted by oxygen nuclei, which have a radius of 3.0 fm, to produce the data shown in Figure 4.8. Calculate the expected angles where the first three diffraction minima should appear.

Solution

The total relativistic energy of the protons is $E = K + mc^2 = 1.00\,\text{GeV} + 0.94\,\text{GeV} = 1.94\,\text{GeV}$ is, so their momentum is

$$p = \frac{1}{c}\sqrt{E^2 - (mc^2)^2}$$
$$= \frac{1}{c}\sqrt{(1.94\,\text{GeV})^2 - (0.94\,\text{GeV})^2} = 1.70\,\text{GeV}/c$$

The corresponding de Broglie wavelength is

$$\lambda = \frac{h}{p} = \frac{hc}{pc} = \frac{1240\,\text{MeV}\cdot\text{fm}}{1700\,\text{MeV}} = 0.73\,\text{fm}$$

We can represent the oxygen nuclei as circular disks, for which the diffraction formula is a bit different from Eq. 4.2: $a \sin\theta = 1.22n\lambda$, where a is the diameter of the diffracting object. Based on this formula, the first diffraction minimum ($n = 1$) should appear at the angle

$$\sin\theta = \frac{1.22n\lambda}{a} = \frac{(1.22)(1)(0.73\,\text{fm})}{6.0\,\text{fm}} = 0.148$$

or $\theta = 8.5^\circ$. Because the sine of the diffraction angle is proportional to the index n, the $n = 2$ minimum should appear at the angle where $\sin\theta = 2 \times 0.148 = 0.296$ ($\theta = 17.2^\circ$), and the $n = 3$ minimum where $\sin\theta = 3 \times 0.148 = 0.444$ ($\theta = 26.4^\circ$).

From the data in Figure 4.8, we see the first diffraction minimum at an angle of about 10°, the second at about 18°, and the third at about 27°, all in very good agreement with the expected values. The data don't exactly follow the formula for diffraction by a disk, because nuclei don't behave quite like disks. In particular, they have diffuse rather than sharp edges, which prevents the intensity at the diffraction minima from falling to zero and also alters slightly the locations of the minima.

Double-Slit Experiments with Particles

The definitive evidence for the wave nature of *light* was deduced from the double-slit experiment performed by Thomas Young in 1801 (discussed in Section 3.1). In principle, it should be possible to do double-slit experiments with *particles* and thereby directly observe their wavelike behavior. However, the technological difficulties of producing double slits for particles are formidable, and such experiments did not become possible until long after the time of de Broglie. The first double-slit experiment with electrons was done in 1961. A diagram of the apparatus is shown in Figure 4.9. The electrons from a hot filament were accelerated through 50 kV (corresponding to $\lambda = 5.4$ pm) and then passed through a double slit of separation 2.0 μm and width 0.5 μm. A photograph of the resulting intensity pattern is shown in Figure 4.10. The similarity with the double-slit pattern for light (Figure 3.2) is striking.

A similar experiment can be done for neutrons. A beam of neutrons from a nuclear reactor can be slowed to a room-temperature "thermal" energy distribution (average $K \approx kT \approx 0.025$ eV), and a specific wavelength can be selected by a scattering process similar to Bragg diffraction (see Eq. 3.18 and Problem 32 at the end of the present chapter). In one experiment, neutrons of kinetic energy 0.00024 eV and de Broglie wavelength 1.85 nm passed through a gap of diameter 148 μm in a material that absorbs virtually all of the neutrons incident on it (Figure 4.11). In the center of the gap was a boron wire (also highly absorptive for neutrons) of diameter 104 μm. The neutrons could pass on either side of the wire through slits of width 22 μm. The intensity of neutrons that pass through this double slit was observed by sliding

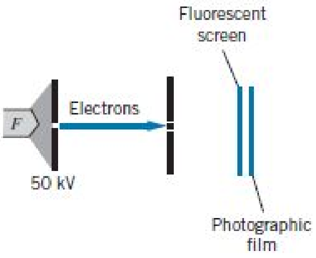


FIGURE 4.9 Double-slit apparatus for electrons. Electrons from the filament F are accelerated through 50 kV and pass through the double slit. They produce a visible pattern when they strike a fluorescent screen (like a TV screen), and the resulting pattern is photographed. A photograph is shown in Figure 4.10. [See C. Jonsson, *American Journal of Physics* **42**, 4 (1974).]

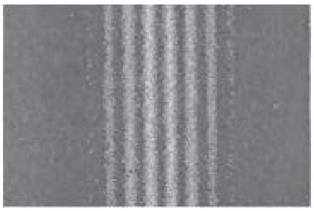


FIGURE 4.10 Double-slit interference pattern for electrons.

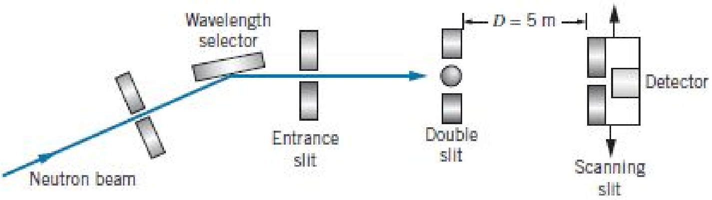


FIGURE 4.11 Double-slit apparatus for neutrons. Thermal neutrons from a reactor are incident on a crystal; scattering through a particular angle selects the energy of the neutrons. After passing through the double slit, the neutrons are counted by the scanning slit assembly, which moves laterally.

another slit across the beam and measuring the intensity of neutrons passing through this "scanning slit." Figure 4.12 shows the resulting pattern of intensity maxima and minima, which leaves no doubt that interference is occurring and that the neutrons have a corresponding wave nature. The wavelength can be deduced from the slit separation using Eq. 3.16 to obtain the spacing between adjacent maxima, $\Delta y = y_{n+1} - y_n$. Estimating the spacing Δy from Figure 4.12 to be about 75 μm, we obtain

$$\lambda = \frac{d\Delta y}{D} = \frac{(126\,\mu\text{m})(75\,\mu\text{m})}{5\,\text{m}} = 1.89\,\text{nm}$$

This result agrees very well with the de Broglie wavelength of 1.85 nm selected for the neutron beam.

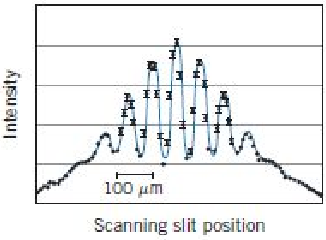


FIGURE 4.12 Intensity pattern observed for double-slit interference with neutrons. The spacing between the maxima is about 75 μm. [*Source:* R. Gahler and A. Zeilinger, *American Journal of Physics* 59, 316 (1991).]

It is also possible to do a similar experiment with atoms. In this case, a source of helium atoms formed a beam (of velocity corresponding to a kinetic energy of 0.020 eV) that passed through a double slit of separation 8 μm and width 1 μm. Again a scanning slit was used to measure the intensity of the beam passing through the double slit. Figure 4.13 shows the resulting intensity pattern. Although the results are not as dramatic as those for electrons and neutrons, there is clear evidence of interference maxima and minima, and the separation of the maxima gives a wavelength that is consistent with the de Broglie wavelength (see Problem 8).

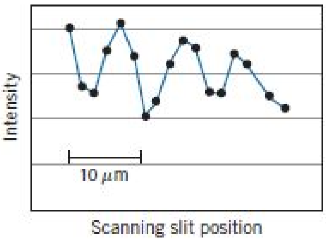


FIGURE 4.13 Intensity pattern observed for double-slit interference with helium atoms. [*Source:* O. Carnal and J. Mlynek, *Physical Review Letters* 66, 2689 (1991).]

Diffraction can be observed with even larger objects. Figure 4.14 shows the pattern produced by fullerene molecules (C_{60}) in passing through a diffraction grating with a spacing of $d = 100$ nm. The diffraction pattern was observed at a distance of 1.2 m from the grating. Estimating the separation of the maxima in Figure 4.14 as 50 μm, we get the angular separation of the maxima to be $\theta \approx \tan\theta = (50\,\mu\text{m})/(1.2\,\text{m}) = 4.2 \times 10^{-5}$ rad, and thus $\lambda = d\sin\theta = 4.2$ pm. For C_{60} molecules with a speed of 117 m/s used in this experiment, the expected de Broglie wavelength is 4.7 pm, in good agreement with our estimate from the diffraction pattern.

In this chapter we have discussed several interference and diffraction experiments using different particles—electrons, protons, neutrons, atoms, and molecules. These experiments are not restricted to any particular type of particle or to any particular type of observation. They are examples of a *general* phenomenon, the wave nature of particles, that was unobserved before 1920 because the necessary experiments had not yet been done. Today this wave nature is used as a basic tool by scientists. For example, neutron diffraction

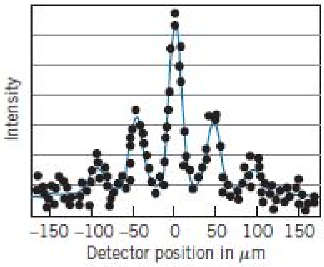


FIGURE 4.14 Diffraction grating pattern produced by C_{60} molecules. [*Source:* O. Nairz, M. Arndt, and A. Zeilinger, *American Journal of Physics* **71**, 319 (2003).]

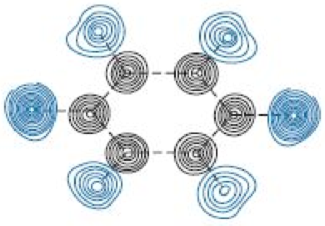


FIGURE 4.15 The atomic structure of solid benzene as deduced from neutron diffraction. The circles indicate contours of constant density. The black circles show the locations of the six carbon atoms that form the familiar benzene ring. The blue circles show the locations of the hydrogen atoms.

FIGURE 4.16 Electron microscope image of bacteria on the surface of a human tongue. The magnification here is about a factor of 5000.

gives detailed information on the structure of solid crystals and of complex molecules (Figure 4.15). The electron microscope uses electron waves to illuminate and form an image of objects; because the wavelength can be made thousands of times smaller than that of visible light, it is possible to resolve and observe small details that are not observable with visible light (Figure 4.16).

Through Which Slit Does the Particle Pass?

When we do a double-slit experiment with particles such as electrons, it is tempting to try to determine through which slit the particle passes. For example, we could surround each slit with an electromagnetic loop that causes a meter to deflect whenever a charged particle or perhaps a particle with a magnetic moment passes through the loop (Figure 4.17). If we fired the particles through the slits at a slow enough rate, we could track each particle as it passed through one slit or the other and then appeared on the screen.

If we performed this imaginary experiment, the result would no longer be an interference pattern on the screen. Instead, we would observe a pattern similar to that shown in Figure 4.17, with "hits" in front of each slit, but no interference fringes. No matter what sort of device we use to determine through which slit the particle passes, the interference pattern will be destroyed. The classical *particle* must pass through one slit or the other; only a *wave* can reveal interference, which depends on parts of the wavefront passing through *both* slits and then recombining.

When we ask through which slit the particle passed, we are investigating only the *particle* aspects of its behavior, and we cannot observe its wave nature (the interference pattern). Conversely, when we study the wave nature, we cannot simultaneously observe the particle nature. The electron will behave as a particle *or* a wave, but we cannot observe *both* aspects of its behavior simultaneously. This curious aspect of quantum mechanics was also discussed for photons in Section 3.6, where we discovered that experiments can reveal either the particle nature of the photon or its wave nature, but not both aspects simultaneously.

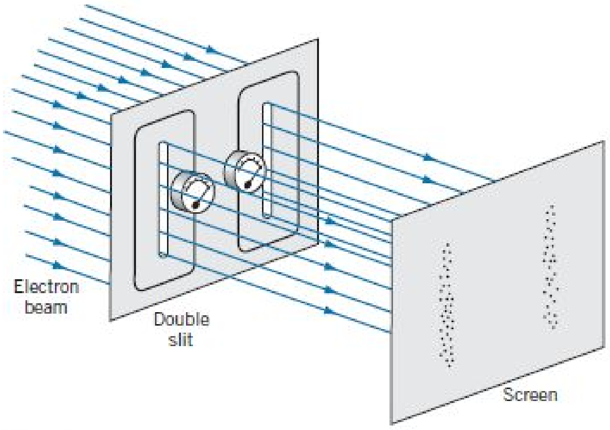


FIGURE 4.17 Apparatus to record passage of electrons through slits. Each slit is surrounded by a loop with a meter that signals the passage of an electron through the slit. No interference fringes are seen on the screen.

