

Chapter 4

Macroscopic and Microscopic States

The treatment given in the preceding chapter was limited to the problems connected with the properties of ideal gases. We now proceed to describe a more general method for dealing with the problems in statistical physics. The method can be applied to a system much more complicated than a gas system, and of macroscopic dimensions, whose constituent particles may interact with one another.

A system is generally characterized by the physical and chemical properties of the substance in the space occupied by it, as well as by the properties of its boundaries. In general, we are concerned with systems consisting of a large number of particles, such as, for example, a gas comprising a large number of molecules contained in a vessel; or a crystal consisting of a large number of atoms in a lattice, etc. The properties of a system depend upon the state in which it exists at the instant at which measurements are made. Our first concern, therefore, will be how to specify the state of a system at any time.

4.1 MACROSCOPIC STATES

Logically, since all systems obey quantum mechanics, we should start with quantum statistical physics and then arrive at classical statistical physics as a particular case. But it is not always necessary to use quantum mechanics in describing the behaviour of a thermodynamic system. Newtonian mechanics is an excellent approximation and, hence, it is convenient to commence the study of statistical mechanics with classical mechanical consideration.

To explain the concept of the state of a system, we will once more refer to a monatomic ideal gas, which is relatively a simple system. By an ideal gas we mean that its constituent particles have a finite mass, undergo elastic collisions with one another and with the boundary walls and have no other interaction of any kind. We further assume that the vessel containing the gas is isolated. That is, it does not exchange energy with the material bodies outside and, hence, all variations in the gas are due to internal reasons. The time spent in collision is very much shorter than the time lapse between collisions. It must be mentioned here that real gases are not ideal in the sense we have described above. They approximate to the ideal gases when the pressure is very low. Under this condition, the particle density is so low that the particles are relatively distant from each other. If a gas is left to itself for a sufficiently long time, the pressure and the temperature of the gas, whatever their initial distribution, stabilize over the entire volume and we say that the gas has attained a steady state. In this state the pressure (P), temperature (T) and volume (V) remain constant in time. This state of the gas, characterized by the parameters, P , T and V , is called its *macroscopic state*. The macroscopic steady state, in this particular case is also called its equilibrium state. Note that a steady state is not always an equilibrium state. For a system that is not isolated, it is possible to establish a

steady state by maintaining the different parts of the walls of the vessel at different temperatures with the help of external sources. We will, however, restrict ourselves in the first few chapters to the study of equilibrium states, as these provide a great simplification. Besides, most systems we encounter are in thermodynamic equilibrium, or else they tend to reach very rapidly to a state of equilibrium.

4.2 MICROSCOPIC STATES

Since a gas consists of a large number of particles, the state of a gas can also be characterized in terms of the states of the constituent particles. The states of the particles can be specified by ascertaining their positions (coordinates q_i) and velocities or momenta p_i . The state of the gas thus described in terms of the properties of its constituent particles is called its *microscopic state*.

We have seen above that the quantities pressure, temperature and volume which describe the macroscopic state remain constant in the steady state of the gas. The constituent particles of the gas, however, are continuously in motion and undergo collisions even in a steady state and, hence, microscopic states keep on changing continuously. Let us be more clear on this point. The microscopic state of a gas at any instant is defined by the instantaneous positions and momenta (in magnitude and direction) of the various molecules. If any change is made in the molecular distribution, a new microscopic state results. Many of the different microscopic states, however, will be indistinguishable from the standpoint of macroscopic observation. For instance, if the positions and momenta of two or more molecules are interchanged, a new microscopic state will result; yet from the macroscopic point of view, no change will be detected in the condition of the gas. Each one of the microscopic states is said to pertain to the same macroscopic state. We are thus led to associate with each macroscopic state a large number of microscopic states, although they widely differ in respect of their dynamical states.

Consider, for example, an ideal gas system in which particles can exist in any one of the states with energy equal to 0, 1, 2, 3 energy units and no other. Suppose the system consists of three distinguishable particles a , b and c and the total energy of the system is three energy units. In how many ways three energy units can be distributed among the three distinguishable particles, consistent with the condition that the total energy of the system is three units? In other words, how many microscopic states correspond to the macroscopic state specified by three energy units?

Figure 4.1 shows the possible configurations of the system and the number of microscopic states.

Configuration	Number of states	
	Distinguishable particles	Indistinguishable particles
3 2 1 0	3 ..	1
3 2 1 0	6 ..	1
3 2 1 0	1 ..	1
	Total number of states	10 ..
		3 ..

The total number of accessible microscopic states corresponding to the single macroscopic state is ten if the particles are distinguishable and three if they are indistinguishable.

One of the aims of statistical mechanics is to correlate the properties of matter in bulk with the properties of its constituent particles, i.e. to establish a relation between a macroscopic state and the corresponding microscopic states.

4.3 PHASE SPACE

We shall now show how the state of a system is specified. We assume that the physical systems obey the most general laws of dynamics as set down by Hamilton. The state of a system is determined when we know the details of its motion and, hence, is fully specified at a particular time by a certain number of generalised coordinates q_i and an equal number of generalized momenta p_i . These coordinates and momenta are governed by the canonical equation of Hamilton:

$$\dot{q}_i = \frac{\partial H}{\partial p_i}; \dot{p}_i = -\frac{\partial H}{\partial q_i} \quad (i = 1, 2, \dots) \quad (4.1)$$

where q_i are the generalized coordinates

p_i are the generalized momenta

and H is the Hamiltonian of the system

A most elementary physical realization of such a system is furnished by a single particle constrained to move in a straight line. The position of the particle at any time can be described by giving its distance from a fixed point on the line — origin of coordinates. This single numerical fact merely gives its 'mechanical configuration' at the time in question. For the complete specification of its state, its momentum at this instant must also be given. Since in elementary physics mass is assumed to be constant, momentum is proportional to the velocity and, hence, it may seem that it matters little which of the two is used! However, it must be emphasized that the generalized momenta used in Hamiltonian equations are very different from the mass-velocity products of elementary mechanics.

