

SEMICONDUCTOR PHYSICS

TOPIC 5: STATISTICAL MECHANICS

PRIMARY READING ASSIGNMENT

"Statistical Mechanics", from *Solid State and Semiconductor Physics*, J.P. McKelvey

SUPPLEMENTARY REFERENCE

"Introduction and Review" from *Statistical Thermodynamics*, D.A. McQuarrie

LECTURE PROGRAM

1. Statistical Mechanics
 - a. Definition
 - b. Six-dimensional phase space
 - c. Basic postulate
2. Distribution Function and Density of States
 - a. Distribution function, $f(\varepsilon)$
 - b. Density of states, $g(\varepsilon)d\varepsilon$
 - c. $N(\varepsilon)d\varepsilon$, distribution of particle density with energy
 - d. $\langle \alpha \rangle$, given $\alpha(\varepsilon)$; $\langle \alpha \rangle$, given $\alpha(\varepsilon)$, $\varepsilon = (1/2m)\sum [p_i^2 + V(q_i)]$
 - e. Schrödinger's equation for free particles; stationary eigenfunction; time-dependent wave function and $\omega-\varepsilon$ relationship; $\langle p \rangle$
 - f. Propagation vector, \mathbf{k} ; relationship between \mathbf{k} , k_x , k_y , k_z and ε ;
 - g. Energy states, $\varepsilon_{n_x n_y n_z}$
 - h. Degeneracy of energy level ε
 - i. Volume of momentum space corresponding to a single quantum state
 - j. Surface in momentum space, surfaces corresponding to constant energy
 - k. Density of states, $g(\varepsilon)d\varepsilon$
3. Maxwell-Boltzmann Distribution
 - a. Conditions
 - b. Equilibrium state of the system
 - c. $Q(N_1, N_2, \dots, N_n)$
 - d. $Q(N_1, N_2, \dots, N_n)$ when degeneracy is included
 - e. Method of Lagrangean multipliers
 - f. Maxwell-Boltzmann distribution
 - g. Connection between Q and entropy, S
4. Maxwell-Boltzmann Statistics of an Ideal Gas
 - a. Boltzmann distribution, $f(e)$; $N(v)$ – for several temperatures
 - b. Equation of state for a Boltzmann gas
5. Fermi-Dirac Statistics
 - a. Conditions
 - b. Fermi-Dirac distribution
 - c. $g(\varepsilon)d\varepsilon$ for a gas of independent particles, such as free electrons
 - d. Temperature variation of the Fermi energy
 - e. Fermi distribution function for different temperatures
 - f. Fermi sphere
 - g. Fermi distribution for $(\varepsilon - \varepsilon_f) \gg kT$
6. Bose-Einstein Distribution
 - a. Conditions
 - b. Bose-Einstein distribution

OUTLINE OF STATISTICAL MECHANICS

5.1 INTRODUCTION

The objective of statistical mechanics is to treat the behavior of a very large assembly of identical particles or systems in a statistical or probabilistic fashion, deriving the most probable values of the properties of the *ensemble* without inquiring in detail what the values of these properties are for any particular particle at any given time.

We shall restrict the present treatment to an assembly of identical systems, which we assume to be independent of one another except that they may interact with each other, or with their external surroundings, only through instantaneous processes which conserve energy and momentum. Such systems might, for example, represent the particles of a free-particle ideal monatomic gas, which could interact with one another or with their surroundings by instantaneous collisions, although systems of other types might equally well be envisioned. External force fields (electric, magnetic, or gravitational) might also exert their influences upon the systems. Under circumstances such as these, we may treat each system independently of the others, and we may describe the behavior of any system in terms of its coordinates in a six-dimensional phase space whose coordinates are (x, y, z, p_x, p_y, p_z) . These coordinates are assumed to be independent and orthogonal, and a complete specification of the state of motion of any particle belonging to the ensemble is given by assigning values to these six-phase space coordinates; its subsequent behavior can then, in principle, be described on the basis of this knowledge.

The basic postulate of statistical mechanics is that the *a priori* probability for a system to be in any given quantum state is the same for all quantum states of the system. This means, for example, that the *a priori* probability that a free particle will be in a region $(\Delta x, \Delta y, \Delta z)$ about (x, y, z) is the same for all values of x, y, z and the probability that its momentum will be in a region $(\Delta p_x, \Delta p_y, \Delta p_z)$ about (p_x, p_y, p_z) is the same for all values of p_x, p_y , and p_z . It should be emphasized that this postulate refers only to the *a priori* probabilities, that is, to the probabilities which prevail in the absence of any dynamical restrictions. The *a priori* probabilities in any given ensemble of systems will generally be modified by external constraints imposed upon the system, such as requirements that the total number and total energy of all the systems belonging to the ensemble remain constant.

We shall first treat the particles belonging to the ensemble as if they were classical "billiard balls," regarding them as distinguishable one from another and allowing any number to occupy a single quantum state. Then, subsequently, we shall investigate what effect the Pauli exclusion principle has, and what the effect of the *indistinguishability* of the elementary particles which we are discussing is. We shall draw freely

upon the results of the preceding chapter, and even in the case of the statistical treatment of classical particles, in the interest of consistency and uniformity, we shall use the quantum framework of energy levels and quantum states, although it is not absolutely necessary to do so.

5.2

THE DISTRIBUTION FUNCTION AND THE DENSITY OF STATES

In order to compute the average properties of an ensemble of particles, it is generally necessary to know how those particles are, on the average, distributed in energy. If we denote by $f(\varepsilon)$ the average number of particles of the system that occupy a single quantum state of energy ε , and if we let $g(\varepsilon)d\varepsilon$ be the number of quantum states of the system whose energy is in a range $d\varepsilon$ about ε , then the number of particles of the system whose energy is in the range $d\varepsilon$ about ε is given by $N(\varepsilon)d\varepsilon$, where

$$N(\varepsilon) d\varepsilon = f(\varepsilon)g(\varepsilon) d\varepsilon. \quad (5.2-1)$$

The quantity $f(\varepsilon)$ as defined above is referred to as the *distribution function* of the system, and depends upon the probabilities associated with the distribution of particles of the system among the available quantum states. The quantity $g(\varepsilon)$, which depends only upon how the quantum states themselves are situated in energy, is called the *density of states*.

If both these quantities are known, then the average value of any quantity α which can be expressed as a function of the total energy ε can be evaluated as

$$\begin{aligned} \langle \alpha \rangle &= \frac{\int \alpha(\varepsilon)N(\varepsilon) d\varepsilon}{\int N(\varepsilon) d\varepsilon} \\ &= \frac{1}{N} \int \alpha(\varepsilon)f(\varepsilon)g(\varepsilon) d\varepsilon, \end{aligned} \quad (5.2-2)$$

the integral of $N(\varepsilon)d\varepsilon$ taken over all possible values of ε giving simply N , the total number of particles in the system. If the quantity α is a function of the system coordinates q_i ($= x, y, z$ for $i = 1, 2, 3$, respectively) and the momenta p_i ($= p_x, p_y, p_z$, in like fashion) which cannot be expressed as a function of the total energy ε alone, then ε may be expressed as a function of the coordinates and momenta as the sum of kinetic and potential energies,

$$\varepsilon = \frac{1}{2m} \sum_i p_i^2 + V(q_i), \quad (5.2-3)$$

and the average value of α evaluated as an integral over the phase space coordinates as

statistical treatment, we shall use though it is not

NSITY

it is generally used in energy. occupy a single state of the particles of the e

(5.2-1)

nction of the n of particles which depends is called the

ity quantity α ted as

(5.2-2)

N , the total e system consists of p_x, p_y, p_z , in ergy ϵ alone, s the sum of

(5.2-3)

ordinates as

$$\langle \alpha \rangle = \frac{\iint \alpha(p_i, q_i) f(p_i, q_i) d\mathbf{p} d\mathbf{q}}{\iint f(p_i, q_i) d\mathbf{p} d\mathbf{q}}. \quad (5.2-4)$$

Here $f(p_i, q_i)$ represents the distribution function $f(\epsilon)$ with ϵ expressed in terms of p_i and q_i by (5.2-3) and the quantities $d\mathbf{p}$ and $d\mathbf{q}$ represent volume elements in momentum and coordinate space expressed in any form (Cartesian, cylindrical, spherical, etc.,) which may be convenient for integration. Equation (5.2-4) contains no density of states factor $g(p_i, q_i)$ because, as we shall soon see, for systems of the type we are considering, the density of quantum states in phase space is constant and will cancel in the numerator and denominator of (5.2-4). We shall often refer to (5.2-2) or (5.2-4) when computing averages.

Since the density of states factor $g(\epsilon)$ as defined in connection with (5.2-1) depends only upon the distribution of quantum states of the system in energy, it should be possible to calculate $g(\epsilon)$ from the solution of Schrödinger's equation. We may do this for free particles as defined in Section 5.1 by assuming that our system is confined within a rigid container whose dimensions are x_0, y_0, z_0 in the x -, y -, and z -directions, respectively. This container is, in effect, an infinitely deep potential well for the particles inside, the potential being zero for each particle within the container, provided that interactions between particles are confined to instantaneous collisions. In this instance, inside the container, Schrödinger's equation, according to (4.8-5) can be written

$$\nabla^2 \psi + k^2 \psi(x, y, z) = 0, \quad (5.2-5)$$

where

$$k^2 = 2me/\hbar^2. \quad (5.2-6)$$

Outside the container, from the results of Sections 4.10 and 4.11, we may conclude that $\psi = 0$. From (4.9-3) we may expect that a valid solution to (5.2-5) would be a plane wave of the form

$$\psi(x, y, z) = A e^{ik_x x} e^{ik_y y} e^{ik_z z} = A e^{i(k_x x + k_y y + k_z z)} = A e^{i(\mathbf{k} \cdot \mathbf{r})}. \quad (5.2-7)$$

It is easy to verify that this solution does actually satisfy (5.2-5). Equation (5.2-7) may indeed be obtained as a solution of (5.2-5) by the separation of variables technique used in connection with the hydrogen atom, assuming initially a product solution of the form $\psi = X(x) Y(y) Z(z)$. In order that (5.2-7) be a solution of (5.2-5), however, the quantities k_x, k_y , and k_z in (5.2-7) must be related in such a way that

$$k_x^2 + k_y^2 + k_z^2 = k^2 = \text{const.} \quad (5.2-8)$$

and these quantities may thus be considered to behave like the components of a vector \mathbf{k} , which is referred to as the propagation vector. From a consideration of the physical character of the time-dependent wave function which is related to (5.2-7) by (4.8-7),

$$\Psi(x, y, z, t) = A e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} \quad (5.2-9)$$

$$(\omega = \epsilon/\hbar),$$

it should be clear that the direction of \mathbf{k} is the same as the direction in which the wave front advances. The actual physical explanation of why this is so is assigned as an exercise for the reader. It can be shown by the methods developed in the last chapter that the expectation value of the vector momentum \mathbf{p} is given by

$$\langle \mathbf{p} \rangle = \hbar \mathbf{k}, \quad (5.2-10)$$

and that \mathbf{p} , (hence \mathbf{k}), is a constant of the motion.