This is the basis for the *principle of complementarity,* which asserts that the complete description of a photon or a particle such as an electron cannot be made in terms of only particle properties or only wave properties, but that both aspects of its behavior must be considered. Moreover, the particle and wave natures cannot be observed simultaneously, and the type of behavior that we observe depends on the kind of experiment we are doing: a particle-type experiment shows only particle like behavior, and a wave-type experiment shows only wavelike behavior.

4.3 UNCERTAINTY RELATIONSHIPS FOR CLASSICAL WAVES

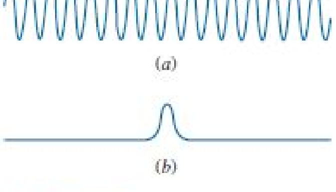


FIGURE 4.18 (*a*) A pure sine wave, which extends from $-\infty$ to $+\infty$. (*b*) A narrow wave pulse.

In quantum mechanics, we want to use de Broglie waves to describe particles. In particular, the amplitude of the wave will tell us something about the location of the particle. Clearly a pure sinusoidal wave, as in Figure 4.18*a*, is not much use in locating a particle—the wave extends from $-\infty$ to $+\infty$, so the particle might be found anywhere in that region. On the other hand, a narrow wave pulse like Figure 4.18*b* does a pretty good job of locating the particle in a small region of space, but this wave does not have an easily identifiable wavelength. In the first case, we know the wavelength exactly but have no knowledge of the location of the particle, while in the second case we have a good idea of the location of the particle but a poor knowledge of its wavelength. Because wavelength is associated with momentum by the de Broglie relationship (Eq. 4.1), a poor knowledge of the wavelength is associated with a poor knowledge of the particle's momentum. For a classical particle, we would like to know both its location and its momentum as precisely as possible. For a quantum particle, we are going to have to make some compromises—the better we know its momentum (or wavelength), the less we know about its location. We can improve our knowledge of its location only at the expense of our knowledge of its momentum.

This competition between knowledge of location and knowledge of wavelength is not restricted to de Broglie waves—classical waves show the same effect. All real waves can be represented as *wave packets*—disturbances that are localized to a finite region of space. We will discuss more about constructing wave packets in Section 4.5. In this section we will examine this competition between specifying the location and the wavelength of classical waves more closely.

Figure 4.19*a* shows a very small wave packet. The disturbance is well localized to a small region of space of length Δx. (Imagine listening to a very short burst of sound, of such brief duration that it is hard for you to recognize the pitch or frequency of the wave.) Let's try to measure the wavelength of this wave packet. Placing a measuring stick along the wave, we have some difficulty defining exactly where the wave starts and where it ends. Our measurement of the wavelength is therefore subject to a small *uncertainty* $\Delta\lambda$. Let's represent this uncertainty as a fraction ε of the wavelength λ, so that $\Delta\lambda \sim \varepsilon\lambda$. The fraction ε is certainly less than 1, but it is probably greater than 0.01, so we estimate that $\varepsilon \sim 0.1$ to within an order of magnitude. (In our discussion of uncertainty, we use the $\sim$ symbol to indicate a rough order-of-magnitude estimate.) That is, the uncertainty in our measurement of the wavelength might be roughly 10% of the wavelength.

The size of this wave disturbance is roughly one wavelength, so $\Delta x \approx \lambda$. For this discussion we want to examine the product of the size of the wave packet and the uncertainty in the wavelength, Δx times $\Delta\lambda$ with $\Delta x \approx \lambda$ and $\Delta\lambda \sim \varepsilon\lambda$:

$$\Delta x \Delta\lambda \sim \varepsilon\lambda^2 \tag{4.4}$$

This expression shows the inverse relationship between the size of the wave packet and the uncertainty in the wavelength: for a given wavelength, the smaller the size of the wave packet, the greater the uncertainty in our knowledge of the wavelength. That is, as Δx gets smaller, $\Delta\lambda$ must become larger.

Making a larger wave packet doesn't help us at all. Figure 4.19*b* shows a larger wave packet with the same wavelength. Suppose this larger wave packet contains

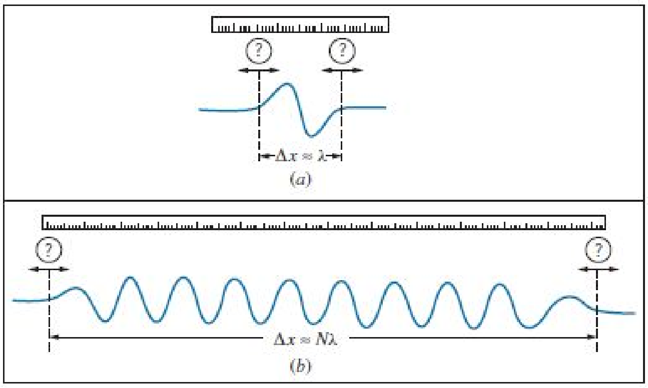


FIGURE 4.19 (*a*) Measuring the wavelength of a wave represented by a small wave packet of length roughly one wavelength. (*b*) Measuring the wavelength of a wave represented by a large wave packet consisting of N waves.

uncertainty in the frequency. Instead, they should be directly related—the better we know the period, the better we know the frequency. Here is how we obtain the relationship: Beginning with $f = 1/T$, we take differentials on both sides:

$$df = -\frac{1}{T^2}\,dT$$

Next we convert the infinitesimal differentials to finite intervals, and because we are interested only in the magnitude of the uncertainties we can ignore the minus sign:

$$\Delta f = \frac{1}{T^2}\Delta T \tag{4.6}$$

Combining Eqs. 4.5 and 4.6, we obtain

$$\Delta f\,\Delta t \sim \varepsilon \tag{4.7}$$

Equation 4.7 shows that the longer the duration of the wave packet, the more precisely we can measure its frequency.

Example 4.4

An electronics salesman offers to sell you a frequency-measuring device. When hooked up to a sinusoidal signal, it automatically displays the frequency of the signal, and to account for frequency variations, the frequency is remeasured once each second and the display is updated. The salesman claims the device to be accurate to 0.01 Hz. Is this claim valid?

Solution

Based on Eq. 4.7, and again estimating ε to be about 0.1, we know that a measurement of frequency in a time $\Delta t = 1\,\text{s}$ must have an associated uncertainty of about

$$\begin{aligned}\Delta f \sim \frac{\varepsilon}{\Delta t} &= \frac{0.1}{1\,\text{s}}\\ &= 0.1\ \text{Hz}\end{aligned}$$

It appears that the salesman may be exaggerating the precision of this device.

4.4 HEISENBERG UNCERTAINTY RELATIONSHIPS

The uncertainty relationships discussed in the previous section apply to *all* waves, and we should therefore apply them to de Broglie waves. We can use the basic de Broglie relationship $p = h/\lambda$ to relate the uncertainty in the momentum Δp to the uncertainty in wavelength $\Delta\lambda$, using the same procedure that we used to obtain Eq. 4.6. Starting with $p = h/\lambda$, we take differentials on both sides and obtain $dp = (-h/\lambda^2)d\lambda$. Now we change the differentials into differences, ignoring the minus sign:

$$\Delta p = \frac{h}{\lambda^2}\Delta\lambda \tag{4.8}$$

Werner Heisenberg (1901–1976, Germany). Best known for the uncertainty principle, he also developed a complete formulation of the quantum theory based on matrices.

An uncertainty in the momentum of the particle is directly related to the uncertainty in the wavelength associated with the particle's de Broglie wave packet.

Combining Eq. 4.8 with Eq. 4.4, we obtain

$$\Delta x \Delta p \sim \varepsilon h \tag{4.9}$$

Just like Eq. 4.4, this equation suggests an inverse relationship between Δx and Δp. The smaller the size of the wave packet of the particle, the larger is the uncertainty in its momentum (and thus in its velocity).

Quantum mechanics provides a formal procedure for calculating Δx and Δp for wave packets corresponding to different physical situations and for different schemes for confining a particle. One outcome of these calculations gives the wave packet with the smallest possible value of the product $\Delta x \Delta p$, which turns out to be $h/4\pi$, as we will discuss in the next chapter. Thus $\varepsilon = 1/4\pi$ in this case. All other wave packets will have larger values for $\Delta x \Delta p$.

The combination $h/2\pi$ occurs frequently in quantum mechanics and is given the special symbol $\hbar$ ("h-bar")

$$\hbar = \frac{h}{2\pi} = 1.05 \times 10^{-34}\ \mathrm{J \cdot s} = 6.58 \times 10^{-16}\ \mathrm{eV \cdot s}$$

In terms of $\hbar$, we can write the uncertainty relationship as

$$\Delta x \Delta p_x \geqslant \tfrac{1}{2}\hbar \tag{4.10}$$

The x subscript has been added to the momentum to remind us that Eq. 4.10 applies to motion in a given direction and relates the uncertainties in position and momentum in that direction only. Similar and independent relationships can be applied in the other directions as necessary; thus $\Delta y \Delta p_y \geqslant \hbar/2$ or $\Delta z \Delta p_z \geqslant \hbar/2$.

Equation 4.10 is the first of the *Heisenberg uncertainty relationships*. It sets the limit of the best we can possibly do in an experiment to measure *simultaneously* the location and the momentum of a particle. Another way of interpreting this equation is to say that the more we try to confine a particle, the less we know about its momentum.

Because the limit of $\hbar/2$ represents the minimum value of the product $\Delta x \Delta p_x$, in most cases we will do worse than this limit. It is therefore quite acceptable to take

$$\Delta x \Delta p_x \sim \hbar \tag{4.11}$$

as a rough estimate of the relationship between the uncertainties in location and momentum.

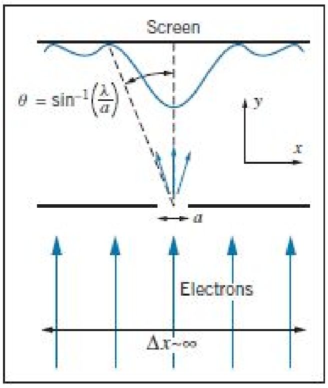


FIGURE 4.21 Single-slit diffraction of electrons. A wide beam of electrons is incident on a narrow slit. The electrons that pass through the slit acquire a component of momentum in the x direction.

As an example, let's consider a beam of electrons incident on a single slit, as in Figure 4.21. We know this experiment as single-slit diffraction, which produces the characteristic diffraction pattern illustrated in Figure 4.1. We'll assume that the particles are initially moving in the y direction and that we know their momentum in that direction as precisely as possible. If the electrons initially have no component of their momentum in the x direction, we know p_x exactly (it is exactly zero), so that $\Delta p_x = 0$; thus we know nothing about the x coordinates of the electrons ($\Delta x = \infty$). This situation represents a very wide beam of electrons, only a small fraction of which pass through the slit.

At the instant that some of the electrons pass through the slit, we know quite a bit more about their x location. In order to pass through the slit, the uncertainty in their x location is no larger than a, the width of the slit; thus $\Delta x = a$. This improvement in our knowledge of the electron's location comes at the expense of our knowledge of its momentum, however. According to Eq. 4.11, the uncertainty in the x component of its momentum is now $\Delta p_x \sim \hbar/a$. Measurements beyond the slit no longer show the particle moving precisely in the y direction (for which $p_x = 0$); the momentum now has a small x component as well, with values distributed about zero but now with a range of roughly $\pm\hbar/a$. In passing through the slit, a particle acquires on the average an x component of momentum of roughly $\hbar/a$, according to the uncertainty principle.

Let us now find the angle θ that specifies where a particle with this value of p_x lands on the screen. For small angles, $\sin\theta \approx \tan\theta$ and so

$$\sin\theta \approx \tan\theta = \frac{p_x}{p_y} = \frac{\hbar/a}{p_y} = \frac{\lambda}{2\pi a}$$

using $\lambda = h/p_y$ for the de Broglie wavelength of the electrons. The first minimum of the diffraction pattern of a single slit is located at $\sin\theta = \lambda/a$, which is larger than the spread of angles into which most of the particles are diffracted. The calculation shows that the distribution of transverse momentum given by the uncertainty principle is roughly equivalent to the spreading of the beam into the central diffraction peak, and it illustrates again the close connection between wave behavior and uncertainty in particle location.