The state of a particle at any time is given by the pair of values $(q, p) = (\text{position, momentum})$ determined by the differential equations of dynamics [Eq. (4.1)]. As the point moves on a straight line, the values of number pair (q, p) will change. Quasi-geometrical terminology makes for convenient and lucid expression and considerably facilitates the study of the changes in the state as time goes on. Any number pair can be plotted as the rectangular coordinates of a point in a Cartesian graph (Fig. 4.2).

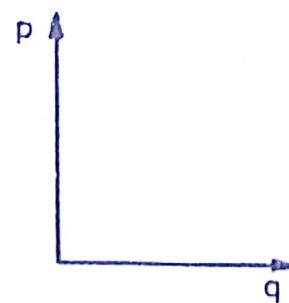


Fig. 4.2

Every such point which corresponds to a certain value of coordinate q_i and momentum p_i of the system, describes both the position and momentum of the system at a particular instant, and, hence, represents a certain state of the system.

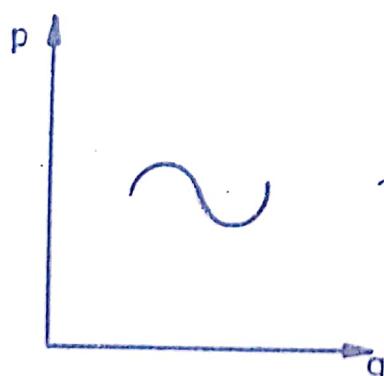


Fig. 4.3

There will be different points for different states. As the particle moves on the straight line, its momentum and coordinate vary in accordance with Hamilton's equation [Eq. (4.1)]. Its evolution through the successive states will be represented by a trajectory in the Cartesian plane (Fig. 4.3).

The state of the system when represented in this way, is referred to as its *phase*, the representative point as the *phase point*, the trajectory as the *phase path* and the Cartesian plane as the *phase space*.

It is not necessary to start with Cartesian coordinates. Polar, cylindrical or any other suitably defined set of coordinates would do as well. Nevertheless, we shall find that the phase space with rectangular axes corresponding to each coordinate and momentum is very convenient, as it has simple properties.

Example 4.1 Determine the phase trajectory of a bullet of unit mass fired

4.9 LIOUVILLE'S THEOREM

Let us now inquire into the rate of change of phase point density in a classical system.

Consider a small volume element in the hyperspace, bounded by q_1 and $q_1 + \delta q_1$, p_1 and $p_1 + \delta p_1$, etc. The number of systems in the element using Eq. (4.23) is

$$\textcircled{3} \quad \delta N = \rho d\Gamma = \rho(q, p, t) \delta q_1 \dots \delta p_f$$

The motion of every system is governed by Hamilton's equations. Representative points will enter and leave through the boundary surfaces of the element. Since the number of phase points entering the element of hyper volume through any 'face' will, in general, be different from the number of points which are leaving the

opposite 'face', the number δN will be changing. The change is equal to $\frac{\partial \rho}{\partial t} dt \delta q_1 \dots \delta p_f$

We can construct a conceptual Euclidean space of $2f$ dimensions. Consider two faces perpendicular to the q_1 -axis which are located at q_1 and $q_1 + \delta q_1$. The number of phase points entering the element in time dt through the face $q_1 = \text{constant}$ is (Fig. 4.7)

$$= \rho(q, p, t) \dot{q}_1 dt \delta q_2 \dots \delta p_f \quad (4.28)$$

where \dot{q}_1 is the velocity in the direction q_1 .

The number leaving the element through the face $q_1 + \delta q_1 = \text{constant}$ is

$$= \left[\rho \dot{q}_1 + \frac{\partial}{\partial q_1} (\rho \dot{q}_1) \delta q_1 \right] dt \delta q_2 \dots \delta p_f \quad (4.29)$$

Hence, the effective number of systems entering the element in the direction q_1 is

$$\begin{aligned} &= \rho \dot{q}_1 dt \delta q_2 \dots \delta p_f - \left[\rho \dot{q}_1 + \frac{\partial}{\partial q_1} (\rho \dot{q}_1) \delta q_1 \right] dt \delta q_2 \dots \delta p_f \\ &= - \frac{\partial}{\partial q_1} (\rho \dot{q}_1) dt \delta q_1 \delta q_2 \dots \delta p_f \end{aligned} \quad (4.30)$$

Macroscopic and Microscopic States

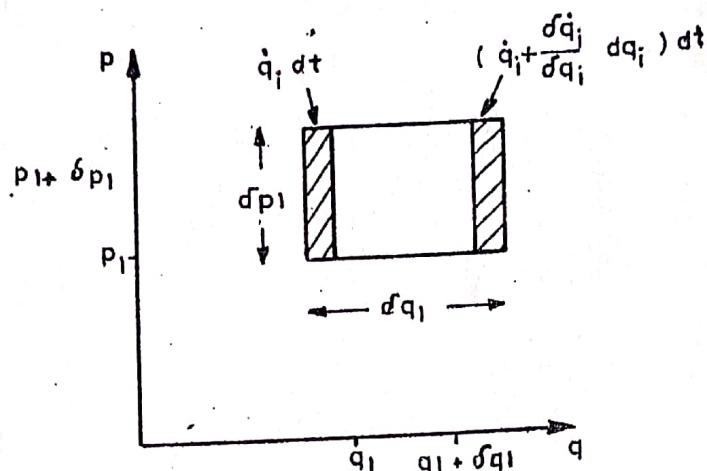


Fig. 4.7

Therefore, the net number entering the volume element through all its faces is given by

$$\frac{\partial p}{\partial t} dt \delta q_1 \dots \delta q_f = \left[- \sum_{i=1}^f \frac{\partial}{\partial q_i} (\rho \dot{q}_i) - \sum_{i=1}^f \frac{\partial}{\partial p_i} (\rho \dot{p}_i) \right] dt \delta q_1 \dots \delta p_f$$

We, thus, obtain the relation

$$\frac{\partial \rho}{\partial t} = - \sum_{i=1}^f \left[\frac{\partial}{\partial q_i} (\rho \dot{q}_i) + \frac{\partial}{\partial p_i} (\rho \dot{p}_i) \right] \quad (4.31)$$

This result can be simplified to yield

$$\frac{\partial \rho}{\partial t} = - \sum_{i=1}^f \left[\rho \left(\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} \right) + \left(\frac{\partial \rho}{\partial q_i} \dot{q}_i + \frac{\partial \rho}{\partial p_i} \dot{p}_i \right) \right] \quad (4.32)$$