For boundary conditions, we shall demand that the wave function on any one face of the container be equal to the wave function on the opposite face. These *periodic* or cyclic boundary conditions have been discussed previously in connection with lattice vibrations (Section 3.3). In a one-dimensional geometry these boundary conditions represent a system which is topologically equivalent to a ring. In three dimensions the periodic boundary conditions divide all space up into exactly similar regions of dimensions (x_0, y_0, z_0) in each of which the wave function is the same, and any one of these regions may be used to represent the interior of the potential well corresponding to the container. Proceeding in this manner and assuming that edges of the container extend along the x -axis from 0 to x_0 , along the y -direction from 0 to y_0 , and along the z -direction from 0 to z_0 , the boundary conditions become

$$\psi(0, y, z) = \psi(x_0, y, z)$$

$$\psi(x, 0, z) = \psi(x, y_0, z) \quad (5.2-11)$$

$$\psi(x, y, 0) = \psi(x, y, z_0),$$

which require that in (5.2-7) we take

$$k_x = 2\pi n_x / x_0 \quad (n_x = 0, \pm 1, \pm 2, \dots)$$

$$k_y = 2\pi n_y / y_0 \quad (n_y = 0, \pm 1, \pm 2, \dots) \quad (5.2-12)$$

$$k_z = 2\pi n_z / z_0 \quad (n_z = 0, \pm 1, \pm 2, \dots),$$

(5.2-7) then reducing to

$$\psi(x, y, z) = A e^{2\pi i \left(\frac{n_x x}{x_0} + \frac{n_y y}{y_0} + \frac{n_z z}{z_0} \right)} \quad (5.2-13)$$

According to (5.2-8) and (5.2-6) this means that only a certain discrete set of energy values, given by

$$\epsilon_{n_x n_y n_z} = \frac{2\pi^2 \hbar^2}{m} \left[\frac{n_x^2}{x_0^2} + \frac{n_y^2}{y_0^2} + \frac{n_z^2}{z_0^2} \right] \quad (5.2-14)$$

are allowed. From (5.2-10), the allowed values of (k_x, k_y, k_z) can be expressed as allowed

which the wave
assigned as an
ie last chapter

(5.2-10)

values of (p_x, p_y, p_z) in the form

$$p_x = hn_x/x_0$$

$$p_y = hn_y/y_0 \quad (5.2-15)$$

$$p_z = hn_z/z_0.$$

If we were to plot all the allowed values of the momentum, corresponding to all possible integer values for (n_x, n_y, n_z) in (5.2-15), as points in an orthogonal momentum space with coordinates (p_x, p_y, p_z) , we would obtain a simple orthogonal lattice of points representing allowed momentum values, with unit cell dimensions $(h/x_0, h/y_0, h/z_0)$ corresponding to unit changes in (n_x, n_y, n_z) in (5.2-15). The volume of momentum space corresponding to a single quantum state of the system is simply the unit cell volume

$$V_p = \frac{h^3}{x_0 y_0 z_0} = h^3/V, \quad (5.2-16)$$

where V is the physical volume of the container. From this (and from the plot of allowed momentum points) we see that the density with which the allowed quantum states are distributed in momentum space is uniform over all momentum space (which was mentioned previously in connection with Equation (5.2-4)). In (5.2-16) we have taken no account of spin; for particles of spin $\frac{1}{2}$ like electrons, there will be two allowed momentum states for each lattice point in momentum space, corresponding to the two possible spin orientations of the particle. Taking this into account, (5.2-16) should be rewritten in the form

$$V_p = \frac{1}{2} h^3/V. \quad (5.2-17)$$

For large values of x_0 , y_0 , and z_0 , such as would be associated with a container of macroscopic dimensions, the volume V_p of (5.2-17) becomes very small, and the momentum states (5.12-15) and energy states (5.2-14) become crowded together very closely compared to energy and momentum intervals which are at all appreciable macroscopically. For a system of reasonable dimensions, then, we shall be justified in regarding the allowed energy and momentum values as substantially *continuously* distributed in momentum space.

Now let us consider a surface in the momentum space (p_x, p_y, p_z) , all points of which are at constant energy ϵ . From (5.2-6), (5.2-8), and (5.2-10), the equation of this surface must be that of a sphere of radius $p = \sqrt{2m\epsilon}$, or

$$p^2 = p_x^2 + p_y^2 + p_z^2 = 2m\epsilon. \quad (5.2-18)$$

This surface is illustrated in Figure 5.1. A similar sphere whose surface represents all points of energy $\epsilon + d\epsilon$ is shown in that figure. The spherical shell between these two spheres represents that region of momentum space corresponding to energies in the range ϵ to $\epsilon + d\epsilon$. The volume of momentum space contained within this shell is just

$$dV_p = 4\pi p^2 dp. \quad (5.2-19)$$

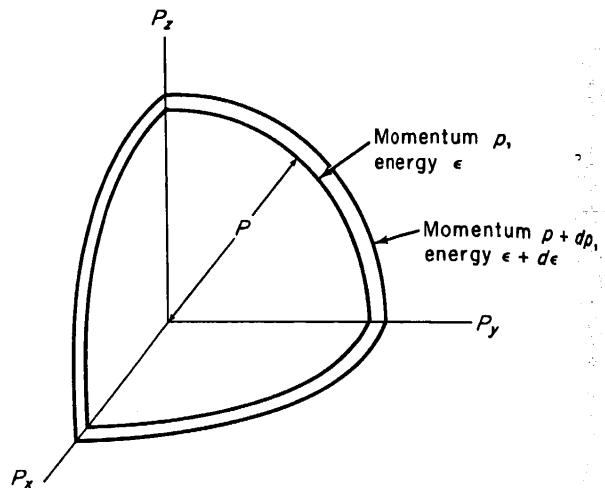
However, from (5.2-18)

$$p \, dp = m \, d\epsilon, \quad (5.2-20)$$

whereby

$$dV_p = 4\pi p \cdot m \, d\epsilon = 4\pi m \sqrt{2m\epsilon} \, d\epsilon. \quad (5.2-21)$$

FIGURE 5.1. Spherical surfaces corresponding to constant energies ϵ and $\epsilon + d\epsilon$ plotted in the momentum space (p_x, p_y, p_z) of a particle.



The number of quantum states to be found in this volume of momentum space is determined simply by dividing this result by the volume of momentum space associated with a single quantum state as given by (5.2-17). The result,

$$g(\epsilon) \, d\epsilon = \frac{8\sqrt{2\pi}V}{h^3} m^{3/2} \sqrt{\epsilon} \, d\epsilon, \quad (5.2-22)$$

is by definition the density of states factor referred to in (5.2-1). This result, of course, refers only to free particles subject to instantaneous collision interactions only. For other systems, particularly where long range mutual interactions between particles are involved, the density of states factor would be much more complex. Also, the result (5.2-22) pertains, strictly speaking, only to a container in the shape of a rectangular solid. However, since only the volume of the container appears in the final result, it is intuitively clear that the same density of states would be obtained for a container of like volume independent of its shape. This can indeed be shown to be true.

5.3

THE MAXWELL-BOLTZMANN DISTRIBUTION

If there are no restrictions on the amount of energy or momentum a particle of the system may possess, then the probability associated with all quantum states is the same, and the average number of particles per quantum state will be independent of energy. This means that the distribution function $f(\epsilon)$ will be constant. This rather simple situation is not very important physically, since a system which is thermally

(5.2-20)

isolated from its surroundings must obey the restriction that the sum of the energies of all the particles of the system be constant, and it is systems of this type with which we shall have to deal. In this more realistic case the proportion of particles occupying states of extremely high energy is reduced and the distribution function is no longer constant with respect to energy.

(5.2-21)

To find out just what the distribution function is under these conditions, we shall proceed initially along classical lines, imagining the particles of the system to be *identifiable* objects such as numbered billiard balls. We shall retain the framework of quantum states and energy levels, although, ignoring the Pauli exclusion principle, we shall permit any number of particles to occupy a given quantum state of the system. We shall suppose that we are dealing with an isolated system of N distinguishable particles, with constant total energy U , which may be distributed among n energy levels $\epsilon_1, \epsilon_2, \epsilon_3, \dots, \epsilon_i, \dots, \epsilon_n$. Statistically, the ensemble will have that energy distribution which corresponds to some chance distribution of particles among levels, the number in each level being given by $N_1, N_2, N_3, \dots, N_i, \dots, N_n$. Of all the chance distributions of N particles among n energy levels, some will occur with relatively high probability on a purely statistical basis, and others will be quite unlikely. This situation is analogous to the simultaneous tossing of two coins, in which the distribution (1 head, 1 tail) is more probable than the distribution (2 heads, 0 tails) or the distribution (0 heads, 2 tails). The distribution which has the maximum probability of occurrence is that distribution of particles among levels which *can be realized in a maximum number of statistically independent ways*. The equilibrium state of the system will then be assumed to correspond closely to this statistical distribution of particles among levels which is of maximum probability under the conditions of the problem.

(5.2-22)

The calculation of the number of ways in which identifiable particles can be distributed among energy levels of a system is equivalent to the calculation of the number of ways in which numbered objects can be distributed among a set of numbered containers. We must thus consider the problem of determining the number of ways of distributing N identifiable objects among n containers such that there are N_1 in the first, N_2 in the second, \dots , N_i in the i th, etc. This number will be proportional to the probability with which a distribution where N_1, N_2, \dots, N_n objects are in containers 1, 2, 3, \dots , n , will occur.

To begin, let us assume that we have only two containers in which the objects may be placed, as shown in Figure 5.2. In general there will be N_1 objects in container 1 and N_2 in container 2, with

$$N_1 + N_2 = N = \text{const.} \quad (5.3-1)$$

Let us denote by $Q(N_1, N_2)$ the number of statistically independent ways of arriving at the distribution (N_1, N_2) objects in containers (1,2). Now $Q(0, N_2)$ is certainly equal to one, because the only way of achieving this distribution is to put all N objects into the second box, as shown in the figure. Likewise, $Q(1, N_2) = N$, because to realize this distribution, we may put object 1 in the first container and the rest in the second, object 2 in the first container, and the rest in the second, and so on through the N objects. There are N ways of choosing the object to be put in the first container and thus N ways of arriving at the distribution in question. If there are two objects in the first box, then $Q(2, N_2) = N(N - 1)/2!$. This result follows because there are N ways of choosing the first object which goes into the first container, but only $N - 1$ ways of choosing the second from the remaining objects; also, the two distributions

where the object numbered α was chosen as the first object and the object numbered β as the second object to be put into the first container and where β was chosen as the first and α as the second are really identical, the same two objects ending up finally in the first box; the factor $N(N - 1)$ must thus be divided by two. For the case where there are three objects in the first box, there are N ways of choosing the first,

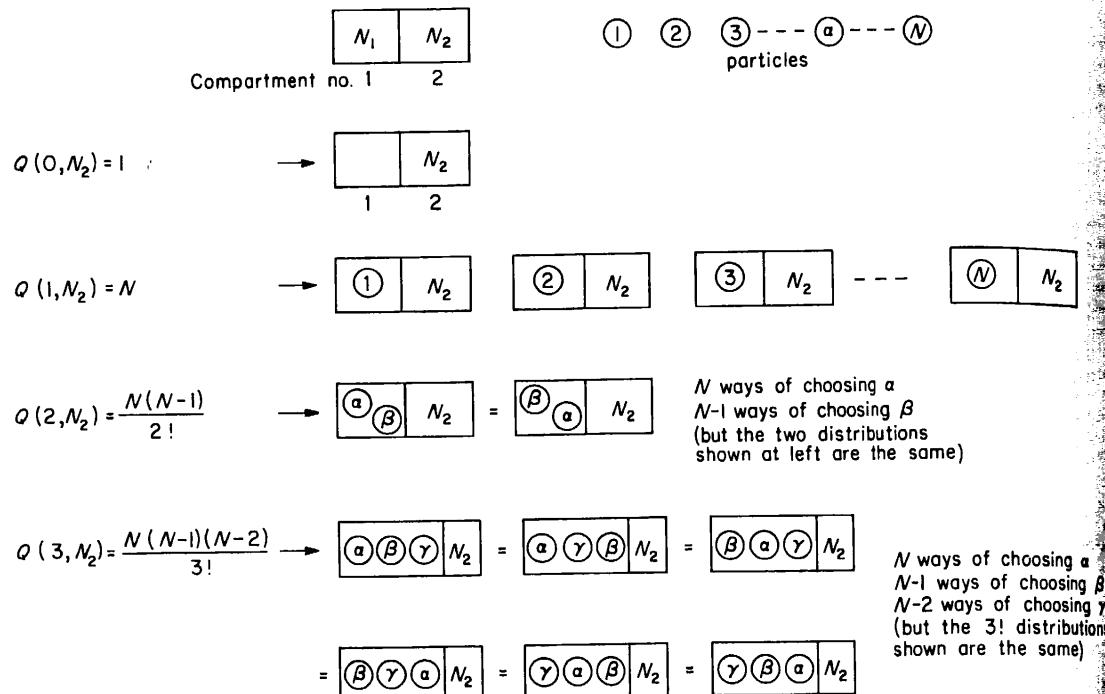


FIGURE 5.2. The possible ways of distributing N identifiable particles among two containers.