The diffraction (spreading) of a beam following passage through a slit is just the effect of the uncertainty principle on our attempt to specify the location of the particle. As we make the slit narrower, p_x increases and the beam spreads even more. In trying to obtain more precise knowledge of the location of the particle by making the slit narrower, we have lost knowledge of the direction of its travel. This trade-off between observations of position and momentum is the essence of the Heisenberg uncertainty principle.

We can also apply the second of our classical uncertainty relationships (Eq. 4.7) to de Broglie waves. If we assume the energy-frequency relationship for light, $E = hf$, can be applied to particles, then we immediately obtain $\Delta E = h\Delta f$. Combining this with Eq. 4.7, we obtain

$$\Delta E \Delta t \sim \varepsilon h \tag{4.12}$$

Once again, the minimum uncertainty wave packet gives $\varepsilon = 1/4\pi$, and so

$$\Delta E \Delta t \geqslant \tfrac{1}{2}\hbar \tag{4.13}$$

This is the second of the Heisenberg uncertainty relationships. It tells us that the more precisely we try to determine the time coordinate of a particle, the less precisely we know its energy. For example, if a particle has a very short lifetime between its creation and decay ($\Delta t \to 0$), a measurement of its rest energy (and thus its mass) will be very imprecise ($\Delta E \to \infty$). Conversely, the rest energy of a stable particle (one with an infinite lifetime, so that $\Delta t = \infty$) can in principle be measured with unlimited precision ($\Delta E = 0$).

As in the case of the first Heisenberg relationship, we can take

$$\Delta E \Delta t \sim \hbar \tag{4.14}$$

as a reasonable estimate for most wave packets.

The Heisenberg uncertainty relationships are the mathematical representations of the *Heisenberg uncertainty principle,* which states:

> *It is not possible to make a simultaneous determination of the position and the momentum of a particle with unlimited precision,*
>
> and
>
> *It is not possible to make a simultaneous determination of the energy and the time coordinate of a particle with unlimited precision.*

These relationships give an estimate of the minimum uncertainty that can result from any experiment; measurement of the position and momentum of a particle will give a spread of values of widths Δx and Δp_x. We may, for other reasons, do much worse than Eqs. 4.10 and 4.13, but *we can do no better.*

These relationships have a profound impact on our view of nature. It is quite acceptable to say that there is an uncertainty in locating the position of a water wave. It is quite another matter to make the same statement about a de Broglie wave, because *there is an implied corresponding uncertainty in the position of the particle.* Equations 4.10 and 4.13 say that *nature imposes a limit on the accuracy with which we can do experiments.* To emphasize this point, the Heisenberg relationships are sometimes called "indeterminacy" rather than "uncertainty" principles, because the idea of uncertainty may suggest an experimental limit that can be reduced by using better equipment or technique. In actuality, these coordinates are *indeterminate* to the limits provided by Eqs. 4.10 and 4.13—no matter how hard we try, it is simply not possible to measure more precisely.

Example 4.5

An electron moves in the x direction with a speed of 3.6×10^6 m/s. We can measure its speed to a precision of 1%. With what precision can we simultaneously measure its x coordinate?

Solution

The electron's momentum is

$$\begin{aligned} p_x = mv_x &= (9.11 \times 10^{-31}\,\mathrm{kg})(3.6 \times 10^6\,\mathrm{m/s}) \\ &= 3.3 \times 10^{-24}\,\mathrm{kg \cdot m/s} \end{aligned}$$

The uncertainty Δp_x is 1% of this value, or 3.3×10^{-26} kg · m/s. The uncertainty in position is then

$$\begin{aligned} \Delta x \sim \frac{\hbar}{\Delta p_x} &= \frac{1.05 \times 10^{-34}\,\mathrm{J \cdot s}}{3.3 \times 10^{-26}\,\mathrm{kg \cdot m/s}} \\ &= 3.2\,\mathrm{nm} \end{aligned}$$

which is roughly 10 atomic diameters.

Example 4.6

Repeat the calculations of the previous example in the case of a pitched baseball ($m = 0.145$ kg) moving at a speed of 95 mi/h (42.5 m/s). Again assume that its speed can be measured to a precision of 1%.

Solution

The baseball's momentum is

$$p_x = mv_x = (0.145\,\mathrm{kg})(42.5\,\mathrm{m/s}) = 6.16\,\mathrm{kg \cdot m/s}$$

The uncertainty in momentum is 6.16×10^{-2} kg · m/s, and the corresponding uncertainty in position is

$$\Delta x \sim \frac{\hbar}{\Delta p_x} = \frac{1.05 \times 10^{-34}\,\mathrm{J \cdot s}}{6.16 \times 10^{-2}\,\mathrm{kg \cdot m/s}} = 1.7 \times 10^{-33}\,\mathrm{m}$$

The Group Speed of deBroglie Waves

Suppose we have a localized particle, represented by a group of de Broglie waves. For each component wave, the energy of the particle is related to the frequency of the de Broglie wave by $E = hf = \hbar\omega$, and so $dE = \hbar d\omega$. Similarly, the momentum of the particle is related to the wavelength of the de Broglie wave by $p = h/\lambda = \hbar k$, so $dp = \hbar dk$. The group speed of the de Broglie wave then can be expressed as

$$v_{\text{group}} = \frac{d\omega}{dk} = \frac{dE/\hbar}{dp/\hbar} = \frac{dE}{dp} \tag{4.29}$$

For a classical particle having only kinetic energy $E = K = p^2/2m$, we can find dE/dp as

$$\frac{dE}{dp} = \frac{d}{dp}\left(\frac{p^2}{2m}\right) = \frac{p}{m} = v \tag{4.30}$$

which is the velocity of the particle.

Combining Eqs. 4.29 and 4.30 we obtain an important result:

$$v_{\text{group}} = v_{\text{particle}} \tag{4.31}$$

The speed of a particle is equal to the group speed of the corresponding wave packet. The wave packet and the particle move together—wherever the particle goes, its de Broglie wave packet moves along with it like a shadow. If we do a wave-type experiment on the particle, the de Broglie wave packet is always there to reveal the wave behavior of the particle. A particle can never escape its wave nature!

The Spreading of a Moving Wave Packet

Suppose we have a wave packet that represents a confined particle at $t = 0$. For example, the particle might have passed through a single-slit apparatus. Its initial uncertainty in position is Δx_0 and its initial uncertainty in momentum is Δp_{x0}. The wave packet moves in the x direction with velocity v_x, but that velocity is not precisely known—the uncertainty in its momentum gives a corresponding uncertainty in velocity: $\Delta v_{x0} = \Delta p_{x0}/m$. Because there is an uncertainty in the velocity of the wave packet, we can't be sure where it will be located at time t. That is, its location at time t is $x = v_x t$, with velocity $v_x = v_{x0} \pm \Delta v_{x0}$. Thus there are two contributions to the uncertainty in its location at time t: the initial uncertainty Δx_0 and an additional amount equal to $\Delta v_{x0} t$ that represents the spreading of the wave packet. We'll assume that these two contributions add quadratically, like experimental uncertainties, so that the total uncertainty in the location of the particle is

$$\Delta x = \sqrt{(\Delta x_0)^2 + (\Delta v_{x0} t)^2} = \sqrt{(\Delta x_0)^2 + (\Delta p_{x0} t/m)^2} \tag{4.32}$$

Taking $\Delta p_{x0} = \hbar/\Delta x_0$ according to the uncertainty principle, we have

$$\Delta x = \sqrt{(\Delta x_0)^2 + (\hbar t/m\Delta x_0)^2} \tag{4.33}$$

If we try to make the wave packet very small at $t = 0$ (Δx_0 is small), then the second term under the square root makes the wave packet expand rapidly, because Δx_0 appears in the denominator of that term. *The more successful we are at confining*

Unit - IV Basics of Quantum Mechanics

Chapter 5

THE SCHRÖDINGER EQUATION

Quantum mechanics provides a mathematical framework in which the description of a process often includes different and possibly contradictory outcomes. A favorite illustration of that situation is the case of Schrödinger's cat. The cat is confined in a chamber with a radioactive atom, the decay of which will trigger the release of poison from a vial. Because we don't know exactly when that decay will occur, until an observation of the condition of the cat is made the quantum-mechanical description of the cat must include both "cat alive" and "cat dead" components.

The future behavior of a particle in a classical (nonrelativistic, nonquantum) situation may be predicted with absolute certainty using Newton's laws. If a particle interacts with its environment through a known force $\vec{F}$ (which might be associated with a potential energy U), we can do the mathematics necessary to solve Newton's second law, $\vec{F} = d\vec{p}/dt$ (a second-order, linear differential equation), and find the particle's location $\vec{r}(t)$ and velocity $\vec{v}(t)$ at all future times t. The mathematics may be difficult, and in fact it may not be possible to solve the equations in closed form (in which case an approximate solution can be obtained with the help of a computer). Aside from any such mathematical difficulties, the physics of the problem consists of writing down the original equation $\vec{F} = d\vec{p}/dt$ and interpreting its solutions $\vec{r}(t)$ and $\vec{v}(t)$. For example, a satellite or planet moving under the influence of a $1/r^2$ gravitational force can be shown, after the equations have been solved, to follow exactly an elliptical path.

In the case of nonrelativistic quantum physics, the basic equation to be solved is a second-order differential equation known as the *Schrödinger equation.* Like Newton's laws, the Schrödinger equation is written for a particle interacting with its environment, although we describe the interaction in terms of the potential energy rather than the force. Unlike Newton's laws, the Schrödinger equation does not give the trajectory of the particle; instead, its solution gives the *wave function* of the particle, which carries information about the particle's wavelike behavior. In this chapter we introduce the Schrödinger equation, obtain some of its solutions for certain potential energies, and learn how to interpret those solutions.

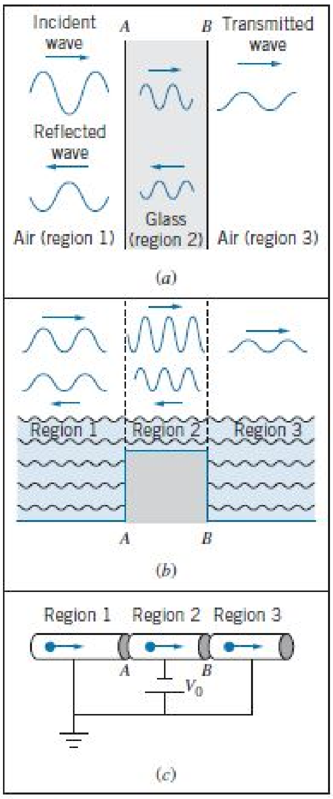


FIGURE 5.1 (*a*) A light wave in air is incident on a slab of glass, showing transmitted and reflected waves at the two boundaries (*A* and *B*). (*b*) A surface wave in water incident on a region of smaller depth similarly has transmitted and reflected waves. (*c*) The de Broglie waves of electrons moving from a region of constant zero potential to a region of constant negative potential V_0 also have transmitted and reflected components.

5.1 BEHAVIOR OF A WAVE AT A BOUNDARY

In studying wave motion, we often must analyze what occurs when a wave moves from one region or medium to a different region or medium in which the properties of the wave may change. For example, when a light wave moves from air into glass, its wavelength and the amplitude of its electric field both decrease. At every such boundary, a portion of the incident wave intensity is transmitted into the second medium and a portion is reflected back into the first medium.

Let's consider the case of a light wave incident on a glass plate, as in Figure 5.1*a*. At boundary *A*, the light wave moves from air (region 1) into glass (region 2), while at *B* the light wave moves from glass into air (region 3). The wavelength in air in region 3 is the same as the original wavelength of the incident wave in region 1, but the amplitude in region 3 is less than the amplitude in region 1, because some of the intensity is reflected at *A* and at *B*.

Other types of waves show similar behavior. For example, Figure 5.1*b* shows a surface water wave that moves into a region of shallower depth. In that region, its wavelength is smaller (but its amplitude is larger) compared with the original incident wave. When the wave enters region 3, in which the depth is the same as in region 1, the wavelength returns to its original value, but the amplitude of the wave is smaller in region 3 than in region 1 because some of the intensity was reflected at the two boundaries.