From Hamilton's equations of motion

$$\frac{\partial \dot{q}_i}{\partial q_i} = \frac{\partial}{\partial q_i} \frac{\partial H}{\partial p_i} = \frac{\partial^2 H}{\partial q_i \partial p_i}$$

$$\frac{\partial \dot{p}_i}{\partial p_i} = - \frac{\partial}{\partial p_i} \frac{\partial H}{\partial q_i} = - \frac{\partial^2 H}{\partial q_i \partial p_i}$$

Therefore, the first term in the bracket vanishes and

$$\left(\frac{\partial \rho}{\partial t} \right)_{q,p} = - \sum_{i=1}^f \left(\frac{\partial \rho}{\partial q_i} \dot{q}_i + \frac{\partial \rho}{\partial p_i} \dot{p}_i \right) \quad (4.33)$$

The symbol of partial differentiation indicates that $\left(\frac{\partial \rho}{\partial t} \right)_{q,p}$ gives the rate of the change in density at the point of interest. This is known as *Liouville's equation* and is of fundamental importance in statistical mechanics.

It is easy to see that the equation can be written in any of the following forms:

$$(i) \quad \frac{\partial \rho}{\partial t} + \sum_{i=1}^f \left(\frac{\partial \rho}{\partial q_i} \dot{q}_i + \frac{\partial \rho}{\partial p_i} \dot{p}_i \right) = 0 \quad (4.34)$$

$$(ii) \quad \frac{\partial \rho}{\partial t} + \sum_{i=1}^f \left(\frac{\partial \rho}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial \rho}{\partial p_i} \frac{\partial H}{\partial q_i} \right) = 0 \quad (4.35)$$

$$(iii) \quad \frac{\partial \rho}{\partial t} + \{\rho, H\} = 0 \quad \dot{q} = \frac{\partial H}{\partial p} \quad \dot{p} = -\frac{\partial H}{\partial q} \quad (4.36)$$

where

$$\{\rho, H\} = \sum_i \left(\frac{\partial \rho}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial \rho}{\partial p_i} \frac{\partial H}{\partial q_i} \right)$$

is known as the Poisson bracket of the functions ρ and H .

The total time derivative of $\rho (q, p, t)$ is given by

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \sum_{i=1}^f \left(\frac{\partial \rho}{\partial q_i} \dot{q}_i + \frac{\partial \rho}{\partial p_i} \dot{p}_i \right) = 0 \quad (4.37)$$

In other words the density of phase points is an integral of motion. This is *Liouville's theorem* in classical presentation.

of conservation of extension.

4.11 CONDITION FOR STATISTICAL EQUILIBRIUM

Liouville's fundamental theorem which controls the temporal behaviour of ensembles, permits us to discuss the conditions for statistical equilibrium.

The ensemble is considered to be in statistical equilibrium if ρ has no explicit dependence on time at all points in phase space, i.e.

$$\frac{\partial \rho}{\partial t} = 0 \quad (4.40)$$

Under the condition of equilibrium, therefore, by Eq. (4.32)

$$\sum_{i=1}^f \left(\frac{\partial \rho}{\partial q_i} \dot{q}_i + \frac{\partial \rho}{\partial p_i} \dot{p}_i \right) = 0 \quad (4.41)$$

This equation will be satisfied if ρ is independent of coordinates (q, p) , i.e.

$$\rho(q, p) = \text{constant} \quad (4.42)$$

In other words, representative points are uniformly distributed in the phase space.

More generally, under statistical equilibrium, the density must be a function of the constants of motion only. Thus, if α is a constant of motion, $\rho = \rho(\alpha)$, so that

$$\frac{\partial \rho}{\partial t} = \frac{\partial \rho}{\partial \alpha} \frac{\partial \alpha}{\partial t} = \frac{\partial \rho}{\partial \alpha} \sum_{i=1}^f \left(\frac{\partial \alpha}{\partial q_i} \dot{q}_i + \frac{\partial \alpha}{\partial p_i} \dot{p}_i \right) = 0 \quad (4.43)$$

It may be noted that when the system is in statistical equilibrium, the observed values of the macroscopic physical quantities are to a high degree of accuracy, equal to their mean values.

Chapter 5

Statistical Ensembles

In the formulation of statistical mechanics, Gibbs introduced three standard ensembles to which real experiments could be approximated. These are: (i) the microcanonical ensemble, (ii) the canonical ensemble and (iii) the grand canonical ensemble. Their classification depends on the manner in which their systems interact. Each of them has its own characteristic distribution. Physical systems can interact in a variety of ways; in particular, they may exchange energy, or matter or both with each other. The ensemble in which the systems exchange energy, but not matter, is called a *canonical ensemble*, that in which both energy and matter are exchanged between the systems is called a *grand canonical ensemble*; and that in which neither energy nor matter is exchanged, is called a *microcanonical ensemble*. Statistical mechanics provides a number of methods for calculating equilibrium thermodynamic properties of macroscopic systems. Explicit calculations of thermodynamic functions can be carried out using microcanonical, canonical or grand canonical ensembles. A specific choice of ensembles may be thought of as corresponding to a particular physical situation.

5.1 MICROCANONICAL ENSEMBLE

This ensemble consists of systems which are isolated from the rest of the world. Such a system is also known as a 'closed isolated system' and has a fixed volume, fixed total energy and a fixed total number of particles. The probability density $P(p, q)$ of such a system differs from zero only on the constant energy hypersurface. It must be noted, however, that in reality, one cannot achieve complete isolation of a system. We must allow for some interaction energy ΔE , though very small. The elements of the microcanonical ensemble, therefore, lie within the range between E and $E + \Delta E$.

The picture of a microcanonical distribution in phase space would be something like a very thin uniform cloud. One can easily see this by referring to Fig. 5.1, which gives the trajectory of a simple harmonic oscillator. The ellipse E represents the evolution of our conservative system with total energy E . This system is just one member of the ensemble. The other members may have energy between E and $E + \delta E$ and, hence their corresponding phase points will be between two ellipses. The thin elliptic cloud is the energy shell. If we conceive of δE as getting smaller and smaller, the energy shell between the two surfaces would become just a surface in the limit as $\delta E \rightarrow 0$.