$N - 1$ ways of choosing the second and $N - 2$ ways of choosing the third, while there are $3!$ ways of permuting numbered objects α , β , and γ among themselves as having been chosen first, second or third, whereby $Q(3, N_2) = N(N - 1)(N - 2)/3!$. By an obvious extension of this process, and recalling (5.3-1), it is readily established that

$$Q(N_1, N_2) = \frac{N(N-1)(N-2) \cdots (N-N_1+1)}{N_1!} = \frac{N!}{N_1!(N-N_1)!} = \frac{N!}{N_1!N_2!}. \quad (5.3-2)$$

Suppose now that the second container is divided into two subcompartments containing v_1 and v_2 objects, where, of course,

$$v_1 + v_2 = N_2 = N - N_1, \quad (5.3-3)$$

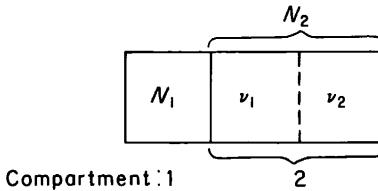
as shown in Figure 5.3. The number of independent ways of realizing the distribution (v_1, v_2) among the sub-compartments of the second box is, from (5.3-3)

$$Q(v_1, v_2) = \frac{N_2!}{v_1! v_2!}. \quad (5.3-4)$$

object numbered was chosen as objects ending up . For the case using the first,

But now we may, if we wish, regard the system as having *three* distinct containers, and the total number of ways of arranging the objects so that there are N_1 in the first container, v_1 in the second and v_2 in the third [which we shall call $Q(N_1, v_1, v_2)$] will be just the product of $Q(N_1, N_2)$ and $Q(v_1, v_2)$, since for *each* arrangement (N_1, N_2)

FIGURE 5.3. The subdivision of the second container into two subcompartments, leading to a system which can be thought of as actually having been divided into *three* distinct containers.



there will be $Q(v_1, v_2)$ ways of arranging particles among the subcompartments, and there are $Q(N_1, N_2)$ basic ways of arriving at a distribution (N_1, N_2) for the two original containers. We shall find thus, that

$$Q(N_1, v_1, v_2) = Q(N_1, N_2)Q(v_1, v_2) = \frac{N!}{N_1! N_2!} \frac{N_2!}{v_1! v_2!} = \frac{N!}{N_1! v_1! v_2!}. \quad (5.3-5)$$

We may now simply relabel the three containers as boxes 1, 2, and 3, and call the number of objects in each N_1 , N_2 , and N_3 , rather than N_1 , v_1 , and v_2 , in which case (5.3-5) may be written

$$Q(N_1, N_2, N_3) = \frac{N!}{N_1! N_2! N_3!}. \quad (5.3-6)$$

Again, one may imagine the *third* compartment to be subdivided as before, and obtain in the same manner an expression for the number of ways of arranging N objects among four boxes so as to obtain a distribution (N_1, N_2, N_3, N_4) . By repeated application of this process, the result may be extended indefinitely by induction to cover the case where there may be n boxes. It is clear from the form of (5.3-3) and (5.3-6) that one will find

$$Q(N_1, N_2, N_3, \dots, N_n) = \frac{N!}{N_1! N_2! N_3! \dots N_n!} = \frac{N!}{\prod_{i=1}^n N_i!}, \quad (5.3-7)$$

where the \prod notation is used to indicate an extended product in the same way that the more familiar \sum notation is used to express a summation.

The actual *probability* associated with a given distribution of N objects among n boxes is the number of independent ways of arranging objects among boxes $Q(N_1, \dots, N_n)$ which lead to that particular distribution, divided by the total number of ways of arranging N objects among n boxes *without regard for what distribution of objects among boxes results*. It is easily shown that this latter factor is simply n^N . The probability associated with a given distribution will then be $Q(N_1, \dots, N_n)/n^N$. We shall find it more convenient, however, to deal exclusively with the quantities $Q(N_1, \dots, N_n)$, which are *proportional* to the actual probabilities, since we shall only be interested in finding the values N_1, \dots, N_n which render the probability a maximum, and if $Q(N_1, \dots, N_n)$ is a maximum, then so is the associated probability.

Now let us identify the i th container with the i th energy level of the system and the number of objects N_i in that container with the number of particles belonging to that level. We must, however, allow for the fact that the energy levels may be degenerate, and we shall assume there are g_i independent quantum states associated with the i th energy level, each of which has the same *a priori* probability of being occupied, according to the fundamental postulate of Section 5.1. Each energy level must then be thought of not as a single container, but as a group of g_i containers, as illustrated in Figure 5.4. For a given energy level, say the i th, containing N_i particles and g_i

FIGURE 5.4. The subdivision of energy levels into separate quantum states according to their degeneracy. The separate quantum states play the role of the containers in the statistical development, each having equal statistical weight.

Degeneracy $g_i = 4$	6	---	2	
Level	1	2	---	n
Population N_i	4	7	---	2
Diagram:				

independent quantum states, there are $g_i^{N_i}$ ways of arranging the particles among the states or containers pertaining to that level, because there are g_i independent choices of where the first particle may be placed, and for each of these there are g_i choices of where the second may be placed, and so on to N_i factors. Each of these arrangements would constitute a separate statistically independent way of distributing particles among the quantum states of the system, since it is the quantum states themselves which are of equal probability, not the energy levels. If the first energy level, ε_1 , were the *only* degenerate level, then there would be $g_1^{N_1}$ times as many independent ways of forming a given distribution of particles among energy levels as is given by (5.3-7), because there are that many ways of permuting particles among separate degenerate states belonging to the first level, and for each of these there still remain the same number of ways of assigning particles to the other levels of the system in such a way that the distribution (N_1, N_2, \dots, N_n) results. If, in addition, the second level were degenerate, there would be $g_2^{N_2}$ possibilities for each one that existed before, the total number of independent ways of achieving a given distribution of particles among levels being now $g_1^{N_1} g_2^{N_2}$ times that given by (5.3-7). It is clear that when the degeneracy of all the levels is included, the result (5.3-7) will have to be multiplied by $g_1^{N_1} g_2^{N_2} g_3^{N_3} \dots g_n^{N_n}$, giving

$$Q(N_1, N_2, \dots, N_n) = \frac{N!}{\prod_{i=1}^n N_i!} \prod_{i=1}^n g_i^{N_i} \quad (5.3-8)$$

when degeneracy is included.

We shall now assume that the actual distribution (N_1, N_2, \dots, N_n) of particles among energy states which is observed in equilibrium is essentially that which can be realized in the *maximum* number of statistically independent ways, in other words, that for which the quantity Q of (5.3-8) is a maximum. Q must then be maximized with respect to the parameters N_1, N_2, \dots, N_n , subject to the restrictions

$$\sum_{i=1}^n N_i = N = \text{const.} \quad (5.3-9)$$

the system and
es belonging to
may be degener-
ated with the
being occupied,
evel must then
s, as illustrated
articles and g_i

and

$$\sum_{i=1}^n \varepsilon_i N_i = U = \text{const.} \quad (5.3-10)$$

This maximization is most easily carried out by the use of a mathematical technique called the method of *Lagrangean multipliers*, and we shall digress briefly to describe this subject. Suppose we are given a function $f(x_1, x_2, \dots, x_n)$ of n variables, and we are asked to find the values of x_1, x_2, \dots, x_n which make f a maximum *under the restriction that some other given function $\phi(x_1, x_2, \dots, x_n)$ remain constant*. For a maximum (or minimum) value of f , $df = 0$, and if ϕ remains constant, $d\phi = 0$, so that if f is a maximum or minimum under the stated conditions, then

$$df + \alpha d\phi = 0 \quad (5.3-11)$$

no matter what value the arbitrary undetermined multiplier α may have. But (5.3-11) may be written

$$\left(\frac{\partial f}{\partial x_1} + \alpha \frac{\partial \phi}{\partial x_1} \right) dx_1 + \left(\frac{\partial f}{\partial x_2} + \alpha \frac{\partial \phi}{\partial x_2} \right) dx_2 + \dots + \left(\frac{\partial f}{\partial x_n} + \alpha \frac{\partial \phi}{\partial x_n} \right) dx_n = 0. \quad (5.3-12)$$

According to this equation, condition (5.3-11) will certainly be fulfilled if for every i

$$\frac{\partial f}{\partial x_i} + \alpha \frac{\partial \phi}{\partial x_i} = 0 \quad (i = 1, 2, \dots, n), \quad (5.3-13)$$

while, in addition, we were given to begin with the fact that

$$\phi(x_1, x_2, \dots, x_n) = \text{const.} \quad (5.3-14)$$

Equations (5.3-13) and (5.3-14), taken together, represent a set of $n + 1$ simultaneous equations which can be solved for the n quantities x_1, x_2, \dots, x_n which maximize (or minimize) f and for the undetermined multiplier α which was introduced in (5.3-11).

If there are *two* auxiliary functions $\phi(x_1, x_2, \dots, x_n)$ and $\psi(x_1, x_2, \dots, x_n)$ which are to remain constant as the function f is maximized, then *two* arbitrary multipliers α and β are introduced and it is required that

$$df + \alpha d\phi + \beta d\psi = 0, \quad (5.3-15)$$

which leads, in the same manner, to a set of $n + 2$ equations

$$\frac{\partial f}{\partial x_i} + \alpha \frac{\partial \phi}{\partial x_i} + \beta \frac{\partial \psi}{\partial x_i} = 0 \quad (i = 1, 2, \dots, n) \quad (5.3-16)$$

$$\phi(x_1, x_2, \dots, x_n) = \phi_0 = \text{const.} \quad (5.3-17)$$

$$\psi(x_1, x_2, \dots, x_n) = \psi_0 = \text{const.} \quad (5.3-18)$$

(5.3-9) for the $n + 2$ unknowns $x_1, x_2, \dots, x_n, \alpha$ and β .

We shall now proceed to apply this method to find the distribution N_1, N_2, \dots, N_n for which (5.3-8) is a maximum. As a matter of mathematical convenience, we shall actually maximize $\ln Q$ rather than Q itself, but since the logarithm is a single-valued monotonic function of all the variables involved, when $\ln Q$ is a maximum, so also is Q . If we take the logarithm of both sides of (5.3-8), we find

$$\ln Q(N_1, N_2, \dots, N_n) = \ln N! + \sum_{i=1}^n N_i \ln g_i - \sum_{i=1}^n \ln N_i!. \quad (5.3-19)$$

We shall assume that our system is so large that for each level $N_i!$ may be approximated by Stirling's approximation, which states that for $x \gg 1$,

$$\ln x! \approx x \ln x - x. \quad (5.3-20)$$

Making use of this approximation, (5.3-19) becomes

$$\ln Q = \ln N! + \sum_i N_i \ln g_i - \sum_i N_i \ln N_i + \sum_i N_i, \quad (5.3-21)$$

which must be maximized under the restrictions

$$\phi(N_1, N_2, \dots, N_n) = \sum_i N_i = N \quad (5.3-22)$$

$$\psi(N_1, N_2, \dots, N_n) = \sum_i \varepsilon_i N_i = U. \quad (5.3-23)$$

From (5.3-16), this requires that

$$\begin{aligned} \frac{\partial(\ln Q)}{\partial N_j} + \alpha \frac{\partial \phi}{\partial N_j} + \beta \frac{\partial \psi}{\partial N_j} &= \frac{\partial}{\partial N_j} \left[\sum_i N_i \ln g_i - \sum_i N_i \ln N_i + \sum_i N_i \right] \\ &+ \alpha \frac{\partial}{\partial N_j} \left(\sum_i N_i \right) + \beta \frac{\partial}{\partial N_j} \left(\sum_i \varepsilon_i N_i \right) = 0 \quad (j = 1, 2, \dots, n). \end{aligned} \quad (5.3-24)$$

Working out the derivatives in (5.3-24), noting that the only terms of the summations whose derivatives with respect to N_j are other than zero are those for which $i=j$, (5.3-24) can be reduced to

$$\ln g_j - \ln N_j + \alpha + \beta \varepsilon_j = 0 \quad (j = 1, 2, \dots, n). \quad (5.3-25)$$

Solving for $\ln(N_j/g_j)$ and exponentializing, this can be rewritten in the form

$$\frac{N_j}{g_j} = e^\alpha e^{\beta \varepsilon_j} = f(\varepsilon_j). \quad (5.3-26)$$

Equation (5.3-26) gives the average number of particles per quantum state of the system and thus, by definition, represents the energy distribution function $f(\varepsilon)$. This

n $N_1, N_2, \dots N_n$,
nience, we shall
a single-valued
mum, so also is

particular energy distribution function obtained as (5.3-26) under the classical assumption of identifiable particles and without the use of the Pauli exclusion principle, is called the *Maxwell-Boltzmann* distribution function.

We must now discuss how the constants α and β are related to the physical properties of the system. To begin with, we shall assign to the constant β the value

(5.3-19)

$$\beta = -1/kT, \quad (5.3-27)$$

ay be approxi-

(5.3-20)

where T is the absolute temperature of the system and k is a constant called Boltzmann's constant, and we shall regard this equation, along with (5.3-26) as *defining* what is meant by temperature. We shall see, in due course, that this definition leads to all the familiar characteristics of temperature as related to an ideal gas, and it will then be clear that the definition of temperature *could* have been postponed until the properties of an ideal gas had been worked out from (5.3-26). An identification of β with the value given above could then have been assigned on the basis of a comparison of the results so obtained and the familiar thermodynamic equation of state for an ideal gas (from which the temperature is more commonly defined). It will be shown that the value of k can be related to the measured ideal gas constant R and Avogadro's number.