The same type of behavior occurs for de Broglie waves that characterize particles. Consider, for example, the apparatus shown in Figure 5.1*c*. Electrons are incident from the left and move inside a narrow metal tube that is at ground potential ($V = 0$). Another narrow tube in region 2 is connected to the negative terminal of a battery, which maintains it at a uniform potential of $-V_0$. Region 3 is connected to region 1 at ground potential. The gaps between the tubes can in principle be made so small that we can regard the changes in potential at A and B as occurring suddenly. In region 1, the electrons have kinetic energy K, momentum $p = \sqrt{2mK}$, and de Broglie wavelength $\lambda = h/p$. In region 2, the potential energy for the electrons is $U = qV = (-e)(-V_0) = +eV_0$. We assume that the original kinetic energy of the electrons in region 1 is greater than eV_0, so that the electrons move into region 2 with a smaller kinetic energy (equal to $K - eV_0$), a smaller momentum, and thus a greater wavelength. When the electrons move from region 2 into region 3, they gain back the lost kinetic energy and move with their original kinetic energy K and thus with their original wavelength. As in the case of the light wave or the water wave, the amplitude of the de Broglie wave in region 3 is smaller than in region 1, meaning that the current of electrons in region 3 is smaller than the incident current, because some of the electrons are reflected at the boundaries at A and B.

We can thus identify a total of 5 waves moving in the three regions: (1) a wave moving to the right in region 1 (the incident wave); (2) a wave moving to the left in region 1 (representing the net combination of waves reflected from boundary A plus waves reflected from boundary B and then transmitted through boundary A back into region 1); (3) a wave moving to the right in region 2 (representing waves transmitted through boundary A plus waves reflected at B and then reflected again at A); (4) a wave moving to the left in region 2 (waves reflected at B); and (5) a wave moving to the right in region 3 (the transmitted waves at boundary B). Because we are assuming that waves are incident from region 1, it is not possible to have a wave moving to the left in region 3.

Penetration of the Reflected Wave

Another property of classical waves that carries over into quantum waves is penetration of a totally reflected wave into a forbidden region. When a light wave is completely reflected from a boundary, an exponentially decreasing wave called the *evanescent wave* penetrates into the second medium. Because 100% of the light wave intensity is reflected, the evanescent wave carries no energy and so cannot be directly observed in the second medium. But if we make the second medium very thin (perhaps equal to a few wavelengths of light) the light wave can emerge on the opposite side of the second medium. We'll discuss this phenomenon in more detail at the end of this chapter.

The same effect occurs with de Broglie waves. Suppose we increase the battery voltage in Figure 5.1*c* so that the potential energy in region 2 (equal to eV_0) is greater than the initial kinetic energy in region 1. The electrons do not have enough energy to enter region 2 (they would have negative kinetic there) and so all electrons are reflected back into region 1.

Like light waves, de Broglie waves can also penetrate into the forbidden region with exponentially decreasing amplitudes. However, because de Broglie waves are associated with the motion of electrons, that means that electrons must also penetrate a short distance into the forbidden region. The electrons cannot be directly observed in that region, because they have negative kinetic energy there. Nor can we do any experiment that would reveal their "real" existence in the forbidden region, such as measuring the speed of their passage through that region or detecting the magnetic field that their motion might produce.

One explanation for the penetration of the electrons into the forbidden region relies on the uncertainty principle—because we can't know exactly the energy of the incident electrons, we can't say with certainty that they don't have enough kinetic energy to penetrate into the forbidden region. For short enough time Δt, the energy uncertainty $\Delta E \sim \hbar/\Delta t$ might allow the electron to travel a short distance into the forbidden region, but this extra energy does not "belong to" the electron in any permanent sense. Later in this chapter we'll discuss a more mathematical approach to this explanation of penetration into the forbidden region.

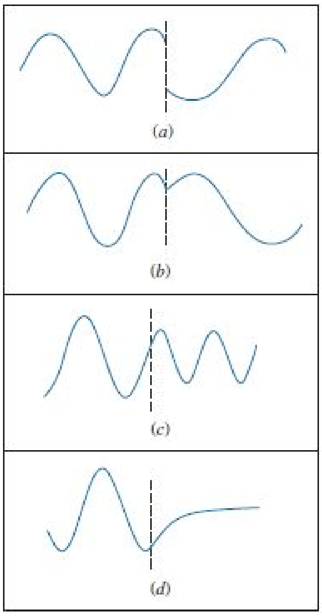


FIGURE 5.2 (*a*) A discontinuous wave. (*b*) A continuous wave with a discontinuous slope. (*c*) Two sine waves join smoothly. (*d*) A sine wave and an exponential join smoothly.

Continuity at the Boundaries

When a wave such as a light wave or a water wave crosses a boundary as in Figure 5.1, the mathematical function that describes the wave must have two properties at each boundary:

1. The wave function must be continuous.
2. The slope of the wave function must be continuous, except when the boundary height is infinite.

Figure 5.2*a* shows a discontinuous wave function; the wave displacement changes suddenly at a single location. This type of behavior is not allowed. Figure 5.2*b* shows a continuous wave function (there are no gaps) with a discontinuous slope. This type of behavior is also not allowed, unless the boundary is of infinite height. Figures 5.2*c*, *d* show how two sine curves and an exponential and a sine can be joined so that both the function and the slope are continuous.

Across any non-infinite boundary, the wave must be smooth—no gaps in the function and no sharp changes in slope. When we solve for the mathematical form of a wave function, there are usually undetermined parameters, such as the amplitude and phase of the wave. In order to make the wave smooth at the boundary, we obtain the values of those coefficients by applying the two *boundary conditions* to make the function and its slope continuous. For example, at boundary A in Figure 5.1, we first evaluate the total wave function in region 1 at A and set it equal to the wave function in region 2 at A. This guarantees that the total wave function is continuous at A. We then take the derivative of the wave function in region 1, evaluate it at A, and set that equal to the derivative of the wave function in region 2 evaluated at A. This step makes the slope in region 1 match the slope in region 2 at boundary A. These two steps give us two equations relating the parameters of the waves and allow us to find relationships between the amplitudes and phases of the waves in regions 1 and 2. The process must be repeated at every boundary, such as at B in Figure 5.1 to match the waves in regions 2 and 3.

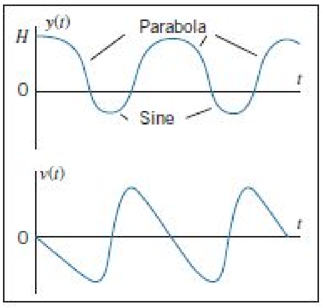


FIGURE 5.3 The position and velocity of a ball dropped from a height H above a springlike rubber sheet at $y = 0$.

We can understand the exception to the continuity of the slope for infinite boundaries with an example from classical physics. Imagine a ball dropped from a height $y = H$ above a stretched rubber sheet at $y = 0$. The ball falls freely under gravity until it strikes the sheet, which we assume behaves like an elastic spring. The sheet stretches as the ball is brought to rest, after which the restoring force propels the ball upward. The motion of the ball might be represented by Figure 5.3. Above the sheet ($y > 0$) the motion is represented by parabolas, and while the ball is in contact with the sheet ($y < 0$) the motion is described by sine curves. Note how the curves join smoothly at $y = 0$, and note how both $y(t)$ and its derivative $v(t)$ are continuous.

On the other hand, imagine a ball hitting a steel surface, which we assume to be perfectly rigid. The ball rebounds elastically, and at the instant it is in contact with the surface its velocity reverses direction. The motion of the ball is represented in Figure 5.4. At the points of contact with the surface, there is a sudden change in the velocity, corresponding to an infinite acceleration and thus to an infinite force. The function $y(t)$ is continuous, but its slope is not—the function has no gaps, but it does have sharp "points" where the slope changes suddenly.

The assumption of the perfectly rigid surface is an idealization that we make to help us understand the situation and also to help simplify the mathematics. In reality the steel surface will flex slightly and ultimately behave somewhat like a much stiffer version of the rubber sheet. In quantum mechanics we will also sometimes use an assumption of a perfectly rigid or impenetrable boundary to help us understand and simplify the analysis of a more complicated physical situation.

In this section we have established several properties of classical waves that also apply to quantum waves:

1. When a wave crosses a boundary between two regions, part of the wave intensity is reflected and part is transmitted.
2. When a wave encounters a boundary to a region from which it is forbidden, the wave will penetrate perhaps by a few wavelengths before reflecting.
3. At a finite boundary, the wave and its slope are continuous. At an infinite boundary, the wave is continuous but its slope is discontinuous.

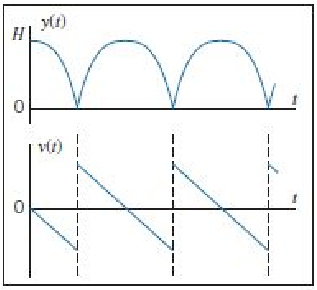


FIGURE 5.4 The position and velocity of a ball dropped from a height H above a rigid surface.

Example 5.1

In the geometry of Figure 5.1, the wave in region 1 is given by $y_1(x) = C_1 \sin(2\pi x/\lambda_1 - \phi_1)$, where $C_1 = 11.5$, $\lambda_1 = 4.97\,\text{cm}$, and $\phi_1 = -65.3^\circ$. In region 2, the wavelength is $\lambda_2 = 10.5\,\text{cm}$. The boundary A is located at $x = 0$, and the boundary B is located at $x = L$, where $L = 20.0\,\text{cm}$. Find the wave functions in regions 2 and 3.

Solution

The general form of the wave in region 2 can be represented in a form similar to that of the wave in region 1: $y_2(x) = C_2 \sin(2\pi x/\lambda_2 - \phi_2)$. To find the complete wave function in region 2, we must find the amplitude C_2 and the phase ϕ_2 by applying the boundary conditions on the function and its slope at boundary A $(x = 0)$. Setting $y_1(x = 0) = y_2(x = 0)$ gives

$$-C_1 \sin\phi_1 = -C_2 \sin\phi_2$$

The slopes can be found from the derivative of the general form $dy/dx = (2\pi/\lambda)C\cos(2\pi x/\lambda - \phi)$ evaluated at $x = 0$:

$$\frac{2\pi}{\lambda_1} C_1 \cos\phi_1 = \frac{2\pi}{\lambda_2} C_2 \cos\phi_2$$

Dividing the first equation by the second eliminates C_2 and allows us to solve for ϕ_2:

$$\begin{aligned}\phi_2 &= \tan^{-1}\left(\frac{\lambda_1}{\lambda_2}\tan\phi_1\right)\\ &= \tan^{-1}\left(\frac{4.97\,\text{cm}}{10.5\,\text{cm}}\tan(-65.3^\circ)\right)\\ &= -45.8^\circ\end{aligned}$$

We can solve for C_2 using the result from applying the first boundary condition:

$$C_2 = C_1\frac{\sin\phi_1}{\sin\phi_2} = 11.5\frac{\sin(-65.3^\circ)}{\sin(-45.8^\circ)} = 14.6$$

To find the wave function in region 3, which we assume to have the same form $y_3(x) = C_3 \sin(2\pi x/\lambda_1 - \phi_3)$, we must apply the boundary conditions on y_2 and y_3 at $x = L$. Applying the two boundary conditions in the same way we did at $x = 0$, we obtain

$$\begin{aligned}C_2 \sin\left(\frac{2\pi L}{\lambda_2} - \phi_2\right) &= C_3 \sin\left(\frac{2\pi L}{\lambda_1} - \phi_3\right)\\ \frac{2\pi}{\lambda_2} C_2 \cos\left(\frac{2\pi L}{\lambda_2} - \phi_2\right) &= \frac{2\pi}{\lambda_1} C_3 \cos\left(\frac{2\pi L}{\lambda_1} - \phi_3\right)\end{aligned}$$

Proceeding as we did before, we divide these two equations to find $\phi_3 = 60.9^\circ$, and then from either equation obtain $C_3 = 7.36$. Our two solutions are then $y_2(x) = 14.6 \sin(2\pi x/10.5 + 45.8^\circ)$ and $y_3(x) = 7.36 \sin(2\pi x/4.97 + 14.6^\circ)$, with x measured in cm. Figure 5.5 shows the wave in all three regions. Note how the waves join smoothly at the boundaries.

How is it possible that the amplitude of y_2 can be greater than the amplitude of y_1? Keep in mind that y_1 represents the total wave in region 1, which includes the incident wave and the reflected wave. Depending on the phase difference between them, when the incident and reflected waves are added to obtain y_1, the amplitude of the resultant can be smaller than the amplitude of either wave.