According to the fundamental postulate of equal a priori probability, under the condition of equilibrium the system is equally likely to be found in one of its accessible states. In the case of microcanonical ensemble all states between E and

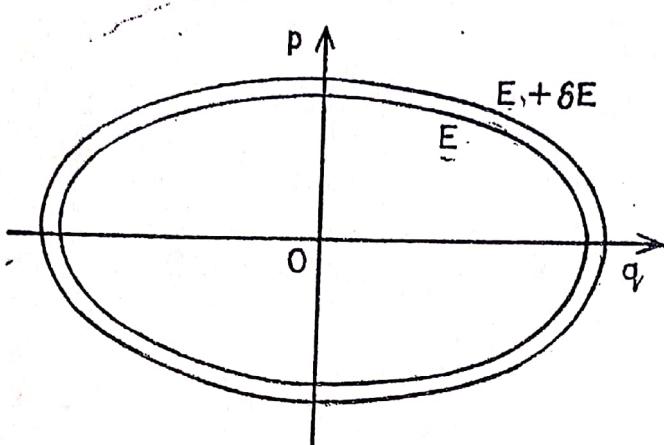


Fig. 5.1

$E + \delta E$ are equally accessible. Therefore, if the system is in a state X , corresponding to the energy E_x , the probability P_x of finding the system in state X is given by

$$P_x = \begin{cases} C & \text{if } E < E_x < E + \delta E \\ 0 & \text{otherwise} \end{cases} \quad (5.1)$$

where C is a constant, the value of which can be determined from the normalization condition $\sum P_x = 1$, when summed over all accessible states in the range between E and $E + \delta E$. It should be more appropriate to write this relation as $P_x = C \delta(E - E_x)$, where δ is the Dirac delta function.

5.2 CANONICAL ENSEMBLE

It is evident that the microcanonical ensemble is fundamental and refers to the simple situation in which the constituent systems are isolated and not influenced in any way by external disturbances. One would, therefore, think it proper to use such ensembles in the development of the basic theory. However, as pointed out above, microcanonical ensembles are best suited for isolated systems, and we seldom deal with a completely isolated system; there is always some energy exchange with surroundings. It would be more realistic, therefore, to use an ensemble in which statistical equilibrium is attained not by isolation but by free interaction. Canonical ensembles have been found to be more appropriate for the description of such systems. Besides, canonical approach gives results which apply even when components interact strongly.

Suppose we desire to study the statistical properties of a thermodynamical system A . To this end we envisage an assembly of a very large number of systems (N) identical in volume and molecular composition with A . Each system of the assembly is considered to be in thermal contact with a huge heat reservoir. A heat reservoir is a system whose heat capacity is so much higher than that of the subsystem in contact with it, that heat flow from or to the heat reservoir does not change its temperature significantly. The assembly can be thought of as contained

within a microcanonical ensemble. Each system of the assembly is in contact with the other elements of the microcanonical ensemble, but is otherwise isolated from the outside world. The walls of the containers of the system of the assembly which separate it from microcanonical ensemble are diathermic walls, which allow heat to flow through them but do not allow the exchange of mass. An assembly of this kind is called a canonical ensemble of systems of type A. The number of systems in a canonical ensemble is constant; but their energy is not fixed.

Let us now find the probability that a canonical system, under the condition of equilibrium, is in a particular state with specific energy E_α .

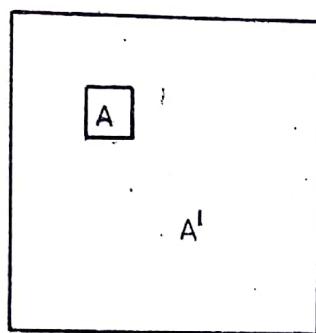


Fig. 5.2

Consider a canonical system A in a microcanonical ensemble A' ($A \ll A'$) (Fig. 5.2). The walls of A are such that A and A' are free to exchange energy. The microcanonical ensemble is isolated and its total energy E_0 is constant. As discussed before, a large number of microscopic states correspond to the macroscopic state of the microcanonical ensemble with energy E_0 . Let Ω_0 be the total number of microstates of the microcanonical ensemble with energy E_0 . By the postulate of equal a priori probability, all these microstates are equally probable. Therefore, the probability

of each microstate is $\frac{1}{\Omega_0}$. It is obvious that different microstates of the

microcanonical ensemble A' will give rise to different microstates of the small subsystem A. Suppose, a certain microstate of A' gives rise to a microstate of A with energy E_α ($E_\alpha \ll E_0$). Hence, the energy of the remaining part of the complete system is $E_0 - E_\alpha$. This is one microstate of A with energy E_α , attained through the other microstates of A'. Let the number of states of the microcanonical ensemble A', through which the states of the small system A corresponding to the energy E_α are attained, be $\Omega_\alpha(E_0 - E_\alpha)$. Therefore, the probability that the small system A is in a state with energy E_α is

$$P_\alpha = \frac{\Omega_\alpha(E_0 - E_\alpha)}{\Omega_0} \quad (5.2)$$

Using the relation

$$a = e^{\ln a}$$

we can express Eq. (5.2) as

$$P_\alpha = \frac{e^{\ln \Omega_\alpha(E_0 - E_\alpha)}}{\Omega_0} \quad (5.3)$$

Expanding $\ln \Omega_\alpha(E_0 - E_\alpha)$ into a Taylor's series

$$\begin{aligned} \ln \Omega_\alpha(E_0 - E_\alpha) &= \ln \Omega_\alpha(E_0) - E_\alpha \frac{\partial \ln \Omega_\alpha(E_0)}{\partial E_0} \\ &\quad + \text{higher order terms} \end{aligned} \quad (5.4)$$

Putting $\beta = \frac{\partial \ln \Omega_\alpha(E_0)}{\partial E_0}$ and neglecting higher order terms

$$\ln \Omega_\alpha(E_0 - E_\alpha) = \ln \Omega_\alpha(E_0) - \beta E_\alpha$$

Hence

$$\begin{aligned} P_\alpha &= \frac{e^{\{\ln \Omega_\alpha(E_0) - \beta E_\alpha\}}}{\Omega_0} = \frac{e^{\ln \Omega_\alpha(E_0)}}{\Omega_0} e^{-\beta E_\alpha} \\ &= \frac{\Omega_\alpha(E_0)}{\Omega_0} e^{-\beta E_\alpha} = \text{const} \times e^{-\beta E_\alpha} \end{aligned} \quad (5.5)$$