Using the value given by (5.3-27) for β , (5.3-26) may now be written

(5.3-22)

$$N_j = g_j e^\alpha e^{-\epsilon_j/kT}. \quad (5.3-28)$$

(5.3-23)

The value of the constant α may now be expressed in terms of the total number of particles N , since from (5.3-9) and (5.3-28)

$$N = \sum_j N_j = e^\alpha \sum_j g_j e^{-\epsilon_j/kT}, \quad (5.3-29)$$

whereby

$$e^\alpha = \frac{N}{\sum_j g_j e^{-\epsilon_j/kT}}, \quad (5.3-30)$$

and

$$N_j = g_j e^\alpha e^{-\epsilon_j/kT} = \frac{N g_j e^{-\epsilon_j/kT}}{\sum_j g_j e^{-\epsilon_j/kT}}. \quad (5.3-31)$$

If the energy levels of the system are crowded very closely together, as are the levels of the gas of free particles which was discussed in the previous section, then the quantity g_j in (5.3-28) may be regarded as $g(\epsilon)d\epsilon$ and the quantity N_j in that equation may be regarded as $N(\epsilon)d\epsilon$, as discussed in connection with Equation (5.2-1). The number of particles in an energy range $d\epsilon$ about ϵ in this limit will, according to (5.3-28), be given by

$$N(\epsilon) d\epsilon = e^\alpha e^{-\epsilon/kT} g(\epsilon) d\epsilon = f(\epsilon) g(\epsilon) d\epsilon. \quad (5.3-32)$$

For an ideal gas of free particles, the density of states factor $g(\epsilon)$ will be given by (5.2-22). Note that the form of (5.3-32) is the same as that of (5.2-1) with $f(\epsilon) = e^\alpha e^{-\epsilon/kT}$. Again, as discussed in connection with (5.3-31), we may evaluate the constant α by the condition that the total number of particles in the system shall be a constant N ,

whereby

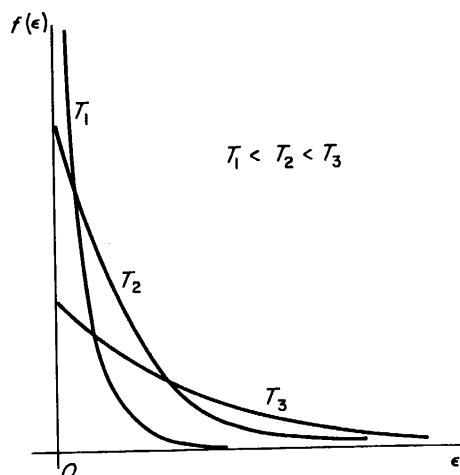
$$N = \int N(\epsilon) d\epsilon = e^\alpha \int g(\epsilon) e^{-\epsilon/kT} d\epsilon, \quad (5.3-33)$$

and

$$e^\alpha = \frac{N}{\int g(\epsilon) e^{-\epsilon/kT} d\epsilon}, \quad (5.3-34)$$

the integrals being taken over all energies available to the particles of the system.

FIGURE 5.5. A schematic representation of the Maxwell-Boltzmann energy distribution function for three different temperatures.



We should note that the quantity Q of (5.3-8) is closely connected with the thermodynamic *entropy* of the system. At equilibrium, the state of the system is such that Q , the number of statistically independent ways of distributing particles among quantum states, is a maximum, as is the entropy of the system. We shall not go into details at this point, but it can indeed be demonstrated¹ that the relation between these two quantities is

$$S = k \ln Q, \quad (5.3-35)$$

where S is the entropy.

5.4 MAXWELL-BOLTZMANN STATISTICS OF AN IDEAL GAS

We shall now discuss the properties of an ideal gas of free particles of the type discussed in Section 4.2. For this system, the density of states factor $g(\epsilon) d\epsilon$ is given by (5.2-22), whereby, according to (5.3-34),

¹ J. E. Mayer and M. Göppert-Mayer, *Statistical Mechanics*, John Wiley and Sons, New York (1940), Chapter 4.

(5.3-33)

(5.3-34)

ie system.

$$e^\alpha = \frac{N}{\frac{8\sqrt{2\pi}Vm^{3/2}}{h^3} \int_0^\infty \sqrt{\varepsilon} e^{-\varepsilon/kT} d\varepsilon}. \quad (5.4-1)$$

The integration is taken between the limits zero and infinity, since the energy of the particles is entirely kinetic, hence positive. The integral in (5.4-1) may be expressed in a simpler form by the substitution

$$x = \varepsilon/kT, \quad (5.4-2)$$

whence (5.4-1) becomes

$$e^\alpha = \frac{N}{\frac{8\sqrt{2\pi}V}{h^2} \left(\frac{mkT}{h^2}\right)^{3/2} \int_0^\infty x^{1/2} e^{-x} dx}. \quad (5.4-3)$$

The integral can now be expressed as a Γ -function,² since $\Gamma(n)$ is defined as

$$\Gamma(n) = \int_0^\infty x^{n-1} e^{-x} dx, \quad (5.4-4)$$

the integral in (5.4-3) thus being equal to $\Gamma(3/2)$, which in turn equals $\sqrt{\pi}/2$. Inserting this value into (5.4-3), we find

$$e^\alpha = \frac{N}{2V} \left(\frac{h^2}{2\pi mkT}\right)^{3/2}, \quad (5.4-5)$$

which makes it possible to write the Maxwell-Boltzmann distribution function for an ideal gas as

$$f(\varepsilon) = e^\alpha e^{-\varepsilon/kT} = \frac{N}{2V} \left(\frac{h^2}{2\pi mkT}\right)^{3/2} e^{-\varepsilon/kT}. \quad (5.4-6)$$

It will be noted that the value of e^α given by (5.4-5) is *temperature dependent*. It should also be emphasized that this particular value for e^α pertains *only* to the free particle density of states function (5.2-22) and that for systems having other density of states functions associated with them the value for e^α will be different from that given by (5.4-5). A plot of the Boltzmann distribution (5.4-6) for several temperatures is shown in Figure 5.5. From (5.4-6) and (5.2-22) it is clear that the actual distribution of particle density with energy is given by

$$N(\varepsilon) d\varepsilon = f(\varepsilon) g(\varepsilon) d\varepsilon = \frac{2\pi N}{(\pi kT)^{3/2}} \sqrt{\varepsilon} e^{-\varepsilon/kT} d\varepsilon. \quad (5.4-7)$$

² See, for instance, I. S. and E. S. Sokolnikoff, *Higher Mathematics for Engineers and Physicists*, McGraw-Hill Book Co., Inc., New York (1941), pp. 273-276.

The total internal energy and specific heat of the gas can be obtained very easily from these results. Since the energy of the particles in the range de about ϵ is simply $\epsilon N(\epsilon) de$, the total internal energy of the gas is

$$U = \int_0^\infty \epsilon N(\epsilon) d\epsilon, \quad (5.4-8)$$

or, using (5.4-7),

$$U = \frac{2\pi N}{(\pi kT)^{3/2}} \int_0^\infty \epsilon^{3/2} e^{-\epsilon/kT} d\epsilon. \quad (5.4-9)$$

This integral can be evaluated in terms of Γ -functions by making the substitution (5.4-2). Working out the integral in this way, noting that $\Gamma(5/2) = (3/2)\Gamma(3/2) = 3\sqrt{\pi}/4$, we find

$$U = \frac{3}{2} NkT, \quad (5.4-10)$$

whence the average internal energy per particle, U/N , is $\frac{3}{2}kT$, an important and familiar result. The heat capacity of the gas at constant volume is the rate of increase of internal energy with respect to temperature, whereby from (5.4-10),

$$C_v = \left(\frac{\partial U}{\partial T} \right)_v = \frac{3}{2} Nk, \quad (5.4-11)$$

independent of temperature. The specific heat c_v is simply the heat capacity per unit volume.

Our objective, eventually, is to derive the equation of state for an ideal Boltzmann gas from the dynamical properties of the particles and from the distribution function. Before doing this, however, we must convert the *energy* distributions (5.4-6) and (5.4-7) into appropriate velocity distributions, and investigate briefly how to use these velocity distribution functions. Table 5.1 gives the value of some of the definite integrals which are frequently encountered when working with the Boltzmann velocity distributions.

In a gas of free particles possessing no internal degrees of freedom, all the energy resides in the kinetic energy of the particles. We may relate this to the velocity by writing

$$\epsilon = \frac{1}{2}mv^2 = \frac{1}{2}m(v_x^2 + v_y^2 + v_z^2) \quad (5.4-12)$$

whence

$$d\epsilon = mv dv, \quad (5.4-13)$$

and the energy distribution (5.4-7) may be written directly as a velocity distribution of the form

$$N(v) dv = 4\pi N \left(\frac{m}{2\pi kT} \right)^{3/2} v^2 e^{-mv^2/2kT} dv. \quad (5.4-14)$$

This function expresses the number of particles of the system whose *speeds* lie in a range dv about v , or the number of particles which lie within a spherical shell of thickness dv and radius v in *velocity* space.

ned very easily
out ϵ is simply

(5.4-8)

$$\int_0^\infty e^{-\alpha x^2} dx = \frac{1}{2} \sqrt{\frac{\pi}{\alpha}} = \frac{1}{2} \sqrt{\frac{2\pi kT}{m}} = \frac{\pi \bar{c}}{4}$$

(5.4-9)

$$\int_0^\infty x e^{-\alpha x^2} dx = \frac{1}{2\alpha} = \frac{kT}{m} = \frac{\pi \bar{c}^2}{8}$$

(5.4-10)

$$\int_0^\infty x^2 e^{-\alpha x^2} dx = \frac{1}{4\alpha} \sqrt{\frac{\pi}{\alpha}} = \frac{\sqrt{\pi}}{4} \left(\frac{2kT}{m}\right)^{3/2} = \frac{\pi^2 \bar{c}^3}{32}$$

$$\int_0^\infty x^3 e^{-\alpha x^2} dx = \frac{1}{2\alpha^2} = \frac{1}{2} \left(\frac{2kT}{m}\right)^2 = \frac{\pi^2 \bar{c}^4}{32}$$

$$\int_0^\infty x^4 e^{-\alpha x^2} dx = \frac{3}{8\alpha^3} \sqrt{\frac{\pi}{\alpha}} = \frac{3\sqrt{\pi}}{8} \left(\frac{2kT}{m}\right)^{5/2} = \frac{3\pi^3 \bar{c}^5}{256}$$

$$\int_0^\infty x^5 e^{-\alpha x^2} dx = \frac{1}{\alpha^3} = \left(\frac{2kT}{m}\right)^3 = \frac{\pi^3 \bar{c}^6}{64}$$

$$\int_0^\infty x^6 e^{-\alpha x^2} dx = \frac{15}{16\alpha^4} \sqrt{\frac{\pi}{\alpha}} = \frac{15\sqrt{\pi}}{16} \left(\frac{2kT}{m}\right)^{7/2} = \frac{15\pi^4 \bar{c}^7}{2^{11}}$$

$$\int_0^\infty x^7 e^{-\alpha x^2} dx = \frac{3}{\alpha^4} = 3 \left(\frac{2kT}{m}\right)^4 = \frac{3\pi^4 \bar{c}^8}{256}$$

$$\int_0^\infty x^8 e^{-\alpha x^2} dx = \frac{105}{32\alpha^5} \sqrt{\frac{\pi}{\alpha}} = \frac{105\sqrt{\pi}}{32} \left(\frac{2kT}{m}\right)^{9/2} = \frac{105\pi^5 \bar{c}^9}{2^{14}}$$

* The second column in the table gives the value of the definite integral shown in the first column. The third column gives the value of the definite integral when α is set equal to $m/2kT$, which is usually the case when working with Boltzmann velocity distributions. The fourth column expresses the result shown in the third column as a multiple of the mean thermal speed \bar{c} , where

$$\bar{c} = \sqrt{8kT/\pi m}$$

Suppose now that we wish to know how many particles of the system have velocities such that the x -component of velocity is in a range dv_x about v_x , the y -component is in a range dv_y about v_y and the z -component is in a range dv_z about v_z , thus the number of particles in a rectangular volume element $(dv_x dv_y dv_z)$ in velocity space centered on the value (v_x, v_y, v_z) . We shall call this number $N(v_x, v_y, v_z) dv_x dv_y dv_z$. We must now proceed more carefully, starting with (5.4-6), which according to (5.4-12) we may write as

$$f(v_x, v_y, v_z) = \frac{N}{2V} \left(\frac{h^2}{2\pi mkT} \right)^{3/2} e^{-m(v_x^2 + v_y^2 + v_z^2)/2kT}. \quad (5.4-15)$$