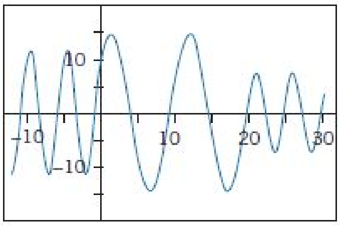


FIGURE 5.5 Example 5.1.

5.2 CONFINING A PARTICLE

A free particle (that is, a particle on which no forces act anywhere) is by definition not confined, so it can be located anywhere. It has, as we discussed in Chapter 4, a definite wavelength, momentum, and energy (for which we can choose any value).

A confined particle, on the other hand, is represented by a wave packet that makes it likely to be found only in a region of space of size Δx. We construct such a wave packet by adding together different sine or cosine waves to obtain the desired mathematical shape.

In quantum mechanics, we often want to analyze the behavior of confined particles, for example an electron that is attached to a specific atom or molecule. We'll consider the properties of atomic electrons beginning in Chapter 6, but for now let's look at a simpler problem: an electron moving in one dimension and confined by a series of electric fields. Figure 5.6 shows how the apparatus of Figure 5.1*c* might be modified for this purpose. The center section is grounded (so that $V = 0$) and the two side sections are connected to batteries so that they are at potentials of $-V_0$ relative to the center section. As before, we assume that the gaps between the center section and the side sections can be made as narrow as possible, so we can regard the potential energy as changing instantaneously at the boundaries A and B. This arrangement is often called a *potential energy well*.

The potential energy of an electron in this situation is then 0 in the center section and $U_0 = qV = (-e)(-V_0) = +eV_0$ in the two side sections as shown in Figure 5.6. To confine the electron, we want to consider cases in which it moves in the center section with a kinetic energy K that is less than U_0. For example, the electron might have a kinetic energy of 5 eV in the center section, and the side sections might have potential energies of 10 eV. The electron thus does not have enough energy to "climb" the potential energy hill between the center section and the side sections, and (at least from the classical point of view) the electron is confined to the center section.

We'll discuss the full solution to this problem later in this chapter, but for now let's simplify even further and consider the case of an infinitely high potential energy barrier at A and B. This is a good approximation to the situation in which the kinetic energy of the electron in the center section is much smaller than the potential energy supplied by the batteries. In this case the penetration

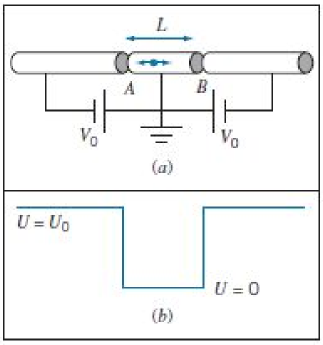


FIGURE 5.6 (*a*) Apparatus for confining an electron to the center region of length L. (*b*) The potential energy of an electron in this apparatus.

into the forbidden region, which we discussed in Section 5.1, cannot occur. The probability to find the electron in either of the side regions is therefore precisely zero everywhere in those regions, and thus the wave amplitude is zero everywhere in those regions, including at the boundaries (locations A and B). For the wave function to be continuous, the wave function in the center section must have values of zero at A and B.

Of all the possible waves that might be used to describe the particle in this center section, the continuity condition restricts us to waves that have zero amplitude at the boundaries. Some of those waves are illustrated in Figure 5.7. Note that the wave function is continuous, but its slope is not (there are sharp points in the function at locations A and B). This is an example of the exception to the second boundary condition—the slope may be discontinuous at an infinite barrier.

In contrast to the free particle for which the wavelength could have any value, *only certain values of the wavelength are allowed*. The de Broglie relationship then tells us that only certain values of the momentum are allowed, and consequently only certain values of the energy are allowed. The energy is not a continuous variable, free to take on any arbitrary value; instead, the energy is a discrete variable that is restricted to a certain set of values. This is known as *quantization of energy*.

You can see directly from Figure 5.7 that the allowed wavelengths are $2L, L, 2L/3, \ldots$, where L is the length of the center section. We can write these wavelengths as

$$\lambda_n = \frac{2L}{n} \qquad n = 1, 2, 3, \ldots \tag{5.1}$$

This set of wavelengths is identical to the wavelengths of the classical problem of standing waves on a string stretched between two points. From the de Broglie relationship $\lambda = h/p$ we obtain

$$p_n = n\frac{h}{2L} \tag{5.2}$$

The energy of the particle in the center section is only kinetic energy $p^2/2m$, and so

$$E_n = n^2\frac{h^2}{8mL^2} \tag{5.3}$$

These are the allowed or quantized values of the energy of the electron.

A wave packet describing the electron in this region must be a combination of waves with the allowed values of the wavelengths. However, it is not necessary to construct a wave packet from a combination of waves to describe this confined particle. Even a single one of these waves represents the confined particle, because the wave function must be zero in the forbidden regions. So the waveforms shown in Figure 5.7 can represent wave packets of this confined electron, each wave packet consisting of only a single wave.

The appearance of energy quantization accompanies every attempt to confine a particle to a finite region of space. Quantization of energy is one of the principal features of the quantum theory, and studying the quantized energy levels of systems (such as by observing the energies of emitted photons) is an important technique of experimental physics that gives us information about the properties of atoms and nuclei.

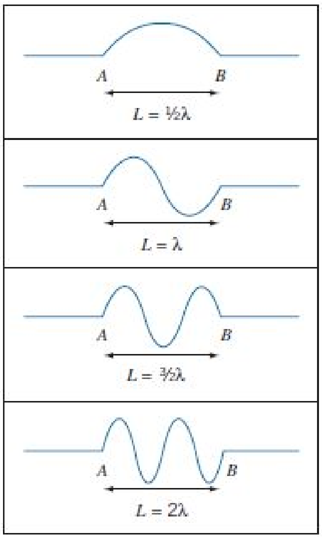


FIGURE 5.7 Some possible waves that might be used to describe an electron confined by an infinite potential energy barrier to a region of length L.

Applying the Uncertainty Principle to a Confined Particle

In Chapter 4 we constructed wave packets and showed how the uncertainty principle related the size of the wave packet to the range of wavelengths that was used in its construction. Let's now see how the Heisenberg uncertainty relationships apply in the case of a confined particle.

In the arrangement of Figure 5.6 (with infinitely high barriers on each side), the particle is known to be somewhere in the center section of the apparatus, and thus $\Delta x \sim L$ is a reasonable estimate of the uncertainty in its location. To find the uncertainty in its momentum, we use the rigorous definition of uncertainty given in Eq. 4.15: $\Delta p_x = \sqrt{(p_x^2)_{\text{av}} - (p_{x,\text{av}})^2}$. The particle moving in the center section can be considered to be moving to the left or to the right with equal probability (just as the classical standing-wave problem can be analyzed as the superposition of identical waves moving to the left and to the right). Thus $p_{x,\text{av}} = 0$. If the particle is moving with a momentum given by Eq. 5.2, $p_x^2 = (nh/L)^2$ and so $\Delta p_x = nh/L$. Combining the uncertainties in position and momentum, we have

$$\Delta x \Delta p_x \sim L\frac{nh}{L} = nh \tag{5.4}$$

The product of the uncertainties is certainly greater than $\hbar/2$, and so the result of confining the particle is entirely consistent with the Heisenberg uncertainty relationship. Note that even the smallest possible value of the product of the uncertainties (which is obtained for $n = 1$) is still much larger than the minimum value given by the uncertainty principle.

Later in this chapter, we will use a more rigorous way to evaluate the uncertainty in position using a formula similar to Eq. 4.15 to find the uncertainty in position, and we will find that the result does not differ very much from the estimate of Eq. 5.4.


Erwin Schrödinger (1887–1961, Austria). Although he disagreed with the probabilistic interpretation that was later given to his work, he developed the mathematical theory of wave mechanics that for the first time permitted the wave behavior of physical systems to be calculated.

5.3 THE SCHRÖDINGER EQUATION

The differential equation whose solution gives us the wave behavior of particles is called the *Schrödinger equation*. It was developed in 1926 by Austrian physicist Erwin Schrödinger. The equation cannot be derived from any previous laws or postulates; like Newton's equations of motion or Maxwell's equations of electromagnetism, it is a new and independent result whose correctness can be determined only by comparing its predictions with experimental results. For nonrelativistic motion, the Schrödinger equation gives results that correctly account for observations at the atomic and subatomic level.

We can justify the form of the Schrödinger equation by examining the solution expected for the free particle, which should give a wave whose shape at any particular time, specified by the *wave function* $\psi(x)$, is that of a simple de Broglie wave, such as $\psi(x) = A \sin kx$, where A is the amplitude of the wave and $k = 2\pi/\lambda$. If we are looking for a differential equation, then we need to take some derivatives:

$$\frac{d\psi}{dx} = kA\cos kx, \qquad \frac{d^2\psi}{dx^2} = -k^2 A \sin kx = -k^2\psi(x)$$

Note that the second derivative gives the original function again. With the kinetic energy $K = p^2/2m = (h/\lambda)^2/2m = \hbar^2 k^2/2m$, we can then write

$$\frac{d^2\psi}{dx^2} = -k^2\psi(x) = -\frac{2m}{\hbar^2}K\psi(x) = -\frac{2m}{\hbar^2}(E-U)\psi(x)$$

where $E = K + U$ is the nonrelativistic total energy of the particle. For a free particle, $U = 0$ so $E = K$; however, we are using the free particle solution to try to extend to the more general case in which there is a potential energy $U(x)$. The equation then becomes

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + U(x)\psi(x) = E\psi(x) \tag{5.5}$$

Equation 5.5 is the *time-independent Schrödinger equation* for one-dimensional motion.

The solution to Eq. 5.5 gives the shape of the wave at time $t = 0$. The mathematical function that describes a one-dimensional *traveling* wave must involve both x and t. This wave is represented by the function $\Psi(x,t)$:

$$\Psi(x,t) = \psi(x)e^{-i\omega t} \tag{5.6}$$

The time dependence is given by the complex exponential function $e^{-i\omega t}$ with $\omega = E/\hbar$. (You can find a few useful formulas involving complex numbers in Appendix B.) We'll discuss the time-dependent part later in this chapter. For now, we'll concentrate on the time-independent function $\psi(x)$.

We assume that we know the potential energy $U(x)$, and we wish to obtain the wave function $\psi(x)$ and the energy E *for that potential energy*. This is a general example of a type of problem known as an *eigenvalue* problem; we find that it is possible to obtain solutions to the equation only for particular values of E, which are known as the *energy eigenvalues*.

The general procedure for solving the Schrödinger equation is as follows:

1. Begin by writing Eq. 5.5 with the appropriate $U(x)$. Note that if the potential energy changes discontinuously [$U(x)$ may be represented by a discontinuous function; $\psi(x)$ may *not*], we may need to write different equations for different regions of space. Examples of this sort are given in Section 5.4.
2. Using general mathematical techniques suited to the form of the equation, find a mathematical function $\psi(x)$ that is a solution to the differential equation. Because there is no one specific technique for solving differential equations, we will study several examples to learn how to find solutions.
3. In general, several solutions may be found. By applying boundary conditions some of these may be eliminated and some arbitrary constants may be determined. It is generally the application of the boundary conditions that selects out the allowed energies.
4. If you are seeking solutions for a potential energy that changes discontinuously, you must apply the continuity conditions on $\psi(x)$ (and usually on $d\psi/dx$) at the boundary between different regions.

Because the Schrödinger equation is linear, any constant multiplying a solution is also a solution. The method to determine the amplitude of the wave function is discussed in the next section.

Unit - V Exactly Soluvable Quantum Systems

The Free Particle

For a free particle, the force is zero and so the potential energy is constant. We may choose any value for that constant, so for convenience we'll choose $U_0 = 0$. The solution is given by Eq. 5.16, $\psi(x) = A\sin kx + B\cos kx$. The energy of the particle is

$$E = \frac{\hbar^2 k^2}{2m} \tag{5.19}$$

Our solution has placed no restrictions on k, so the energy is permitted to have any value (in the language of quantum physics, we say that the energy is *not* quantized). We note that Eq. 5.19 is the kinetic energy of a particle with momentum $p = \hbar k$ or, equivalently, $p = h/\lambda$. This is as we would have expected, because the free particle can be represented by a de Broglie wave with any wavelength.