From the normalization condition $\sum P_\alpha = \text{const} \times \sum e^{-\beta E_\alpha} = 1$

Therefore, $\text{const} = \frac{1}{\sum e^{-\beta E_\alpha}}$ (5.6)

and $P_\alpha = \frac{e^{-\beta E_\alpha}}{\sum e^{-\beta E_\alpha}}$ (5.7)

This is the Gibbs canonical distribution. It gives the probability that the small system A be in a state with energy E_α . There may be other microstates of A with the same energy E_α . If this number is g_α :

$$P_\alpha = \frac{g_\alpha e^{-\beta E_\alpha}}{\sum e^{-\beta E_\alpha}} \quad (5.8)$$

The denominator is usually represented by Z , i.e.

$$Z = \sum_\alpha e^{-\beta E_\alpha} \quad (5.9)$$

Therefore $P_\alpha = \frac{g_\alpha e^{-\beta E_\alpha}}{Z}$ (5.10)

If E is a continuous variable

$$Z = \int_0^\infty e^{-\beta E} dE \quad (5.11)$$

We have yet to identify the parameter β , which we will shortly do. Meanwhile, we shall discuss here an example which gives a clue towards the identification of β .

Consider a composite system A^0 made up of two systems A and A' in thermal contact and weakly interacting with each other. Let A be in a state r with energy E_r and A' in a state s with energy E_s . The energy of the composite system A^0 , therefore, is $E^0 = E_r + E_s$.

5.5 GRAND CANONICAL ENSEMBLE

We now consider a situation slightly different from the one considered in the case of canonical distribution. Imagine as before a sub-system A in a system A^1 with a difference that the walls of the sub-systems are such that A and A^1 can exchange both energy and particles. Such systems are known as *open systems*. A^1 is, of course, a member of a microcanonical ensemble. If E_0, N_0 , are the energy and number of particles of the entire system

$$E_0 = E_\alpha + E^1 \quad (5.39)$$

$$N_0 = N_\alpha + N^1$$

where E_α, E^1 and N_α, N^1 are the energies and number of particles in A and A^1 respectively, and none of them is constant. In deriving the probability distribution, we adopt the same technique as in the case of canonical distribution.

The probability that the sub-system A is in a state with energy E_α and number of particles N_α , is given by

$$P_\alpha = C \Omega_\alpha(E_0 - E_\alpha, N_0 - N_\alpha) \quad (5.40)$$

where C is a constant and Ω_α has the same meaning as in the case of canonical distribution.

We expand $\ln \Omega_\alpha(E_0 - E_\alpha, N_0 - N_\alpha)$ in a power series

$$\begin{aligned} \ln \Omega_\alpha(E_0 - E_\alpha, N_0 - N_\alpha) &= \ln \Omega_\alpha(E_0, N_0) \\ &\quad - \frac{\partial(\ln \Omega_\alpha)}{\partial E_0} E_\alpha - \frac{\partial(\ln \Omega_\alpha)}{\partial N_0} N_\alpha + \dots \\ &= \ln \Omega_\alpha(E_0, N_0) - \beta E_\alpha + \alpha N_\alpha \end{aligned} \quad (5.41)$$

where we have put

$$\beta = \frac{\partial(\ln \Omega_\alpha)}{\partial E_0}, \quad \alpha = \frac{\partial(\ln \Omega_\alpha)}{\partial N_0} \quad (5.42)$$

Hence

$$\Omega_\alpha(E_0 - E_\alpha, N_0 - N_\alpha) = \Omega_\alpha(E_0, N_0) e^{-\beta E_\alpha + \alpha N_\alpha}$$

Therefore

$$P_\alpha = C e^{-\beta E_\alpha + \alpha N_\alpha} \quad (5.43)$$

The normalization condition gives

$$C = \frac{1}{\sum_\alpha e^{-\beta E_\alpha + \alpha N_\alpha}}$$

putting $\alpha = \mu \beta$

and hence

$$P_\alpha = \frac{e^{-\beta E_\alpha + \mu N_\alpha \beta}}{\sum_\alpha e^{-\beta E_\alpha + \mu N_\alpha \beta}} \quad (5.45)$$

This distribution is called the *grand canonical distribution*. We now introduce here a quantity

$$z = e^{-\beta \mu} \quad (5.46)$$

which is generally known as the *fugacity* of the system. In the chemical literature it is known as the *absolute activity*.

In terms of fugacity the relation (5.45) becomes

$$P_\alpha = \frac{z^{N_\alpha} e^{-\beta E_\alpha}}{\sum_\alpha z^{N_\alpha} e^{-\beta E_\alpha}} \quad (5.47)$$

The denominator in Eq. (5.47) is called the *grand partition function*.

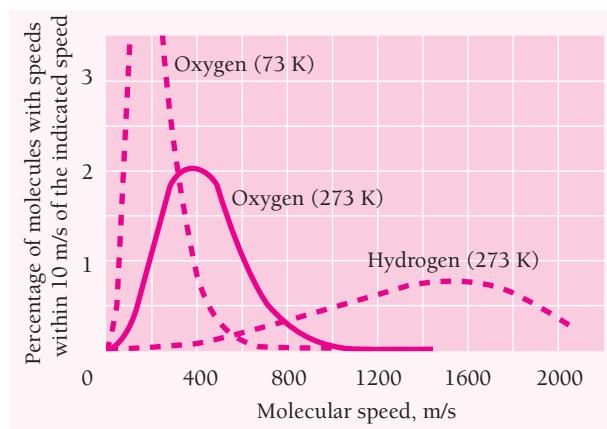
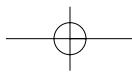


Figure 9.4 The distributions of molecular speeds in oxygen at 73 K, in oxygen at 273 K, and in hydrogen at 273 K.