Then, as required by (5.2-1),

$$N(v_x, v_y, v_z) dv_x dv_y dv_z = f(v_x, v_y, v_z) g(v_x, v_y, v_z) dv_x dv_y dv_z. \quad (5.4-16)$$

where $g(v_x, v_y, v_z) dv_x dv_y dv_z$ is the number of quantum states within the velocity space element $dv_x dv_y dv_z$. From (5.2-17) the number of states in a volume element $dp_x dp_y dp_z$ of momentum space is $(2V/h^3)(dp_x dp_y dp_z)$, whereby

$$\frac{2V}{h^3} dp_x dp_y dp_z = \frac{2m^3 V}{h^3} dv_x dv_y dv_z = g(v_x, v_y, v_z) dv_x dv_y dv_z. \quad (5.4-17)$$

Since the density of states in momentum space is constant, so also is the density of states in velocity space. Substituting (5.4-15) and (5.4-17) into (5.4-16) and simplifying, we find

$$N(v_x, v_y, v_z) dv_x dv_y dv_z = N \left(\frac{m}{2\pi k T} \right)^{3/2} e^{-m(v_x^2 + v_y^2 + v_z^2)/2kT} dv_x dv_y dv_z. \quad (5.4-18)$$

A third velocity distribution function, $N(v_x)dv_x$, representing the number of particles whose x -component of velocity lies in a range dv_x about v_x , regardless of what values the y - and z -components of velocity for those particles may have, is closely related to the distribution function (5.4-18). The distribution function $N(v_x)dv_x$ can be calculated from (5.4-18) by integrating over all possible values of v_y and v_z , the result being

$$N(v_x)dv_x = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [N(v_x, v_y, v_z)dv_x] dv_y dv_z = N \sqrt{\frac{m}{2\pi k T}} e^{-mv_x^2/2kT} dv_x. \quad (5.4-19)$$

The values of the definite integrals required in integrating (5.4-18) have been taken from Table 5.1. A plot of the distributions $N(v)$ and $N(v_x)$ as functions of the appropriate velocity coordinate is shown in Figure 5.6.

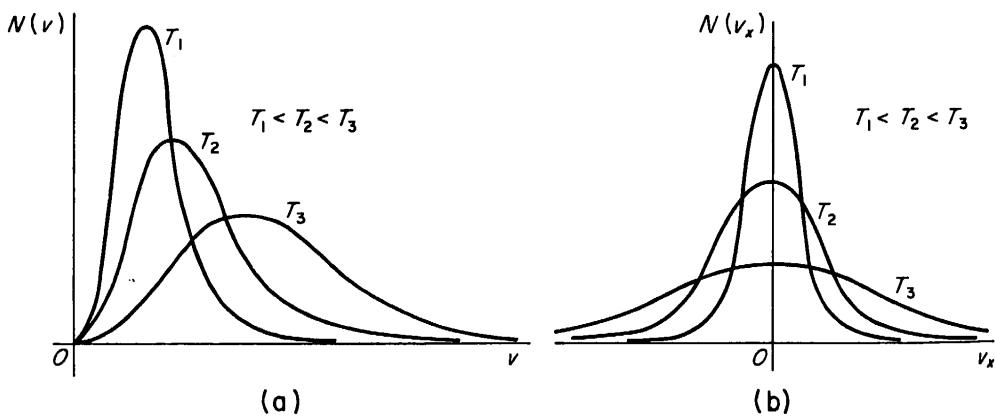


FIGURE 5.6. (a) A schematic representation of the Maxwell-Boltzmann distribution of particle speeds, $N(v)$, for three different temperatures. (b) The corresponding distribution of x -components of velocity, $N(v_x)$.

Knowing these three velocity distribution functions it is an easy matter to evaluate averages over the velocity distributions. For example, the average thermal speed \bar{v}

velocity space
t $dp_x dp_y dp_z$

for a particle in a Boltzmann distribution is obtained quite simply from (5.4-14) by writing

(5.4-17)

$$\langle v \rangle = \frac{\int_0^\infty v N(v) dv}{\int_0^\infty N(v) dv} = \frac{1}{N} \cdot 4\pi N \left(\frac{m}{2\pi kT} \right)^{3/2} \int_0^\infty v^3 e^{-mv^2/2kT} dv$$

density of
simplifying,

 v_z .

(5.4-18)

of particles
what values
 v related to
calculated
being

 dv_x .

(5.4-19)

been taken
the appro-

 $v_2 < T_3$ v_x stribu-
onding

o evaluate
al speed \bar{c}

$$= \sqrt{8kT/\pi m}, \quad (5.4-20)$$

the integral having been evaluated with the help of Table 5.1.

We are now in a position to discuss the equation of state of an ideal gas. Consider the particles striking unit area of the wall of a container filled with such a gas, as shown in Figure 5.7. If the wall is a plane oriented normal to the x -axis, the momentum components p_y and p_z of the particles striking the wall are conserved if the collisions of the

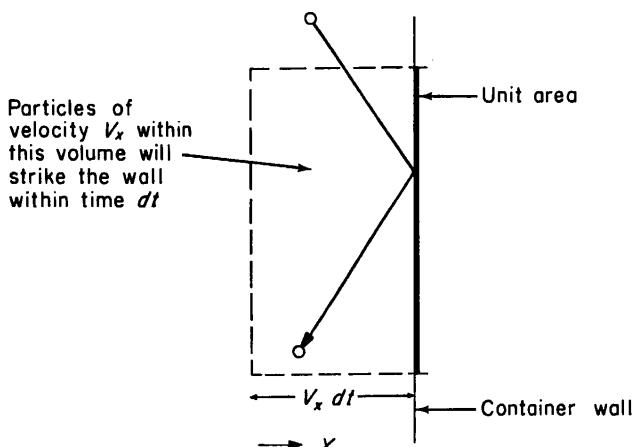


FIGURE 5.7. The elastic collision of a particle whose x -component of velocity is v_x with the wall of a rigid, fixed container.

particles with the walls are elastic, which we shall assume to be the case. In each collision the velocity component v_x directed toward the wall is changed to $-v_x$, directed away from the wall. The transfer of momentum from the particle to the wall, per collision, is thus $2mv_x$. The transfer of momentum to the wall in time dt is this quantity times the number of particles which collide with the wall during that time; for particles whose x -component of velocity is v_x , this is simply the number of such particles in a volume extending a distance $d = v_x dt$ behind the wall, or the number per unit volume times $v_x dt$, thus $v_x dt \cdot N(v_x)dv_x/V$. The momentum transfer in time dt for particles of velocity in the range dv_x about v_x is then

$$(dp_x)_{v_x} = 2mv_x \frac{v_x dt \cdot N(v_x)dv_x}{V}$$

or,

$$\left(\frac{dp_x}{dt} \right)_{v_x} = \frac{2m}{V} v_x^2 N(v_x)dv_x. \quad (5.4-21)$$

The total rate of momentum transfer to the walls by collisions involving particles with all possible values of v_x can be obtained by integrating over v_x , giving, with the help of (5.4-19) and Table 5.1,

$$\frac{dp_x}{dt} = \frac{2m}{V} N \sqrt{\frac{m}{2\pi kT}} \int_0^\infty v_x^2 e^{-mv_x^2/2kT} dv_x = \frac{NkT}{V}. \quad (5.4-22)$$

The time rate of transfer of momentum to unit area of the container wall, however, is according to Newton's law equal to the force experienced by a unit area of the wall, which by definition is the pressure P . Equation (5.4-22) then reduces to

$$PV = NkT, \quad (5.4-23)$$

which is the equation of state for a Boltzmann gas of independent particles.

This equation has the form of the familiar ideal gas law, which is more commonly written in the form

$$PV = nRT, \quad (5.4-24)$$

where n is the number of moles of gas in the system and R is an experimentally measured "molar gas constant" which is the same for all "ideal" gases. If there are n moles of gas present, then the number of particles must be given by $N = nN_A$, where N_A is Avogadro's number, equal to 6.026×10^{23} molecules per mole. Under these conditions (5.4-23) takes the form

$$PV = nN_A kT, \quad (5.4-25)$$

whence, comparing (5.4-24) with (5.4-25), it is clear that Boltzmann's constant k must be given by

$$k = R/N_A. \quad (5.4-26)$$

Boltzmann's constant is thus simply the gas constant *per particle* of the system. Its value can be derived from N_A and R , and is equal to 1.380×10^{-16} ergs/ $^{\circ}$ K, or 8.615×10^{-5} eV/ $^{\circ}$ K.

It is clear now that the value for the constant β in the distribution function which was assumed in (5.3-27) was chosen correctly. Had we carried through all our calculations up to this point without ever having assumed any value for β , we should have found for equation (5.4-23), instead,

$$PV = -N/\beta, \quad (5.4-27)$$

and in order that our results be in agreement with the experimentally established gas law (5.4-24) we should have been *forced* to choose

$$-N/\beta = nRT = \frac{N}{N_A} RT,$$

or,

$$\beta = -\frac{N_A}{RT} = -1/kT. \quad (5.4-28)$$

5.5 FERMI-DIRAC STATISTICS

In the development of the Maxwell-Boltzmann distribution function, the particles are considered to be distinguishable, while in actual fact it is quite impossible to distinguish one electron or other elementary particle from another. Furthermore, we permitted any number of particles to occupy the same quantum state of the system, in spite of the fact that many particles, *electrons* in particular, obey the Pauli exclusion principle, which allows each quantum state to accept no more than one particle. If these additional conditions are imposed upon the system, the calculations of Section 5.4 must be modified, resulting in another distribution function, which is called the Fermi-Dirac distribution function. This energy distribution is of the greatest importance, since it describes the statistical behavior of free electrons in metals and semiconductors, and since many of the electrical and thermal properties of solids which could not be understood at all on the basis of classical statistics follow as a direct consequence of the Fermi-Dirac statistics.

If the particles of the system are indistinguishable, they cannot be identified by number in the manner which was adopted when discussing the various possibilities illustrated in Figure 5.2. As a matter of fact, all the various distributions in each row of Figure 5.2 leading to $Q(N_1, N_2)$ for a particular value of N_1 are *the same* if the numbered labels are removed from the particles, in which case the factor (5.3-7) reduces to unity. It is still possible, however, to permute the N_i particles in the i th energy level among the g_i quantum states belonging to that level in many ways, each of which constitutes a statistically independent way of achieving an arrangement wherein N_i particles are in the i th energy level of the system, as shown in Figure 5.8. The product

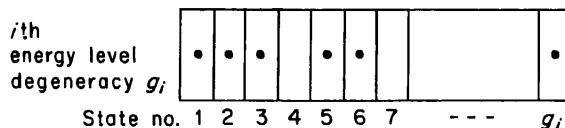


FIGURE 5.8. A possible distribution of particles among quantum states in the i th energy level of a system wherein the Pauli exclusion principle is applicable.

of the possible numbers of permutations of particles among quantum states over all energy levels of the system then gives the number of independent ways of realizing a given distribution of particles among energy levels.