Solving for A and B presents some difficulties because the normalization integral, Eq. 5.9, cannot be evaluated from $-\infty$ to $+\infty$ for this wave function. We therefore cannot determine probabilities for the free particle from the wave function of Eq. 5.16.

It is also instructive to write the wave function in terms of complex exponentials, using $\sin kx = (e^{ikx} - e^{-ikx})/2i$ and $\cos kx = (e^{ikx} + e^{-ikx})/2$:

$$\psi(x) = A\left(\frac{e^{ikx} - e^{-ikx}}{2i}\right) + B\left(\frac{e^{ikx} + e^{-ikx}}{2}\right) = A'e^{ikx} + B'e^{-ikx} \tag{5.20}$$

where $A' = A/2i + B/2$ and $B' = -A/2i + B/2$. To interpret this solution in terms of waves we form the complete time-dependent wave function using Eq. 5.6:

$$\Psi(x,t) = (A'e^{ikx} + B'e^{-ikx})e^{-i\omega t} = A'e^{i(kx-\omega t)} + B'e^{-i(kx+\omega t)} \tag{5.21}$$

The dependence of the first term on $kx - \omega t$ identifies this term as representing a wave moving to the right (in the positive x direction) with amplitude A', and the second term involving $kx + \omega t$ represents a wave moving to the left (in the negative x direction) with amplitude B'.

If we want the wave to represent a beam of particles moving in the $+x$ direction, then we must set $B' = 0$. The probability density associated with this wave is then, according to Eq. 5.7,

$$P(x) = |\psi(x)|^2 = |A'|^2 e^{ikx}e^{-ikx} = |A'|^2 \tag{5.22}$$

The probability density is constant, meaning the particles are equally likely to be found anywhere along the x axis. This is consistent with our discussion of the free-particle de Broglie wave in Chapter 4—a wave of precisely defined wavelength extends from $x = -\infty$ to $x = +\infty$ and thus gives a completely unlocalized particle.

Infinite Potential Energy Well

Now we'll consider the formal solution to the problem we discussed in Section 5.2: a particle is trapped in the region between $x = 0$ and $x = L$ by infinitely high potential energy barriers. Imagine an apparatus like that of Figure 5.6, in which the particle moves freely in this region and makes elastic collisions with the perfectly rigid barriers that confine it. This problem is sometimes called "a particle in a box." For now we'll assume that the particle moves in only one dimension; later we'll expand to two and three dimensions.

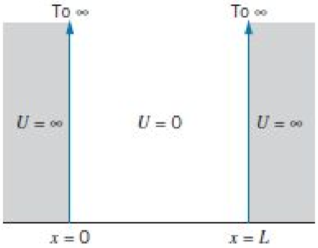


FIGURE 5.9 The potential energy of a particle that moves freely ($U = 0$) in the region $0 \leq x \leq L$ but is completely excluded ($U = \infty$) from the regions $x < 0$ and $x > L$.

The potential energy may be expressed as:

$$\begin{aligned} U(x) &= 0 \qquad 0 \leq x \leq L \\ &= \infty \qquad x < 0, x > L \end{aligned} \tag{5.23}$$

The potential energy is shown in Figure 5.9. We are free to choose any constant value for U in the region $0 \leq x \leq L$; we choose it to be zero for convenience.

Because the potential energy is different in the regions inside and outside the well, we must find separate solutions in each region. We can analyze the outside region in either of two ways. If we examine Eq. 5.5 for the region outside the well, we find that the only way to keep the equation from becoming meaningless when $U \to \infty$ is to require $\psi = 0$, so that $U\psi$ will not become infinite. Alternatively, we can go back to the original statement of the problem. If the walls at the boundaries of the well are perfectly rigid, the particle must always be in the well, and the probability for finding it outside must be zero. To make the probability zero everywhere outside the well, we must make $\psi = 0$ everywhere outside. Thus we have

$$\psi(x) = 0 \qquad x < 0, x > L \tag{5.24}$$

The Schrödinger equation for $0 \leq x \leq L$, when $U(x) = 0$, is identical with Eq. 5.14 with $U_0 = 0$ and has the same solution:

$$\psi(x) = A \sin kx + B \cos kx \qquad 0 \leq x \leq L \tag{5.25}$$

with

$$k = \sqrt{\frac{2mE}{\hbar^2}} \tag{5.26}$$

Our solution is not yet complete, for we have not evaluated A or B, nor have we found the allowed values of the energy E. To do this, we must apply the requirement that $\psi(x)$ is continuous across any boundary. In this case, we require that our solutions for $x < 0$ and $x > 0$ match up at $x = 0$; similarly, the solutions for $x > L$ and $x < L$ must match at $x = L$.

Let us begin at $x = 0$. At $x < 0$, we have found that $\psi = 0$, and so we must set $\psi(x)$ of Eq. 5.25 to zero at $x = 0$.

$$\psi(0) = A \sin 0 + B \cos 0 = 0 \tag{5.27}$$

which gives $B = 0$. Because $\psi = 0$ for $x > L$, the second boundary condition is $\psi(L) = 0$, so

$$\psi(L) = A \sin kL + B \cos kL = 0 \tag{5.28}$$

We have already found $B = 0$, so we must now have $A \sin kL = 0$. Either $A = 0$, in which case $\psi = 0$ *everywhere*, $\psi^2 = 0$ everywhere, and there is no particle (a meaningless solution) or else $\sin kL = 0$, which is true only when $kL = \pi, 2\pi, 3\pi, \ldots$, or

$$kL = n\pi \qquad n = 1, 2, 3, \ldots \tag{5.29}$$

With $k = 2\pi/\lambda$, we have $\lambda = 2L/n$; this is identical with the result obtained in introductory mechanics for the wavelengths of the standing waves in a string of length L fixed at both ends, which we already obtained in Section 5.2 (Eq. 5.1). *Thus the solution to the Schrödinger equation for a particle trapped in a linear region of length L is a series of standing de Broglie waves!* Not all wavelengths are permitted; only certain values, determined from Eq. 5.29, may occur.

From Eq. 5.26 we find that, because only certain values of k are permitted by Eq. 5.29, only certain values of E may occur—*the energy is quantized!* Solving Eq. 5.29 for k and substituting into Eq. 5.26, we obtain

$$E_n = \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2 \pi^2 n^2}{2mL^2} = \frac{h^2 n^2}{8mL^2} \qquad n = 1, 2, 3, \ldots \qquad (5.30)$$

For convenience, let $E_0 = \hbar^2\pi^2/2mL^2 = h^2/8mL^2$; this unit of energy is determined by the mass of the particle and the width of the well. Then $E_n = n^2 E_0$, and the only allowed energies for the particle are $E_0, 4E_0, 9E_0, 16E_0$, etc. All intermediate values, such as $3E_0$ or $6.2E_0$, are forbidden. Figure 5.10 shows the allowed energy levels. The lowest energy state, for which $n = 1$, is known as the *ground state,* and the states with higher energies ($n > 1$) are known as *excited states.*

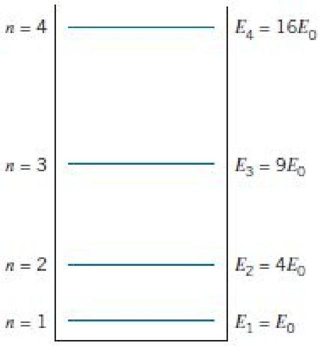


FIGURE 5.10 The first four energy levels in a one-dimensional infinite potential energy well.

Because the energy is purely kinetic in this case, our result means that only certain speeds are permitted for the particle. This is very different from the case of the classical trapped particle, in which the particle can be given any initial velocity and will move forever, back and forth, at the same speed. In the quantum case, this is not possible; only certain initial speeds can result in sustained states of motion; these special conditions represent the "stationary states." Average values calculated according to Eq. 5.13 likewise do not change with time.

From one energy state, the particle can make jumps or transitions to another energy state by absorbing or releasing an amount of energy equal to the energy difference between the two states. By absorbing energy the particle will move to a higher energy state, and by releasing energy it moves to a lower energy state. A similar effect occurs for electrons in atoms, in which the absorbed or released energy is usually in the form of a photon of visible light or other electromagnetic radiation. For example, from the state with $n = 3$ ($E_3 = 9E_0$), the particle might absorb an energy of $\Delta E = 7E_0$ and jump upward to the $n = 4$ state ($E_4 = 16E_0$) or might release energy of $\Delta E = 5E_0$ and jump downward to the $n = 2$ state ($E_2 = 4E_0$).

Example 5.2

An electron is trapped in a one-dimensional region of length 1.00×10^{-10} m (a typical atomic diameter). (*a*) Find the energies of the ground state and first two excited states. (*b*) How much energy must be supplied to excite the electron from the ground state to the second excited state? (*c*) From the second excited state, the electron drops down to the first excited state. How much energy is released in this process?

Solution

(*a*) The basic quantity of energy needed for this calculation is

$$E_0 = \frac{h^2}{8mL^2} = \frac{(hc)^2}{8mc^2L^2}$$

$$= \frac{(1240\,\mathrm{eV\cdot nm})^2}{8(511{,}000\,\mathrm{eV})(0.100\,\mathrm{nm})^2} = 37.6\,\mathrm{eV}$$

With $E_n = n^2 E_0$, we can find the energy of the states:

$$n = 1: \quad E_1 = E_0 = 37.6\,\text{eV}$$
$$n = 2: \quad E_2 = 4E_0 = 150.4\,\text{eV}$$
$$n = 3: \quad E_3 = 9E_0 = 338.4\,\text{eV}$$

(*b*) The energy difference between the ground state and the second excited state is

$$\Delta E = E_3 - E_1 = 338.4\,\text{eV} - 37.6\,\text{eV} = 300.8\,\text{eV}$$

This is the energy that must be absorbed for the electron to make this jump.

(*c*) The energy difference between the second and first excited states is

$$\Delta E = E_3 - E_2 = 338.4\,\text{eV} - 150.4\,\text{eV} = 188.0\,\text{eV}$$

This is the energy that is released when the electron makes this jump.

To complete the solution for $\psi(x)$, we must determine the constant A by using the normalization condition given in Eq. 5.9, $\int_{-\infty}^{+\infty} |\psi(x)|^2 dx = 1$. The integrand is zero in the regions $-\infty < x \leq 0$ and $L \leq x < +\infty$, so all that remains is

$$\int_0^L A^2 \sin^2 \frac{n\pi x}{L}\, dx = 1 \tag{5.31}$$

from which we find $A = \sqrt{2/L}$. The complete wave function for $0 \leq x \leq L$ is then

$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L} \qquad n = 1, 2, 3, \ldots \tag{5.32}$$

In Figure 5.11, the wave functions and probability densities ψ^2 are illustrated for the lowest several states.

In the ground state, the particle has the greatest probability to be found near the middle of the well ($x = L/2$), and the probability falls off to zero between the center and the sides of the well. This is very different from the behavior of a classical particle—a classical particle moving at constant speed would be found with equal probability at every location inside the well. The quantum particle also has constant speed but yet is still found with differing probability at various locations in the well. It is the wave nature of the quantum particle that is responsible for this very nonclassical behavior.

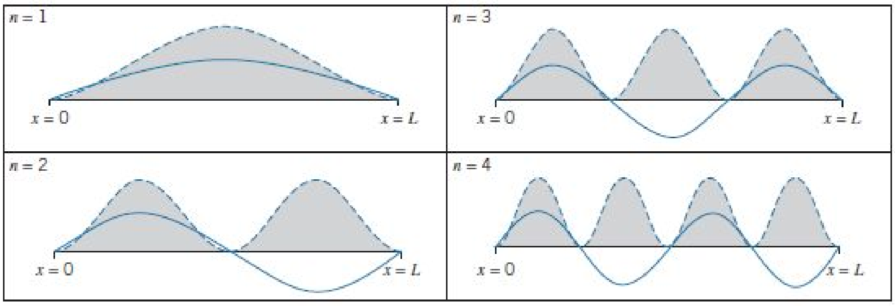


FIGURE 5.11 The wave functions (solid lines) and probability densities (shaded regions) of the first four states in the one-dimensional infinite potential energy well.