Molecular speeds in a gas vary considerably on either side of v_p . Figure 9.4 shows the distribution of speeds in oxygen at 73 K (-200°C), in oxygen at 273 K (0°C), and in hydrogen at 273 K. The most probable speed increases with temperature and decreases with molecular mass. Accordingly molecular speeds in oxygen at 73 K are on the whole less than at 273 K, and at 273 K molecular speeds in hydrogen are on the whole greater than in oxygen at the same temperature. The average molecular energy is the same in both oxygen and hydrogen at 273 K, of course.

Example 9.4

Find the rms speed of oxygen molecules at 0°C .

Solution

Oxygen molecules have two oxygen atoms each. Since the atomic mass of oxygen is 16.0 u, the molecular mass of O_2 is 32.0 u which is equivalent to

$$m = (32.0 \text{ u})(1.66 \times 10^{-27} \text{ kg/u}) = 5.31 \times 10^{-26} \text{ kg}$$

At an absolute temperature of 273 K, the rms speed of an O_2 molecule is

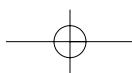
$$v_{\text{rms}} = \sqrt{\frac{3kT}{m}} = \sqrt{\frac{3(1.38 \times 10^{-23} \text{ J/K})(273 \text{ K})}{5.31 \times 10^{-26} \text{ kg}}} = 461 \text{ m/s}$$

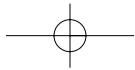
This is a little over 1000 mi/h.

9.4 QUANTUM STATISTICS

Bosons and fermions have different distribution functions

As mentioned in Sec. 9.1, the Maxwell-Boltzmann distribution function holds for systems of identical particles that can be distinguished one from another, which means particles whose wave functions do not overlap very much. Molecules in a gas fit this





description and obey Maxwell-Boltzmann statistics. If the wave functions do overlap appreciably, the situation changes because the particles cannot now be distinguished, although they can still be counted. The quantum-mechanical consequences of indistinguishability were discussed in Sec. 7.3, where we saw that systems of particles with overlapping wave functions fall into two categories:

- 1 Particles with 0 or integral spins, which are **bosons**. Bosons do not obey the exclusion principle, and the wave function of a system of bosons is not affected by the exchange of any pair of them. A wave function of this kind is called **symmetric**. Any number of bosons can exist in the same quantum state of the system.
- 2 Particles with odd half-integral spins ($\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots$), which are **fermions**. Fermions obey the exclusion principle, and the wave function of a system of fermions changes sign upon the exchange of any pair of them. A wave function of this kind is called **antisymmetric**. Only one fermion can exist in a particular quantum state of the system.

We shall now see what difference all this makes in the probability $f(\epsilon)$ that a particular state of energy ϵ will be occupied.

Let us consider a system of two particles, 1 and 2, one of which is in state a and the other in state b . When the particles are distinguishable there are two possibilities for occupancy of the states, as described by the wave functions

$$\psi_I = \psi_a(1)\psi_b(2) \quad (9.17)$$

$$\psi_{II} = \psi_a(2)\psi_b(1) \quad (9.18)$$

When the particles are not distinguishable, we cannot tell which of them is in which state, and the wave function must be a combination of ψ_I and ψ_{II} to reflect their equal likelihoods. As we found in Sec. 7.3, if the particles are bosons, the system is described by the symmetric wave function

$$\text{Bosons} \quad \psi_B = \frac{1}{\sqrt{2}} [\psi_a(1)\psi_b(2) + \psi_a(2)\psi_b(1)] \quad (9.19)$$

and if they are fermions, the system is described by the antisymmetric wave function

$$\text{Fermions} \quad \psi_F = \frac{1}{\sqrt{2}} [\psi_a(1)\psi_b(2) - \psi_a(2)\psi_b(1)] \quad (9.20)$$

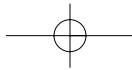
The $1/\sqrt{2}$ factors are needed to normalize the wave functions.

Now we ask what the likelihood in each case is that both particles be in the same state, say a . For distinguishable particles, both ψ_I and ψ_{II} become

$$\psi_M = \psi_a(1)\psi_a(2) \quad (9.21)$$

to give a probability density of

$$\text{Distinguishable particles} \quad \psi_M^* \psi_M = \psi_a^*(1)\psi_a^*(2)\psi_a(1)\psi_a(2) \quad (9.22)$$



For bosons the wave function becomes

$$\psi_B = \frac{1}{\sqrt{2}} [\psi_a(1)\psi_a(2) + \psi_a(1)\psi_a(2)] = \frac{2}{\sqrt{2}} \psi_a(1)\psi_a(2) = \sqrt{2} \psi_a(1)\psi_a(2) \quad (9.23)$$

to give a probability density of

$$\text{Bosons} \quad \psi_B^* \psi_B = 2\psi_a^*(1)\psi_a^*(2)\psi_a(1)\psi_a(2) = 2\psi_M^* \psi_M \quad (9.24)$$

Thus the probability that both bosons be in the same state is twice what it is for distinguishable particles!

For fermions the wave function becomes

$$\text{Fermions} \quad \psi_F = \frac{1}{\sqrt{2}} [\psi_a(1)\psi_a(2) - \psi_a(1)\psi_a(2)] = 0 \quad (9.25)$$

It is impossible for both particles to be in the same state, which is a statement of the exclusion principle.

These results can be generalized to apply to systems of many particles:

- 1 In a system of bosons, the presence of a particle in a certain quantum state *increases* the probability that other particles are to be found in the same state;
- 2 In a system of fermions, the presence of a particle in a certain state *prevents* any other particles from being in that state.

Bose-Einstein and Fermi-Dirac Distribution Functions

The probability $f(\epsilon)$ that a boson occupies a state of energy ϵ turns out to be

$$\text{Bose-Einstein distribution function} \quad f_{BE}(\epsilon) = \frac{1}{e^{\alpha e^{\epsilon/kT}} - 1} \quad (9.26)$$

and the probability for a fermion turns out to be

$$\text{Fermi-Dirac distribution function} \quad f_{FD}(\epsilon) = \frac{1}{e^{\alpha e^{\epsilon/kT}} + 1} \quad (9.27)$$

Names of the Functions

The Indian physicist S. N. Bose in 1924 derived Planck's radiation formula on the basis of the quantum theory of light with indistinguishable photons whose number is not conserved. His paper was rejected by a leading British journal. He then sent it to Einstein, who translated the paper into German and submitted it to a German journal where it was published. Because Einstein extended Bose's treatment to material particles whose number is conserved, both names are attached to Eq. 9.26. Two years later Enrico Fermi and Paul Dirac independently realized that Pauli's exclusion principle would lead to different statistics for electrons, and so Eq. 9.27 is named after them.