For the Fermi-Dirac case, in which the Pauli exclusion principle is assumed to apply, a maximum of one particle per quantum state is allowed. Referring to Figure 5.8, which depicts the i th energy level of the system, there are g_i ways of choosing where to insert a first particle, $g_i - 1$ ways of choosing where to insert a second, since the second cannot occupy the same quantum state as the first, $g_i - 2$ ways of choosing where to insert a third, and so forth. The total number of ways of arranging N_i particles in the i th level obtained on this basis is

$$g_i(g_i - 1)(g_i - 2) \cdots (g_i - N_i + 1) = \frac{g_i!}{(g_i - N_i)!}. \quad (5.5-1)$$

However, since the particles are indistinguishable, the $N_i!$ ways of permuting the

particles among themselves in any given arrangement of particles among states do not count as separate arrangements. The actual number of independent ways of realizing a distribution of N_i particles in the i th level is thus arrived at by dividing (5.5-1) by $N_i!$, giving

$$\frac{g_i!}{N_i!(g_i - N_i)!}. \quad (5.5-2)$$

The total number of independent ways of realizing a distribution of (N_1, N_2, \dots, N_n) indistinguishable particles among n energy levels, no more than one per quantum state, is just the product of individual factors of the form (5.5-2) over all the levels, that is,

$$Q_f(N_1, N_2, \dots, N_n) = \prod_{i=1}^n \frac{g_i!}{N_i!(g_i - N_i)!}. \quad (5.5-3)$$

We now proceed to maximize the logarithm of this quantity with respect to the variables N_1, N_2, \dots, N_n by the method of Lagrangean multipliers used in Section 5.3. We find from (5.5-3) that

$$\ln Q_f = \sum_i \ln g_i! - \sum_i \ln N_i! - \sum_i \ln(g_i - N_i)!, \quad (5.5-4)$$

which by using Stirling's approximation (5.3-20) can be written

$$\ln Q_f = \sum_i [g_i \ln g_i - N_i \ln N_i - (g_i - N_i) \ln(g_i - N_i)]. \quad (5.5-5)$$

Again, it is required that the total number of particles in the system and the total energy of the system be constant, which means that Equations (5.3-22) and (5.3-23) must hold here also, in which case one may write, as before,

$$\frac{\partial(\ln Q_f)}{\partial N_j} + \alpha \frac{\partial \phi}{\partial N_j} + \beta \frac{\partial \psi}{\partial N_j} = 0 \quad (j = 1, 2, \dots, n), \quad (5.5-6)$$

where ϕ and ψ are given by (5.3-22) and (5.3-23). Substituting (5.5-5), (5.3-22), and (5.3-23) into this equation, one obtains

$$-\frac{\partial}{\partial N_j} \left[\sum_i N_i \ln N_i + \sum_i (g_i - N_i) \ln(g_i - N_i) \right] + \alpha \frac{\partial}{\partial N_j} \left(\sum_i N_i \right) + \beta \frac{\partial}{\partial N_j} \left(\sum_i \varepsilon_i N_i \right) = 0. \quad (5.5-7)$$

Working out the derivatives as in Section 5.3, we find

$$\ln(g_j - N_j) - \ln N_j = -\alpha - \beta \varepsilon_j, \quad (5.5-8)$$

or, rearranging, exponentializing and solving for $N_j/g_j = f(\varepsilon_j)$,

$$f(\varepsilon_j) = N_j/g_j = \frac{1}{1 + e^{-\alpha - \beta \varepsilon_j}}. \quad (5.5-9)$$

tates do not
realizing a
g (5.5-1) by

This is the Fermi-Dirac distribution function.

As in Section 5.4, we shall take the value of β to be

$$\beta = -1/kT, \quad (5.5-10)$$

(5.5-2)

postponing until later the justification for this step. It is customary to write α in the form

$$\alpha = \varepsilon_f/kT, \quad (5.5-11)$$

where ε_f is a parameter with the dimensions of energy, which is called the *Fermi energy*, or the Fermi level, of the system. Equation (5.5-9) then becomes

(5.5-3)

$$N_j = \frac{g_j}{1 + e^{(\varepsilon_j - \varepsilon_f)/kT}}, \quad (5.5-12)$$

or, if the levels are assumed to crowd together into a continuum, so that $g_j(\varepsilon_j) \rightarrow g(\varepsilon)d\varepsilon$, then

(5.5-4)

$$N(\varepsilon)d\varepsilon = g(\varepsilon)f(\varepsilon)d\varepsilon = \frac{g(\varepsilon)d\varepsilon}{1 + e^{(\varepsilon - \varepsilon_f)/kT}}. \quad (5.5-13)$$

For a gas of independent particles such as free electrons, $g(\varepsilon)d\varepsilon$ is represented by (5.2-22), just as for a Maxwell-Boltzmann gas.

The Fermi energy ε_f is in general a function of the temperature, whose form and temperature dependence are critically dependent upon the density of states function for the system, just as for the corresponding parameter e^α of the Maxwell-Boltzmann distribution. Its value is determined by the condition (5.3-22), or for a continuum of levels, by

(5.5-6)

$$N = \text{const.} = \int \frac{g(\varepsilon)d\varepsilon}{1 + e^{(\varepsilon - \varepsilon_f)/kT}} = \int g(\varepsilon)f(\varepsilon)d\varepsilon. \quad (5.5-14)$$

The integral is taken over all energies available to the particles of the system. For a Fermi gas of independent particles, $g(\varepsilon)$ is given by (5.2-22) and ε_f is determined by

$$\varepsilon_i N_i = 0. \quad (5.5-7)$$

$$N = \frac{8\sqrt{2\pi}Vm^{3/2}}{h^3} \int_0^\infty \frac{\sqrt{\varepsilon} d\varepsilon}{1 + e^{(\varepsilon - \varepsilon_f)/kT}}. \quad (5.5-15)$$

Unfortunately, this integral cannot be evaluated in closed analytic form, so that ε_f cannot be determined as a simple function of the temperature. For a *two-dimensional* independent particle gas, however, the density of states function can be shown by the methods used to derive (5.2-22) to be

$$g(\varepsilon)d\varepsilon = \frac{4\pi m A}{h^2} d\varepsilon, \quad (5.5-16)$$

(5.5-9)
independent of energy (Exercise 3, Chapter 5). In this formula, A represents the

area of the two-dimensional "container." For this system (5.5-14) becomes

$$N = \frac{4\pi mA}{h^2} \int_0^\infty \frac{d\varepsilon}{1 + e^{(\varepsilon - \varepsilon_f)/kT}}, \quad (5.5-17)$$

which can be evaluated in closed form, allowing one to solve for ε_f and obtain

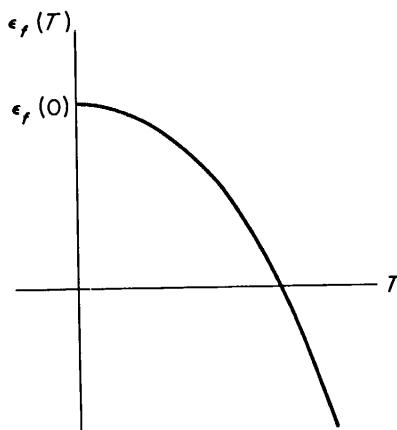
$$\varepsilon_f(T) = kT \ln(e^{\varepsilon_f(0)/kT} - 1), \quad (5.5-18)$$

where

$$\varepsilon_f(0) = \frac{Nh^2}{4\pi mA} \quad (5.5-19)$$

is the value which $\varepsilon_f(T)$ as given by (5.5-18) assumes as T approaches zero. The details of calculating this result are very instructive, and are assigned as an exercise for the reader. The variation of the Fermi energy of the two-dimensional independent particle gas with temperature is shown in Figure 5.9. It will be noted for this case that the Fermi energy is a monotonic decreasing function of temperature. The Fermi energy for the three-dimensional Fermi gas, with the density of states function (5.2-22) will be found to exhibit the same general behavior, except that in this case the variation of the Fermi energy with temperature is *linear* with temperature at reasonably low temperatures, while for the two-dimensional example the variation is much more complex at low temperatures. For many systems, including these two, the variation of the Fermi level with temperature is quite small over the range of physically realizable temperatures; in Figure 5.9 the temperature for which $\varepsilon_f = 0$ would be of the order of

FIGURE 5.9. Schematic representation of the variation of Fermi energy with temperature in a two-dimensional free electron gas, according to (5.5-18).



75 000°K for an electron gas with the free-electron density of metallic copper. For this reason, in many applications the temperature dependence of the Fermi energy may either be neglected or approximated by a linear or other appropriate function of temperature.

The Fermi distribution function itself,

$$f(\varepsilon) = \frac{1}{1 + e^{(\varepsilon - \varepsilon_f)/kT}}, \quad (5.5-20)$$

is plotted in Figure 5.10 for several values of the temperature. Since only one particle

SEC. 5.5

SEC. 5.5

OUTLINE OF STATISTICAL MECHANICS

153

nes

(5.5-17)

may occupy a given quantum state, the value of $f(\epsilon)$ for a Fermi distribution at a particular energy is just equal to the probability that a quantum state of that energy will be occupied. At absolute zero, it is easily seen from Figure 5.10 and from (5.5-20) that the Fermi distribution function becomes simply the step function

$$\begin{aligned} f(\epsilon) &= 1 & (\epsilon < \epsilon_f) \\ &= 0 & (\epsilon > \epsilon_f). \end{aligned} \quad (5.5-21)$$

(5.5-18)

(5.5-19)

The details
cise for the
ndependent
or this case
The Fermi
ion (5.2-22)
ie variation
onably low
nuch more
e variation
y realizable
he order of

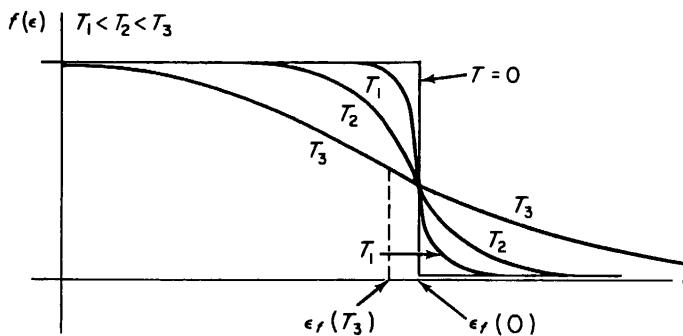


FIGURE 5.10. Schematic representation of the Fermi distribution function for four different temperatures. Note the variation of the Fermi energy with temperature. The temperature dependence of the Fermi energy depicted here is typical of a three-dimensional free-electron gas, but the actual variation in any particular system will depend critically upon the density of states function (or level degeneracies) for that system.

As the temperature increases, the edges of the step are rounded off, and the distribution function varies rapidly from nearly unity to nearly zero over an energy range of a few times kT around the value $\epsilon = \epsilon_f$. At the same time, the value of ϵ_f itself changes, the variation illustrated in Figure 5.10 being approximately that associated with the three-dimensional electron gas whose density of states is given by (5.2-22). At very high temperatures, the distribution function loses its step-like character and varies much more slowly with energy. From (5.5-20), it is clear that the value of $f(\epsilon)$ at $\epsilon = \epsilon_f$ is just $\frac{1}{2}$, that is,

$$f(\epsilon_f) = \frac{1}{2}, \quad (5.5-22)$$

hence a quantum state at the Fermi level has a probability of occupation of $\frac{1}{2}$.

Figure 5.11 shows the actual distribution of particle density $N(\epsilon)$ as a function of energy for a Fermi gas of independent particles, as given by (5.5-13) with the density of states function (5.2-22). Again, at $T = 0$ the curve has a step-like character, the portion for which ($\epsilon < \epsilon_f$) being the density of states parabola (5.2-22) and that for which ($\epsilon > \epsilon_f$) being zero. As the temperature increases this step-like aspect becomes less and less pronounced, as shown in the drawing. At low temperatures, when the Fermi distribution function is step-like, the distribution is said to be highly *degenerate*.

At low temperatures, the Fermi-Dirac distribution may be represented as a sphere in momentum space in which all or most of the quantum states of energy less than ϵ_f

(5.5-20)

e particle

are filled, while all or most of the states of energy greater than ϵ_f are empty. From (5.2-18), the equation of the surface of this "Fermi sphere" must be

$$p_x^2 + p_y^2 + p_z^2 = 2me_f, \quad (5.5-23)$$

the radius therefore being $\sqrt{2me_f}$, as shown in Figure 5.12. At very high temperatures the surface of the Fermi sphere becomes poorly defined, due to the disappearance of the step-like aspect of $f(\epsilon)$, and the concept loses some of its usefulness.

FIGURE 5.11. Schematic representation of electron density as a function of energy for a three-dimensional free-electron gas.

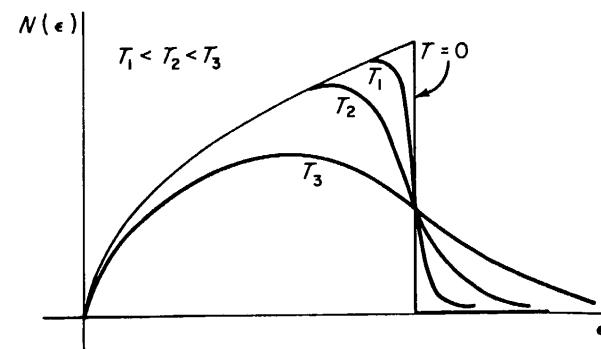
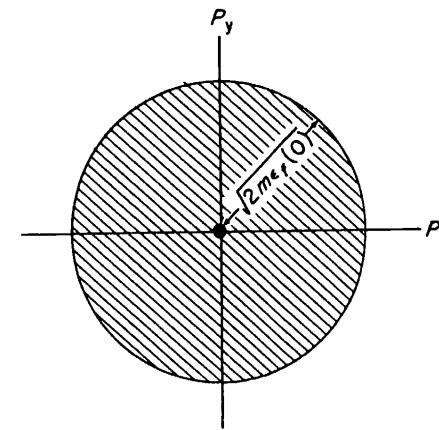


FIGURE 5.12. The representation of the absolute zero Fermi distribution for a free-electron gas as a sphere of electrons in momentum space. This representation is a useful one so long as $T \ll T_F$, where $T_F = \epsilon_f(0)/k$.