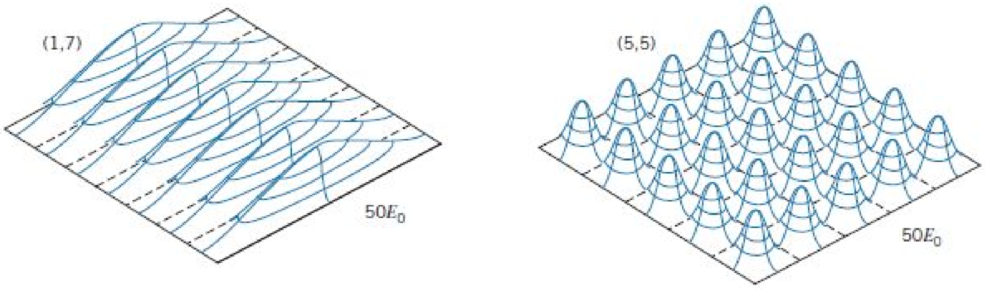


FIGURE 5.19 Two very different probability densities with exactly the same energy.

$E = 13E_0$. This degeneracy arises from interchanging n_x and n_y (which is the same as interchanging the x and y axes), so the probability distributions in the two cases are not very different. However, consider the state with $E = 50E_0$, for which there are three sets of quantum numbers: $n_x = 7, n_y = 1$; $n_x = 1, n_y = 7$; and $n_x = 5, n_y = 5$. The first two sets of quantum numbers result from the interchange of n_x and n_y and so have similar probability distributions, but the third represents a *very* different state of motion, as shown in Figure 5.19. The level at $E = 13E_0$ is said to be *two-fold* degenerate, while the level at $E = 50E_0$ is *three-fold* degenerate; we could also say that one level has a degeneracy of 2, while the other has a degeneracy of 3.

Degeneracy occurs in general whenever a system is labeled by two or more quantum numbers; as we have seen in the above calculation, different combinations of quantum numbers often can give the same value of the energy. The number of different quantum numbers required by a given physical problem turns out to be exactly equal to the number of dimensions in which the problem is being solved—one-dimensional problems need only one quantum number, two-dimensional problems need two, and so forth. When we get to three dimensions, as in Problem 19 at the end of this chapter and especially in the hydrogen atom in Chapter 7, we find that the effects of degeneracy become more significant; in the case of atomic physics, the degeneracy is a major contributor to the structure and properties of atoms.

5.5 THE SIMPLE HARMONIC OSCILLATOR

Another situation that can be analyzed using the Schrödinger equation is the one-dimensional simple harmonic oscillator. The classical oscillator is an object of mass m attached to a spring of force constant k. The spring exerts a restoring force $F = -kx$ on the object, where x is the displacement from its equilibrium position. Using Newton's laws, we can analyze the oscillator and show that it has a (circular or angular) frequency $\omega_0 = \sqrt{k/m}$ and a period $T = 2\pi\sqrt{m/k}$. The maximum distance of the oscillating object from its equilibrium position is x_0, the amplitude of the oscillation. The oscillator has its maximum kinetic energy at $x = 0$; its kinetic energy vanishes at the *turning points* $x = \pm x_0$. At the turning points the oscillator comes to rest for an instant and then reverses its direction of motion. The motion is, of course, confined to the region $-x_0 \leq x \leq +x_0$.

Why analyze the motion of such a system using quantum mechanics? Although we never find in nature an example of a one-dimensional quantum oscillator, there are systems that behave approximately as one—a vibrating diatomic molecule, for example. In fact, any system in a smoothly varying potential energy well near its minimum behaves approximately like a simple harmonic oscillator.

A force $F = -kx$ has the associated potential energy $U = \frac{1}{2}kx^2$, and so we have the Schrödinger equation:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + \frac{1}{2}kx^2\psi = E\psi \tag{5.46}$$

(Because we are working in one dimension, U and ψ are functions only of x.) There are no boundaries between different regions of potential energy here, so the wave function must fall to zero for both $x \to +\infty$ and $x \to -\infty$. The simplest function that satisfies these conditions, which turns out to be the correct ground state wave function, is $\psi(x) = Ae^{-ax^2}$. The constant a and the energy E can be found by substituting this function into Eq. 5.46. We begin by evaluating $d^2\psi/dx^2$.

$$\frac{d\psi}{dx} = -2ax(Ae^{-ax^2})$$

$$\frac{d^2\psi}{dx^2} = -2a(Ae^{-ax^2}) - 2ax(-2ax)Ae^{-ax^2} = (-2a + 4a^2x^2)Ae^{-ax^2}$$

Substituting into Eq. 5.46 and canceling the common factor Ae^{-ax^2} yields

$$\frac{\hbar^2 a}{m} - \frac{2a^2\hbar^2}{m}x^2 + \frac{1}{2}kx^2 = E \tag{5.47}$$

Equation 5.47 is *not* an equation to be solved for x, because we are looking for a solution that is valid for *any* x, not just for one specific value. In order for this to hold for *any* x, the coefficients of x^2 must cancel and the remaining constants must be equal. (That is, consider the equation $bx^2 = c$. It will be true for *any* and *all* x only if both $b = 0$ and $c = 0$.) Thus

$$-\frac{2a^2\hbar^2}{m} + \frac{1}{2}k = 0 \qquad \text{and} \qquad \frac{\hbar^2 a}{m} = E \tag{5.48}$$

which yield

$$a = \frac{\sqrt{km}}{2\hbar} \qquad \text{and} \qquad E = \frac{1}{2}\hbar\sqrt{k/m} \tag{5.49}$$

We can also write the energy in terms of the classical frequency $\omega_0 = \sqrt{k/m}$ as

$$E = \frac{1}{2}\hbar\omega_0 \tag{5.50}$$

The coefficient A is found from the normalization condition (see Problem 20 at the end of the chapter). The result, which is valid *only* for this ground-state wave function, is $A = (m\omega_0/\hbar\pi)^{1/4}$. The complete wave function of the ground state is then

$$\psi(x) = \left(\frac{m\omega_0}{\hbar\pi}\right)^{1/4} e^{-(\sqrt{km}/2\hbar)x^2} \tag{5.51}$$

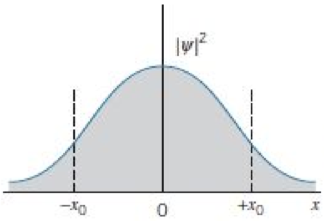


FIGURE 5.20 The probability density for the ground state of the simple harmonic oscillator. The classical turning points are at $x = \pm x_0$.

The probability density for this wave function is illustrated in Figure 5.20. Note that, as in the case of the finite potential energy well, the probability density can penetrate into the forbidden region beyond the classical turning points at $x = \pm x_0$ (in this region the potential energy is greater than E).

The solution we have found corresponds only to the *ground state* of the oscillator. The general solution is of the form $\psi_n(x) = Af_n(x)e^{-ax^2}$, where $f_n(x)$ is a polynomial in which the highest power of x is x^n. The corresponding energies are

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega_0 \qquad n = 0, 1, 2, \ldots \tag{5.52}$$

These levels are shown in Figure 5.21. Note that they are *uniformly spaced*, in contrast to the one-dimensional infinite potential energy well. Probability densities are shown in Figure 5.22. All of the solutions have the property of penetration of probability density into the forbidden region beyond the classical turning points. The probability density oscillates, somewhat like a sine wave, between the turning points, and decreases like e^{-ax^2} to zero beyond the turning points. Note the great similarity between the probability densities for the quantum oscillator and those of the finite potential energy well (Figure 5.14).

A sequence of vibrational excited states similar to Figure 5.21 is commonly found in diatomic molecules such as HCl (see Chapter 9). The spacing between the states is typically 0.1 − 1 eV; the states are observed when photons (in the infrared region of the spectrum) are emitted or absorbed as the molecule jumps from one state to another. A similar sequence is observed in nuclei, where the spacing is 0.1 − 1 MeV and the radiations are in the gamma-ray region of the spectrum.

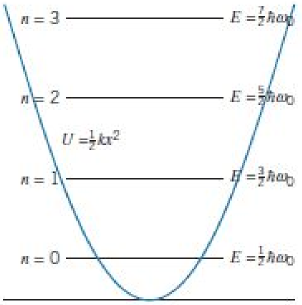


FIGURE 5.21 Energy levels of the simple harmonic oscillator. Note that the levels have equal spacings and that the distance between the classical turning points increases with energy.

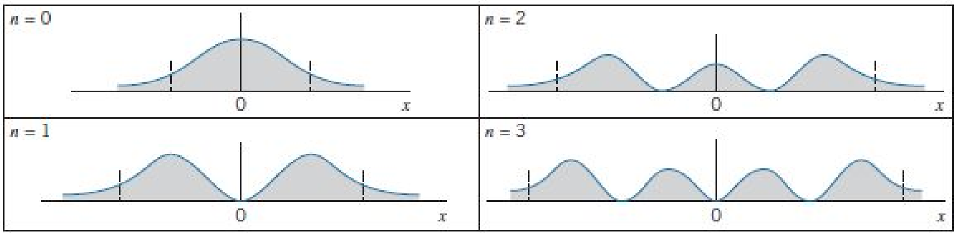


FIGURE 5.22 Probability densities for the simple harmonic oscillator. Note how the distance between the classical turning points (marked by the short vertical lines) increases with energy. Compare with the probability densities for the finite potential energy well (Figure 5.14).

Example 5.5

An electron is bound to a region of space by a springlike force with an effective spring constant of $k = 95.7\ \text{eV/nm}^2$. (*a*) What is its ground-state energy? (*b*) How much energy must be absorbed for the electron to jump from the ground state to the second excited state?

Solution

(*a*) The ground-state energy is

$$E = \frac{1}{2}\hbar\omega_0 = \frac{1}{2}\hbar\sqrt{\frac{k}{m}} = \frac{1}{2}\hbar c\sqrt{\frac{k}{mc^2}}$$

$$= \frac{1}{2}(197\ \text{eV}\cdot\text{nm})\sqrt{\frac{95.7\ \text{eV/nm}^2}{0.511 \times 10^6\ \text{eV}}}$$

$$= 1.35\ \text{eV}$$

(*b*) The difference between adjacent energy levels is $\hbar\omega_0 = 2.70\ \text{eV}$ for all energy levels, so the energy that must be absorbed to go from the ground state to the second excited state is $\Delta E = 2 \times 2.70\ \text{eV} = 5.40\ \text{eV}$.

the other forms of the uncertainty principle; for example, reducing the uncertainty in x is always accompanied by an increase in the uncertainty in p_x.

From this discussion you can see why the length of the angular momentum is defined according to Eq. 7.5 and why, for example, we could not have simply defined the length as $|\vec{\mathbf{L}}| = l\hbar$. If this were possible, then when m_l had its maximum value ($m_l = +l$), we would have $L_z = m_l\hbar = l\hbar$; the length of the vector would then be equal to its z component, and so it must lie along the z axis with $L_x = L_y = 0$. However, this simultaneous exact knowledge of all three components of $\vec{\mathbf{L}}$ violates the angular momentum form of the uncertainty principle, and therefore this situation is not permitted to occur. It is therefore necessary for the length of $\vec{\mathbf{L}}$ to be greater than $l\hbar$.

7.3 THE HYDROGEN ATOM WAVE FUNCTIONS

To find the complete spatial description of the electron in a hydrogen atom, we must obtain three-dimensional wave functions. The Schrödinger equation in three-dimensional Cartesian coordinates has the following form:

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2}\right) + U(x,y,z)\psi(x,y,z) = E\psi(x,y,z) \qquad (7.9)$$

where ψ is a function of x, y, and z. The usual procedure for solving a partial differential equation of this type is to separate the variables by replacing a function of three variables with the product of three functions of one variable—for example, $\psi(x,y,z) = X(x)Y(y)Z(z)$. However, the Coulomb potential energy (Eq. 7.1) written in Cartesian coordinates, $U(x,y,z) = -e^2/4\pi\varepsilon_0\sqrt{x^2+y^2+z^2}$, does not lead to a separable solution.