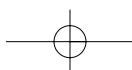
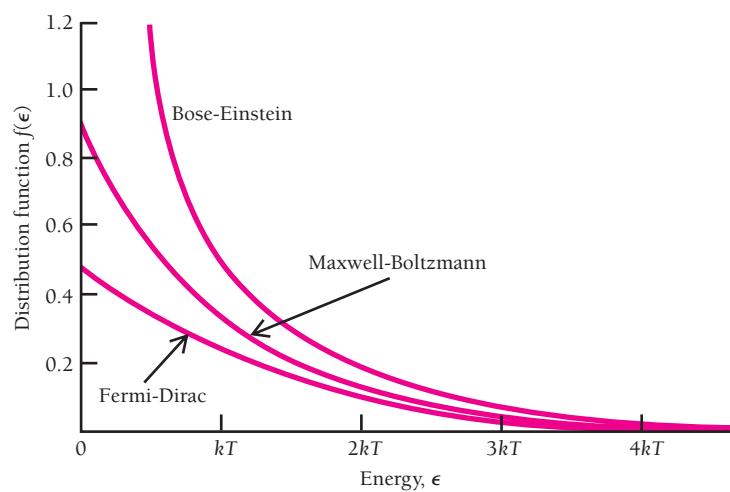


Figure 9.5 A comparison of the three distribution functions for the same value of α . The Bose-Einstein function is always higher than the Maxwell-Boltzmann one, which is a pure exponential, and the Fermi-Dirac function is always lower. The functions give the probability of occupancy of a state of energy ϵ at the absolute temperature T .



The quantity α depends on the properties of the particular system and may be a function of T . Its value is determined by the normalization condition that the sum over all energy states of $n(\epsilon) = g(\epsilon)f(\epsilon)$ be equal to the total number of particles in the system. If the number of particles is not fixed, as in the case of a photon gas, then from the way α is defined in deriving Eqs. (9.26) and (9.27), $\alpha = 0$, $e^\alpha = 1$.

The -1 term in the denominator of Eq. (9.26) expresses the increased likelihood of multiple occupancy of an energy state by bosons compared with the likelihood for distinguishable particles such as molecules. The $+1$ term in the denominator of Eq. (9.27) is a consequence of the uncertainty principle: No matter what the values of α , ϵ , and T , $f(\epsilon)$ can never exceed 1. In both cases, when $\epsilon \gg kT$ the functions $f(\epsilon)$ approach that of Maxwell-Boltzmann statistics, Eq. (9.2). Figure 9.5 is a comparison of the three distribution functions. Clearly $f_{BE}(\epsilon)$ for bosons is always greater at a given ratio of ϵ/kT than it is for molecules, and $f_{FD}(\epsilon)$ for fermions is always smaller.

From Eq. (9.27) we see that $f_{FD}(\epsilon) = \frac{1}{e^{(\epsilon-\epsilon_F)/kT} + 1}$ for an energy of

$$\text{Fermi energy} \quad \epsilon_F = -\alpha kT \quad (9.28)$$

This energy, called the **Fermi energy**, is a very important quantity in a system of fermions, such as the electron gas in a metal. In terms of ϵ_F the Fermi-Dirac distribution function becomes

$$\text{Fermi-Dirac} \quad f_{FD}(\epsilon) = \frac{1}{e^{(\epsilon-\epsilon_F)/kT} + 1} \quad (9.29)$$

To appreciate the significance of the Fermi energy, let us consider a system of fermions at $T = 0$ and investigate the occupancy of states whose energies are less than ϵ_F and greater than ϵ_F . What we find is this:

$$T = 0, \epsilon < \epsilon_F: \quad f_{FD}(\epsilon) = \frac{1}{e^{(\epsilon-\epsilon_F)/kT} + 1} = \frac{1}{e^{-\infty} + 1} = \frac{1}{0 + 1} = 1$$

$$T = 0, \epsilon > \epsilon_F: \quad f_{FD}(\epsilon) = \frac{1}{e^{(\epsilon-\epsilon_F)/kT} + 1} = \frac{1}{e^{\infty} + 1} = 0$$

Bose-Einstein Condensate

Under ordinary conditions, the wave packets that correspond to individual atoms in a gas of atoms are sufficiently small in size relative to their average spacing for the atoms to move independently and be distinguishable. If the temperature of the gas is reduced, the wave packets grow larger as the atoms lose momentum, in accord with the uncertainty principle. When the gas becomes very cold, the dimensions of the wave packets exceed the average atomic spacing so that the wave packets overlap. If the atoms are bosons, the eventual result is that all the atoms fall into the lowest possible energy state and their separate wave packets merge into a single wave packet. The atoms in such a **Bose-Einstein condensate** are barely moving, are indistinguishable, and form one entity—a superatom.

Although such condensates were first visualized by Einstein in 1924, not until 1995 was one actually created. The problem was to achieve a cold enough gas without it becoming a liquid or solid first. This was accomplished by Eric Cornell, Carl Wieman, and their coworkers in Colorado using a gas of rubidium atoms. The atoms were first cooled and trapped by six intersecting beams of laser light. The frequency of the light was adjusted so that the atoms moving against one of the beams would “see” light whose frequency was doppler-shifted to that of one of rubidium’s absorption lines. Thus the atoms would only absorb photons coming toward them, which would slow the atoms and thereby cool the assembly as well as pushing the atoms together and away from the warm walls of the chamber. To get the assembly still colder, the lasers were turned off and a magnetic field held the slower atoms together while allowing the faster ones to escape. (Such evaporative cooling is familiar in everyday life when the faster molecules of a liquid, for instance perspiration, leave its surface and so reduce the average energy of the remaining molecules.) Finally, when the temperature was down to under 10^{-7} K—a tenth of a millionth of a degree above absolute zero—about 2000 rubidium atoms came together in a Bose-Einstein condensate $10 \mu\text{m}$ long that lasted for 10 s.