(z - axis \perp to paper)

Although, as we have already seen, the fact that the integral (5.5-15) cannot be evaluated in closed form prevents us from finding a simple expression for the Fermi energy of a three-dimensional Fermi gas at all temperatures, the simple character (5.5-21) which the function assumes at $T = 0$ permits one to evaluate ϵ_f very simply at absolute zero. Using (5.5-21) to represent $f(\epsilon)$ at $T = 0$, we may rewrite (5.5-14), using the density of states function (5.2-22), as

$$\frac{N}{V} = \frac{8\sqrt{2}\pi m^{3/2}}{h^3} \int_0^{\epsilon_f(0)} \sqrt{\epsilon} d\epsilon = \frac{16\sqrt{2}\pi m^{3/2} [\epsilon_f(0)]^{3/2}}{3h^3}. \quad (5.5-24)$$

Solving this equation for $\epsilon_f(0)$, we may obtain

$$\epsilon_f(0) = \frac{h^2}{8m} \left(\frac{3N}{\pi V} \right)^{2/3}. \quad (5.5-25)$$

empty. From

(5.5-23)

temperatures
pearance of

In a somewhat similar manner, it is possible to obtain the internal energy of a Fermi gas at absolute zero, the result being

$$\frac{U_0}{V} = \frac{\pi h^2}{40m} \left(\frac{3N}{\pi V} \right)^{5/3} = \frac{3}{5} \frac{N}{V} \varepsilon_f(0). \quad (5.5-26)$$

The details of this calculation are assigned as an exercise.

For energies which are much greater than ε_f , $e^{(\varepsilon - \varepsilon_f)/kT}$ is much larger than unity and for such energies the Fermi-Dirac distribution function (5.5-20) may be written approximately as

$$f(\varepsilon) \cong e^{\varepsilon_f/kT} e^{-\varepsilon/kT}. \quad (5.5-27)$$

If all the energies available to the system satisfy the condition

$$\varepsilon - \varepsilon_f \gg kT, \quad (5.5-28)$$

that is, if ε_f is many kT units smaller than any energy a particle belonging to the system may have, then (5.5-27) will be a good approximation to (5.5-20) for all particles of the system. The approximate distribution function (5.5-27), however, is simply the Maxwell-Boltzmann distribution function of (5.4-6) with $\alpha = \varepsilon_f/kT$. If the condition (5.5-28) holds, then, for all particles of the system, the Fermi-Dirac distribution and the Maxwell-Boltzmann distribution are very nearly the same.

For the two-dimensional Fermi gas, where the Fermi energy is given by (5.5-18), if T is so large that $kT \gg \varepsilon_f(0)$, the exponent $\varepsilon_f(0)/kT$ will be small, so that the exponential can be approximated by $1 + [\varepsilon_f(0)/kT]$, giving

$$\begin{aligned} \varepsilon_f(T) &= kT \ln \frac{\varepsilon_f(0)}{kT} \\ &= \varepsilon_f(0) \left[\frac{kT}{\varepsilon_f(0)} \ln \frac{\varepsilon_f(0)}{kT} \right] \\ &= -\varepsilon_f(0) \left[\frac{kT}{\varepsilon_f(0)} \ln \frac{kT}{\varepsilon_f(0)} \right]. \end{aligned} \quad (5.5-29)$$

From this, we see that as T becomes large, $\varepsilon_f(T) \rightarrow -\infty$. Since the lowest energy any particle of the system may have is zero, it is clear that condition (5.5-28) will be fulfilled for sufficiently high temperatures, and the distribution function will therefore be approximately the same as the Maxwell-Boltzmann distribution at very high temperatures. The same result can be shown to hold for the three-dimensional case. These results are to be expected on physical grounds, since at high temperatures the particles are distributed over a very wide range of energy states, the number of particles in every range of available energies being so small that there are always many more available quantum states than there are particles to occupy them. Under these circumstances, the probability of *two or more* particles occupying the same quantum state becomes vanishingly small in any case, so that there is not much difference

(5.5-24)

(5.5-25)

between the distribution function for which the Pauli principle is obeyed (the Fermi-Dirac distribution) and that for which it is not (the Maxwell-Boltzmann distribution).³ Since the quantity $\epsilon_f(0)$ which is required to be much less than kT for the two distributions to coincide is proportional to the density of particles N/A by (5.5-19), the reduction of the Fermi distribution to a Maxwell-Boltzmann distribution will take place at lower temperatures in less dense gases. By the same token, according to (5.5-19) it will take place at lower temperatures in gases where the particle mass m is large. It is for these reasons that ordinary gaseous substances obey the Maxwell-Boltzmann statistics at normal temperatures rather than the Fermi-Dirac (or Bose-Einstein) statistics. For a dense gas of very light particles, such as the free electrons in a metal, however, the Fermi energy at absolute zero is quite large, and the condition (5.5-28) can be satisfied for all particles of the system only at temperatures so high as to be unrealizable physically. A dense free electron gas must therefore be treated using Fermi-Dirac statistics. In semiconductors, however, the peculiar form of the density of states function is such that the Maxwell-Boltzmann distribution is virtually *always* a good approximation to the Fermi-Dirac distribution. We shall examine this situation in considerable detail in a later chapter.

Had we made a choice for the undetermined multiplier β other than that given by (5.5-10) we should *not* in general have found the correspondence between the Fermi-Dirac and Maxwell-Boltzmann systems in the high-temperature limit, which, as we have seen, we have every right to expect on physical grounds. We must therefore conclude that the value for β as given by (5.5-10) is physically justified.

5.6 THE BOSE-EINSTEIN DISTRIBUTION

In the previous section we showed that the statistical distribution which characterizes the behavior of an ensemble of indistinguishable particles which obey the Pauli exclusion principle is the Fermi-Dirac distribution (5.5-20). Since not all elementary particles obey the Pauli principle (photons being the most conspicuous exception) it is necessary to consider the statistical behavior of particles which, though indistinguishable, do not obey the Pauli exclusion principle. In this case again, since the particles are not numbered, the factor (5.3-7) reduces to unity and we need only consider the possible permutations of N_i identical particles among the g_i quantum states of the i th energy level, but now with no restrictions with regard to the number of particles which may occupy any given quantum state.

Consider a linear array of N_i particles and $g_i - 1$ partitions which would be necessary to divide these particles into g_i groups, as shown in Figure 5.13. It is not difficult to see that the number of ways of permuting the N_i particles among g_i levels is equal to the number of independent permutations of objects and partitions in Figure 5.13.

³ It might be contended that the indistinguishability of particles renders the Fermi-Dirac distribution distinct from the Maxwell-Boltzmann distribution even in this limit. It can, however, be shown, as we shall see in the next section, that the distribution function for *indistinguishable* particles which do not obey the Pauli exclusion principle (the Bose-Einstein distribution) also approaches the Maxwell-Boltzmann distribution function in this limit. A more accurate way of explaining the situation would be to say that in the high-temperature limit the Fermi-Dirac distribution approaches the Bose-Einstein distribution, which in turn approaches the Maxwell-Boltzmann distribution.

(the Fermi-distribution).³ In a metal, the reduced form of the distribution function (5.5-19) is large. It is Boltzmann-like (Bose-Einstein) in a metal, and (5.5-28) shows that it is calculated using the density matrix, always giving the same result as the situation

that gives the result, which, it therefore

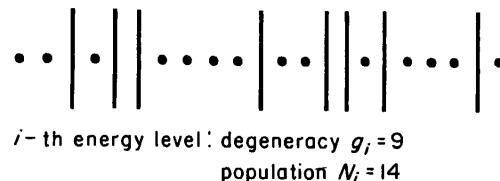
characterizes the Pauli exclusion principle (it is indistinguishable particles) and consider the behavior of the i th group of particles

be necessary, but difficult. This is equal to figure 5.13.

Fermi-Dirac however, because particles approaches the g the situation approaches the

Since there are a total of $N_i + g_i - 1$ particles plus partitions, these can be arranged linearly in $(N_i + g_i - 1)!$ ways, but since permutations of particles among themselves or of partitions among themselves do not count as independent arrangements, we must divide by the number of ways of permuting particles among themselves ($N_i!$)

FIGURE 5.13. A possible distribution of particles among quantum states in the i th energy level of a system to which the Pauli exclusion principle does not apply.



and again by the number of ways of permuting partitions among themselves $((g_i - 1)!)!$, giving

$$\frac{(N_i + g_i - 1)!}{N_i!(g_i - 1)!} \quad (5.6-1)$$

ways of realizing a distribution of $N_i!$ indistinguishable particles among g_i states which may accommodate any number of particles. The number Q_b of statistically independent ways of achieving a distribution (N_1, N_2, \dots, N_n) particles among the energy levels of the system according to these rules is then just the product of factors of the form (5.6-1) over all levels of the system, whence

$$Q_b(N_1, N_2, \dots, N_n) = \prod_{i=1}^n \frac{(N_i + g_i - 1)!}{N_i!(g_i - 1)!}. \quad (5.6-2)$$

It is now possible to maximize this quantity with respect to the variables N_1, N_2, \dots, N_n , using the method of Lagrangean multipliers, under the restrictions (5.3-22) and (5.3-23). The actual calculations will not be set forth here, but will be left as an exercise for the reader. The result is

$$f(\varepsilon_j) = N_j/g_j = \frac{1}{e^{-\alpha} e^{-\beta\varepsilon_j} - 1}. \quad (5.6-3)$$

This formula is referred to as the Bose-Einstein distribution function. It is again possible to identify β as

$$\beta = -1/kT, \quad (5.6-4)$$

while α may be determined in terms of the number of particles in the system, just as in the Maxwell-Boltzmann and Fermi-Dirac cases. For the case of a continuum of closely spaced levels, $g_j \rightarrow g(\varepsilon)d\varepsilon$ and $N_j \rightarrow N(\varepsilon)d\varepsilon$, giving

$$f(\varepsilon) = \frac{1}{e^{-\alpha} e^{+\varepsilon/kT} - 1}. \quad (5.6-5)$$

For the independent particle density of states function (5.5-16) for a two-dimensional free particle gas, the parameter α may be explicitly calculated in the same way as ε_f for

this system under Fermi-Dirac statistics. It may then be shown that as T becomes large, $\alpha \rightarrow -\infty$, in which case the exponential factor in the denominator of (5.6-5) becomes much larger than unity, and the Bose-Einstein distribution (5.6-5) approaches

$$f(\varepsilon) \cong e^\alpha e^{-\varepsilon/kT}, \quad (5.6-6)$$

which is a distribution function of the Maxwell-Boltzmann type. The same general behavior is obtained for a three-dimensional independent-particle Bose-Einstein gas, although in this case it is not possible to obtain an expression for α in closed form.

In the limit of very high temperatures, the particles of the system will be distributed over a very wide range of energies, and the number of particles in every available energy range will become much smaller than the number of quantum states in that range, whence for all states $g_i \gg N_i$. In this situation, we may write, approximately,

$$\frac{(N_i + g_i - 1)!}{(g_i - 1)!} \cong g_i^{N_i}, \quad (5.6-7)$$

whereby (5.6-2) becomes

$$Q_b(N_1, N_2, \dots, N_n) \cong \prod_{i=1}^n \frac{g_i^{N_i}}{N_i!}. \quad (5.6-8)$$

But this, apart from a constant factor $N!$, is just equal to $Q(N_1, N_2, \dots, N_n)$ as given by (5.3-8) for a Maxwell-Boltzmann system! We should then expect the Bose-Einstein and Maxwell-Boltzmann distributions to coincide in the high-temperature limit on purely physical grounds. The choice of the value given by (5.6-4) for β is thus justified, for this choice, as we have seen above, leads directly to the correspondence between the Bose-Einstein and Maxwell-Boltzmann Statistics shown by (5.6-6).

In the case of the two-dimensional and three-dimensional independent particle Bose-Einstein gases, as the temperature approaches zero, the value of α tends toward zero, the result being that all the particles of the system tend to condense into the lowest energy state of the system at absolute zero. This phenomenon, called the *Bose condensation*, is characteristic of systems obeying Bose-Einstein statistics.