For this calculation, it is more convenient to work in spherical polar coordinates (r, θ, ϕ) instead of Cartesian coordinates (x, y, z). The variables of spherical polar coordinates are illustrated in Figure 7.5. This simplification in the solution is at the expense of an increased complexity of the Schrödinger equation, which becomes:

$$-\frac{\hbar^2}{2m}\left[\frac{\partial^2\psi}{\partial r^2} + \frac{2}{r}\frac{\partial\psi}{\partial r} + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2}\right] \qquad (7.10)$$

$$+U(r)\psi(r,\theta,\phi) = E\psi(r,\theta,\phi)$$

where now ψ is a function of the spherical polar coordinates r, θ, and ϕ. When the potential energy depends only on r (and not on θ or ϕ), as is the case for the Coulomb potential energy, we can find solutions that are separable and can be factored as

$$\psi(r,\theta,\phi) = R(r)\Theta(\theta)\Phi(\phi) \qquad (7.11)$$

where the *radial function* $R(r)$, the *polar function* $\Theta(\theta)$, and the *azimuthal function* $\Phi(\phi)$ are each functions of a single variable. This procedure gives three differential equations, each of a single variable (r, θ, ϕ).

The quantum state of a particle that moves in a potential energy that depends only on r can be described by angular momentum quantum numbers l and m_l.

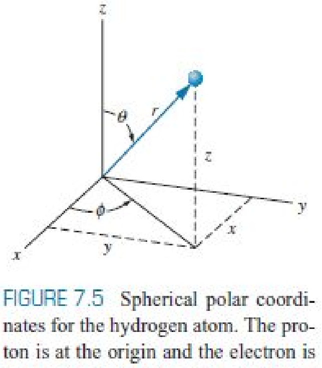


FIGURE 7.5 Spherical polar coordinates for the hydrogen atom. The proton is at the origin and the electron is at a radius r, in a direction determined by the polar angle θ and the azimuthal angle ϕ.

The polar and azimuthal solutions are given by combinations of standard trigonometric functions. The remaining radial function is then obtained from solving the radial equation:

$$-\frac{\hbar^2}{2m}\left(\frac{d^2R}{dr^2}+\frac{2}{r}\frac{dR}{dr}\right)+\left(-\frac{e^2}{4\pi\varepsilon_0 r}+\frac{l(l+1)\hbar^2}{2mr^2}\right)R(r)=ER(r) \qquad (7.12)$$

The mass that appears in this equation is the *reduced mass* of the proton-electron system defined in Eq. 6.44.

Quantum Numbers and Wave Functions

When we solve a three-dimensional equation such as the Schrödinger equation, three parameters emerge in a natural way as indices or labels for the solutions, just as the single index n emerged from our solution of the one-dimensional infinite well in Section 5.4. These indices are the three *quantum numbers* that label the solutions. The three quantum numbers that emerge from the solutions and their allowed values are:

n	principal quantum number	$1, 2, 3, \ldots$
l	angular momentum quantum number	$0, 1, 2, \ldots, n-1$
m_l	magnetic quantum number	$0, \pm 1, \pm 2, \ldots, \pm l$

The principal quantum number n is identical to the quantum number n that we obtained in the Bohr model. It determines the quantized energy levels:

$$E_n = -\frac{me^4}{32\pi^2\varepsilon_0^2\hbar^2}\frac{1}{n^2} \qquad (7.13)$$

which is identical to Eq. 6.30. Note that the energy depends only on n and not on the other quantum numbers l or m_l. The permitted values of the angular momentum quantum number l are limited by n (l ranges from 0 to $n-1$) and those of the magnetic quantum number m_l are limited by l.

Complete with quantum numbers, the separated solutions of Eq. 7.10 can be written

$$\psi_{n,l,m_l}(r,\theta,\phi) = R_{n,l}(r)\Theta_{l,m_l}(\theta)\Phi_{m_l}(\phi) \qquad (7.14)$$

The indices (n, l, m_l) are the three quantum numbers that are necessary to describe the solutions. Wave functions corresponding to some values of the quantum numbers are shown in Table 7.1. The wave functions are written in terms of the Bohr radius a_0 defined in Eq. 6.29.

For the ground state ($n = 1$), only $l = 0$ and $m_l = 0$ are allowed. The complete set of quantum numbers for the ground state is then $(n, l, m_l) = (1, 0, 0)$, and the wave function for this state is given in the first line of Table 7.1. The first excited state ($n = 2$) can have $l = 0$ or $l = 1$. For $l = 0$, only $m_l = 0$ is allowed. This state has quantum numbers (2, 0, 0), and its wave function is given in the second line of Table 7.1. For $l = 1$, we can have $m_l = 0$ or ± 1. There are thus three possible sets of quantum numbers: (2, 1, 0) and (2, 1, ± 1). The wave functions for these states are given in the third and fourth lines of Table 7.1. The second excited state ($n = 3$) can have $l = 0$ ($m_l = 0$), $l = 1$ ($m_l = 0, \pm 1$), or $l = 2$ ($m_l = 0, \pm 1, \pm 2$).

For the $n = 2$ level, there are four different possible sets of quantum numbers and correspondingly four different wave functions. All of these wave functions

TABLE 7.1 Some Hydrogen Atom Wave Functions

n	l	m_l	$R(r)$	$\Theta(\theta)$	$\Phi(\phi)$
1	0	0	$\frac{2}{a_0^{3/2}}e^{-r/a_0}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2\pi}}$
2	0	0	$\frac{1}{(2a_0)^{3/2}}\left(2-\frac{r}{a_0}\right)e^{-r/2a_0}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2\pi}}$
2	1	0	$\frac{1}{\sqrt{3}(2a_0)^{3/2}}\frac{r}{a_0}e^{-r/2a_0}$	$\sqrt{\frac{3}{2}}\cos\theta$	$\frac{1}{\sqrt{2\pi}}$
2	1	±1	$\frac{1}{\sqrt{3}(2a_0)^{3/2}}\frac{r}{a_0}e^{-r/2a_0}$	$\mp\frac{\sqrt{3}}{2}\sin\theta$	$\frac{1}{\sqrt{2\pi}}e^{\pm i\phi}$
3	0	0	$\frac{2}{(3a_0)^{3/2}}\left(1-\frac{2r}{3a_0}+\frac{2r^2}{27a_0^2}\right)e^{-r/3a_0}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2\pi}}$
3	1	0	$\frac{8}{9\sqrt{2}(3a_0)^{3/2}}\left(\frac{r}{a_0}-\frac{r^2}{6a_0^2}\right)e^{-r/3a_0}$	$\sqrt{\frac{3}{2}}\cos\theta$	$\frac{1}{\sqrt{2\pi}}$
3	1	±1	$\frac{8}{9\sqrt{2}(3a_0)^{3/2}}\left(\frac{r}{a_0}-\frac{r^2}{6a_0^2}\right)e^{-r/3a_0}$	$\mp\frac{\sqrt{3}}{2}\sin\theta$	$\frac{1}{\sqrt{2\pi}}e^{\pm i\phi}$
3	2	0	$\frac{4}{27\sqrt{10}(3a_0)^{3/2}}\frac{r^2}{a_0^2}e^{-r/3a_0}$	$\sqrt{\frac{5}{8}}(3\cos^2\theta-1)$	$\frac{1}{\sqrt{2\pi}}$
3	2	±1	$\frac{4}{27\sqrt{10}(3a_0)^{3/2}}\frac{r^2}{a_0^2}e^{-r/3a_0}$	$\mp\sqrt{\frac{15}{4}}\sin\theta\cos\theta$	$\frac{1}{\sqrt{2\pi}}e^{\pm i\phi}$
3	2	±2	$\frac{4}{27\sqrt{10}(3a_0)^{3/2}}\frac{r^2}{a_0^2}e^{-r/3a_0}$	$\frac{\sqrt{15}}{4}\sin^2\theta$	$\frac{1}{\sqrt{2\pi}}e^{\pm 2i\phi}$

correspond to the same energy, so the $n = 2$ level is *degenerate*. (Degeneracy was introduced in Section 5.4.) The $n = 3$ level is degenerate with nine possible sets of quantum numbers. In general, the level with principal quantum number n has a degeneracy equal to n^2. Figure 7.6 illustrates the labeling of the first three levels.

If different combinations of quantum numbers have exactly the same energy, what is the purpose of listing them separately? First, as we discuss in the last section of this chapter, the levels are not precisely degenerate, but are separated

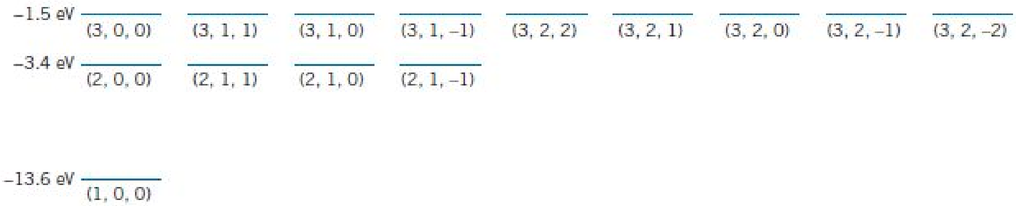


FIGURE 7.6 The lower energy levels of hydrogen, labeled with the quantum numbers (n, l, m_l). The first excited state is four-fold degenerate and the second excited state is nine-fold degenerate.

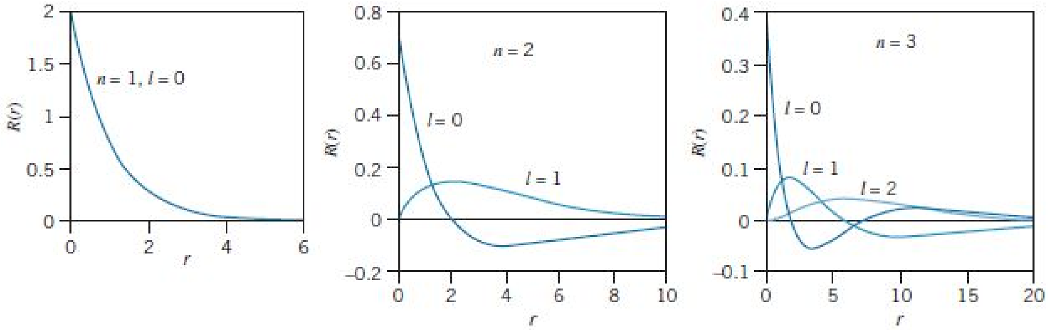


FIGURE 7.7 The radial wave functions of the $n = 1$, $n = 2$, and $n = 3$ states of hydrogen. The radius coordinate is measured in units of a_0.

by a very small energy (about 10^{-5} eV). Second, in the study of the transitions between the levels, we find that the intensities of the individual transitions depend on the quantum numbers of the particular level from which the transition originates. Third, and perhaps most important, *each of these sets of quantum numbers corresponds to a very different wave function, and therefore represents a very different state of motion of the electron.* These states have different spatial probability distributions for locating the electron, and thus can affect many atomic properties—for example, the way two atoms can form molecular bonds.

The radial wave functions for the states listed in Table 7.1 are plotted in Figure 7.7. You can readily see the differences in the motion of the electron for the different states. For example, in the $n = 2$ level, the $l = 0$ and $l = 1$ wave functions have the same energy but their behavior is very different: the $l = 1$ wave function falls to zero at $r = 0$, but the $l = 0$ wave function remains nonzero at $r = 0$. The $l = 0$ electron thus has a much greater probability of being found close to (or even inside) the nucleus, which turns out to play a large role in determining the rates for certain radioactive decay processes.

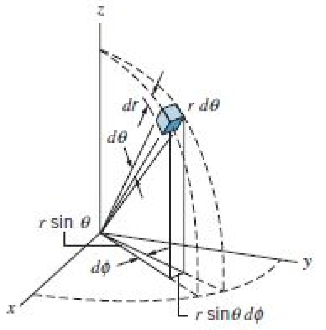


FIGURE 7.8 The volume element in spherical polar coordinates.

Probability Densities

As we learned in Chapter 5, the probability of finding the electron in any spatial interval is determined by the square of the wave function. For the hydrogen atom, $|\psi(r,\theta,\phi)|^2$ gives the *volume probability density* (probability per unit volume) at the location (r,θ,ϕ). To compute the actual probability of finding the electron, we multiply the probability per unit volume by the volume element dV located at (r,θ,ϕ). In spherical polar coordinates (see Figure 7.8) the volume element is

$$dV = r^2 \sin\theta \, dr \, d\theta \, d\phi \tag{7.15}$$

and therefore the probability to find the electron in the volume element at that location is

$$|\psi_{n,l,m_l}(r,\theta,\phi)|^2 \, dV = |R_{n,l}(r)|^2 |\Theta_{l,m_l}(\theta)|^2 |\Phi_{m_l}(\phi)|^2 r^2 \sin\theta \, dr \, d\theta \, d\phi \tag{7.16}$$