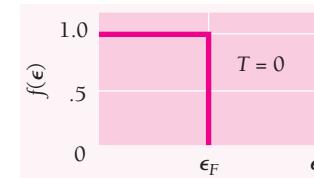
Soon after this achievement other groups succeeded in creating Bose-Einstein condensates in lithium and sodium. One condensate in sodium contained about 5 million atoms, was shaped like a pencil $8 \mu\text{m}$ across and $150 \mu\text{m}$ long, and lasted for 20 s. Still larger condensates were later produced, including one that consisted of 10^8 hydrogen atoms. It proved possible to extract from condensates beams of atoms whose behavior confirmed that they were coherent, with all the atomic wave functions in phase just like the light waves in the coherent beam from a laser. Bose-Einstein condensates are extremely interesting from a number of points of view both fundamental and applied—for example, for possible use in ultrasensitive measurements of various kinds.

Thus at absolute zero all energy states up to ϵ_F are occupied, and none above ϵ_F (Fig. 9.6a). If a system contains N fermions, we can calculate its Fermi energy ϵ_F by filling up its energy states with the N particles in order of increasing energy starting from $\epsilon = 0$. The highest state to be occupied will then have the energy $\epsilon = \epsilon_F$. This calculation will be made for the electrons in a metal in Sec. 9.9.

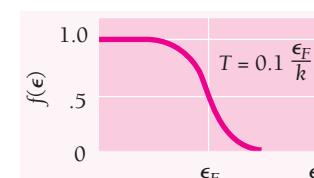
As the temperature is increased above $T = 0$ but with kT still smaller than ϵ_F , fermions will leave states just below ϵ_F to move into states just above it, as in Fig. 9.6b. At higher temperatures, fermions from even the lowest state will begin to be excited to higher ones, so $f_{FD}(0)$ will drop below 1. In these circumstances $f_{FD}(\epsilon)$ will assume a shape like that in Fig. 9.6c, which corresponds to the lowest curve in Fig. 9.5.

The properties of the three distribution functions are summarized in Table 9.1. It is worth recalling that to find the *actual number* $n(\epsilon)$ of particles with an energy ϵ , the functions $f(\epsilon)$ must be multiplied by the number of states $g(\epsilon)$ with this energy:

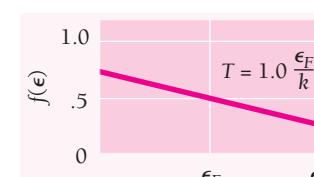
$$n(\epsilon) = g(\epsilon)f(\epsilon) \quad (9.1)$$



(a)

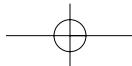


(b)



(c)

Figure 9.6 Distribution function for fermions at three different temperatures. (a) At $T = 0$, all the energy states up to the Fermi energy ϵ_F are occupied. (b) At a low temperature, some fermions will leave states just below ϵ_F and move into states just above ϵ_F . (c) At a higher temperature, fermions from any state below ϵ_F may move into states above ϵ_F .

**Table 9.1** The Three Statistical Distribution Functions

	Maxwell-Boltzmann	Bose-Einstein	Fermi-Dirac
Applies to systems of	Identical, distinguishable particles	Identical, indistinguishable particles that do not obey exclusion principle	Identical, indistinguishable particles that obey exclusion principle
Category of particles	Classical	Bosons	Fermions
Properties of particles	Any spin; particles far enough apart so wave functions do not overlap	Spin 0, 1, 2, . . . ; wave functions are symmetric to interchange of particle labels	Spin $\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots$; wave functions are antisymmetric to interchange of particle labels
Examples	Molecules of a gas	Photons in a cavity; phonons in a solid; liquid helium at low temperatures	Free electrons in a metal; electrons in a star whose atoms have collapsed (white dwarf stars)
Distribution function (number of particles in each state of energy ϵ at the temperature T)	$f_{MB}(\epsilon) = Ae^{-\epsilon/kT}$	$f_{BE}(\epsilon) = \frac{1}{e^{\alpha\epsilon/kT} - 1}$	$f_{FD}(\epsilon) = \frac{1}{e^{(\epsilon-\epsilon_F)/kT} + 1}$
Properties of distribution	No limit to number of particles per state	No limit to number of particles per state; more particles per state than f_{MB} at low energies; approaches f_{MB} at high energies	Never more than 1 particle per state; fewer particles per state than f_{MB} at low energies; approaches f_{MB} at high energies



Paul A. M. Dirac (1902–1984) was born in Bristol, England, and studied electrical engineering there. He then switched his interest to mathematics and finally to physics, obtaining his Ph.D. from Cambridge in 1926. After reading Heisenberg's first paper on quantum mechanics in 1925, Dirac soon devised a more general theory and the next year formulated Pauli's exclusion principle in quantum-mechanical terms. He investigated the statistical behavior of particles that obey the Pauli principle, such as electrons, which Fermi had done independently a little earlier, and the result is called Fermi-Dirac statistics in honor of both. In 1928 Dirac joined special relativity to quantum theory to give a theory of the electron that not only permitted its spin and magnetic moment to be calculated but also predicted the existence of positively charged electrons, or positrons, which were discovered by Carl Anderson in the United States in 1932.

In an attempt to explain why charge is quantized, Dirac in 1931 found it necessary to postulate the existence of **magnetic monopoles**, isolated N or S magnetic poles. More recent theories show that magnetic monopoles should have been created in profusion just after the Big Bang that marked the beginning of the universe; the predicted monopole mass is $\sim 10^{16} \text{ GeV}/c^2$ ($\sim 10^{-8} \text{ g}$!). As Dirac said in 1981, "From the theoretical point of view one would think that monopoles should exist, because of the prettiness of the mathematics. Many attempts to find them have been made, but all have been unsuccessful. One should conclude that pretty mathematics by itself is not an adequate reason for nature to have made use of a theory."

In 1932 Dirac became Lucasian Professor of Mathematics at Cambridge, the post Newton had held two and a half centuries earlier, and in 1933 shared the Nobel Prize in physics with Schrödinger. He remained active in physics for the rest of his life, after 1969 in the warmer climate of Florida, but as is often the case in science he will be remembered for the brilliant achievements of his youth.