In some applications it is of interest to obtain the Bose-Einstein distribution function without making the restriction that the number of particles in the system be constant. It can be seen from (5.3-15) and (5.3-16) that this result can be obtained from (5.6-5) if α is taken to be identically zero. In this case (5.6-5) reduces to

$$f(\varepsilon) = \frac{1}{e^{\varepsilon/kT} - 1}. \quad (5.6-9)$$

EXERCISES

1. Show that the wave front associated with the plane wave $\Psi = e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)}$ advances along the \mathbf{k} direction.
2. Show for the free particle in three dimensions, whose wave function is given by

T becomes
r of (5.6-5)
approaches
(5.6-6)

me general
instein gas,
osed form.
ll be distri-
tely available
ites in that
roximately,
(5.6-7)

(5.6-8)

) as given
se-Einstein
e limit on
is justified,
 \propto between

nt particle
ids toward
the lowest
use condens-

distribution
the system
e obtained
s to
(5.6-9)

' advances
s given by

(5.2-7) that the expectation value of the vector momentum \mathbf{p} is equal to $\hbar\mathbf{k}$, and that \mathbf{p} (hence \mathbf{k}) is a constant of the motion.

3. Calculate the density of states factor $g(\epsilon)d\epsilon$ for a two-dimensional system of free particles, with instantaneous collision interactions only, contained within a rigid container of area A and dimensions x_0 and y_0 . Start with Schrödinger's equation.

4. Suppose that 4 coins are tossed simultaneously; what are the probabilities associated with the distributions (0 heads, 4 tails), (1 head, 3 tails) ... (4 heads, 0 tails)?

5. Find the dimensions and area of the rectangle of maximum area with sides parallel to the coordinate axes which can be inscribed within an ellipse whose major axis is $2a$ and whose minor axis is $2b$. The axes of the ellipse may be taken to be parallel with the coordinate axes. Use the method of Lagrangean multipliers.

6. Find the root-mean-square speed and the most probable speed of a particle in an ideal Boltzmann gas.

7. Show that the *flux* of particles in an ideal Boltzmann gas, whose x -components of velocity are positive, per unit area across a plane normal to the x -axis is $\frac{1}{4}N\bar{c}/V$. Hint: The flux or current density is defined as the number of particles per unit volume times their velocity component along the normal to the plane across which the flux is observed.

8. Show that the Fermi energy of a two-dimensional Fermi gas of free particles, whose density of states function is given by the result of Exercise 3, is $\epsilon_f(T) = kT \ln(e^{\epsilon_f(0)/kT} - 1)$, where $\epsilon_f(0) = Nh^2/(4\pi m A)$.

9. Calculate the Fermi energy, in electron volts, for the free electrons in copper at absolute zero, assuming one free electron per copper atom. For what temperature would kT be equal to $\epsilon_f(0)$?

10. Show that the internal energy per unit volume of a three-dimensional Fermi gas of free particles at absolute zero is $\frac{3}{5} \frac{N}{V} \epsilon_f(0)$; show that the corresponding result for a two-dimensional Fermi gas is $\frac{1}{2} \frac{N}{V} \epsilon_f(0)$.

11. Show by the method of Lagrangean multipliers that if $Q(N_1, N_2, \dots, N_n)$ is given by Equation (5.6-2), the distribution function corresponding to the most probable values of N_1, N_2, \dots, N_n is given by (5.6-3).

12. Using the density of states function calculated in Exercise 3, show that the parameter α of a two-dimensional Bose-Einstein gas of free particles is given by

$$\alpha = \ln(1 - e^{-h^2 N / (2\pi A m k T)})$$

Discuss the properties of the resulting distribution function for $T \rightarrow \infty$ and for $T \rightarrow 0$.

GENERAL REFERENCES

- W. Band, *An Introduction to Quantum Statistics*, D. Van Nostrand, Princeton, N.J. (1955).
- R. W. Gurney, *Introduction to Statistical Mechanics*, McGraw-Hill Book Co., Inc., New York (1949).
- D. ter Haar, *Elements of Statistical Mechanics*, Holt, Rinehart-and Winston, Inc., New York (1954).
- R. B. Lindsay, *Introduction to Physical Statistics*, John Wiley and Sons, New York (1941).
- J. E. Mayer and Maria Göppert-Mayer, *Statistical Mechanics*, John Wiley and Sons, New York (1940).
- R. C. Tolman, *The Principles of Statistical Mechanics*, Oxford University Press, London (1938).

CHAPTER 1

INTRODUCTION AND REVIEW

1-1 INTRODUCTION

Statistical mechanics is that branch of physics which studies macroscopic systems from a microscopic or molecular point of view. The goal of statistical mechanics is the understanding and prediction of macroscopic phenomena and the calculation of macroscopic properties from the properties of the individual molecules making up the system.

Present-day research in statistical mechanics varies from mathematically sophisticated discussions of general theorems to almost empirical calculations based upon simple, but nevertheless useful, molecular models. An example of the first type of research is the investigation of the question of whether statistical mechanics, as it is formulated today, is even capable of predicting the existence of a first-order phase transition. General questions of this kind are by their nature mathematically involved and are generally beyond the level of this book. We shall, however, discuss such questions to some extent later on. On the other hand, for many scientists statistical mechanics merely provides a recipe or prescription which allows them to calculate the properties of the physical systems which they are studying.

The techniques of statistical mechanics have been used in attacking a wide variety of physical problems. A quick glance through this text will show that statistical mechanics has been applied to gases, liquids, solutions, electrolytic solutions, polymers, adsorption, metals, spectroscopy, transport theory, the helix-coil transition of DNA, the electrical properties of matter, and cell membranes, among others.

Statistical mechanics may be broadly classified into two parts, one dealing with systems in equilibrium and the other with systems not in equilibrium. The treatment of systems in equilibrium is usually referred to as *statistical thermodynamics*, since it forms a bridge between thermodynamics (often called classical thermodynamics) and molecular physics.

Thermodynamics provides us with mathematical relations between the various

2 INTRODUCTION AND REVIEW

experimental properties of macroscopic systems in equilibrium. An example of such a thermodynamic relation is that between the molar heat capacities at constant pressure and at constant volume,

$$C_p - C_v = \left[p + \left(\frac{\partial E}{\partial V} \right)_T \right] \left(\frac{\partial V}{\partial T} \right)_p \quad (1-1)$$

Another, and one that we shall use in the next chapter, is

$$\left(\frac{\partial E}{\partial V} \right)_{N,T} - T \left(\frac{\partial p}{\partial T} \right)_{N,V} = -p \quad (1-2)$$

Note that thermodynamics provides connections between many properties, but does not supply information concerning the magnitude of any one. Neither does it attempt to base any relation on molecular models or interpretations. This, in fact, is both the power and weakness of thermodynamics. It is a general discipline which does not need to recognize or rely upon the existence of atoms and molecules. Its many relations would remain valid even if matter were continuous. In addition, there are many systems (such as biological systems) which are too complicated to be described by an acceptable molecular theory, but here again the relations given by thermodynamics are exact. This great generality, however, is paid for by its inability to calculate physical properties separately or to supply physical interpretations of its equations. When one seeks a molecular theory which can do just this, one then enters the field of statistical thermodynamics. Thus thermodynamics and statistical thermodynamics treat the same systems. Thermodynamics provides general relations without the need of ever considering the ultimate constitution of matter, while statistical thermodynamics, on the other hand, assumes the existence of atoms and molecules to calculate and interpret thermodynamic quantities from a molecular point of view.

Statistical thermodynamics itself may be further divided into two areas: first, the study of systems of molecules in which molecular interactions may be neglected (such as dilute gases), and second, the study of systems in which the molecular interactions are of prime importance (such as liquids). We shall see that the neglect of intermolecular interactions enormously simplifies our problem. Chapters 4 through 14 of the book are devoted to the treatment of systems in which these interactions either may be ignored or highly simplified. This is the kind of statistical thermodynamics to which most undergraduates have been exposed and, to some extent, represents typical statistical thermodynamical research done in the 1930s. The more interesting and challenging problems, however, concern systems in which these molecular interactions cannot be neglected; Chapters 15 through 20 of the expanded volume, *Statistical Mechanics*, are devoted to the study of such systems. It is in this latter area that a great deal of the research of the 1940s, 1950s, and 1960s was carried out. There are, of course, many important problems of this sort still awaiting attack. The theory of concentrated electrolyte solutions and the proof for the existence of first-order phase transitions are just two examples.

The most difficult branch of statistical mechanics, both mathematically and conceptually, is the study of systems not in equilibrium. This field is often referred to as *nonequilibrium statistical mechanics*. This is presently a very active area of research.

There are still some important unsolved conceptual problems in nonequilibrium statistical mechanics. Nevertheless, in the 1950s great strides were made toward the establishment of a firm basis for nonequilibrium statistical mechanics, commensurate with that of equilibrium statistical mechanics, or what we have called statistical

thermodynamics. Chapters 21 through 29 of *Statistical Mechanics* present an introduction to some of the more elementary of these fairly new and useful concepts and techniques.

In Chapter 2 we shall introduce and discuss the basic concepts and assumptions of statistical thermodynamics. We shall present these ideas in terms of quantum mechanical properties such as energy states, wave functions, and degeneracy. Although it may appear at this point that quantum mechanics is a prerequisite for statistical thermodynamics, it will turn out that a satisfactory version of statistical thermodynamics can be presented by using only a few quantum mechanical ideas and results. We assume that the student is familiar with only the amount of quantum mechanics taught in most present-day physical chemistry courses. About the only requirement of the first few chapters is an understanding that the Schrödinger equation determines the possible energy values E_j available to the system and that these may have a degeneracy associated with them which we denote by $\Omega(E_j)$.

Before discussing the principles, however, we shall present in this chapter a discussion of some of the terms or concepts that are particularly useful in statistical thermodynamics. In Section 1-2 we shall treat classical mechanics, including an introduction to the Lagrangian and Hamiltonian formalisms. In Section 1-3 we shall briefly review the main features of quantum mechanics and give the solutions of the Schrödinger equation for some important systems. The only new material in this section to most students probably will be the discussion of the eigenvalues or energy levels of a many-body system. Then in Section 1-4 we shall review thermodynamics briefly, since it is assumed that the reader is familiar with the three laws of thermodynamics and the tedious manipulations of partial derivatives. Two important topics that are not usually discussed in elementary physical chemistry texts are introduced, however. These two topics are the Legendre transformation and Euler's theorem, both of which are useful in studying statistical thermodynamics. Finally, in Section 1-5 we shall discuss some mathematical techniques and results that are particularly useful in statistical thermodynamics. Much of this section may be new material to the reader.

1-2 CLASSICAL MECHANICS

NEWTONIAN APPROACH

Everyone knows the equation $F = ma$. What this equation really says is that the rate of change of momentum is equal to the applied force. If we denote the momentum by \mathbf{p} , we have then a more general version of Newton's second law, namely,

$$\frac{d\mathbf{p}}{dt} \equiv \dot{\mathbf{p}} = \mathbf{F} \quad (1-3)$$

If the mass is independent of time, then $d\mathbf{p}/dt = m d\mathbf{r}/dt = m\ddot{\mathbf{r}} = m\mathbf{a}$. If \mathbf{F} is given as a function of position $\mathbf{F}(x, y, z)$, then Eq. (1-3) represents a set of second-order differential equations in x , y , and z whose solutions give x , y , and z as a function of time if some initial conditions are known. Thus Eq. (1-3) is called an equation of motion. We shall consider three applications of this equation.

Example 1. Solve the equation of motion of a body of mass m shot vertically upward with an initial velocity v_0 in a gravitational field.

If we choose the x -axis (positive in the upward direction) to be the height of the body, then we have

$$m\ddot{x} = -mg$$

4 INTRODUCTION AND REVIEW

where mg is the magnitude of the force. The negative sign indicates that the force is acting in a downward direction. The solution to this differential equation for x is then

$$x(t) = -\frac{1}{2}gt^2 + v_0 t + x_0 \quad (1-4)$$

with x_0 equal to $x(0)$, which in our case is 0. This then gives the position of the body at any time after it was projected. The extension of this problem to two dimensions (i.e., a shell shot out of a cannon) and the inclusion of viscous drag on the body are discussed in Problems 1-1 and 1-2.

Example 2. Set up and solve the equation of motion of a simple harmonic oscillator.

Let x_0 be the length of the unstrained spring. Hooke's law says that the force on the mass attached to the end of the spring is $F = -k(x - x_0)$. If we let $\xi = x - x_0$, we can write

$$\frac{d^2\xi}{dt^2} + \frac{k}{m}\xi = 0 \quad (1-5)$$

whose solution is

$$\xi(t) = A \sin \omega t + B \cos \omega t \quad (1-6)$$

The quantity

$$\omega = (k/m)^{1/2} \quad (1-7)$$

is the natural vibrational frequency of the system. Equation (1-6) can be written in an alternative form (see Problem 1-5)

$$\xi(t) = C \sin(\omega t + \phi) \quad (1-8)$$

This shows more clearly that the mass undergoes simple harmonic motion with frequency ω . Problems 1-3 through 1-5 illustrate some of the basic features of simple harmonic motion.

Example 3. Two-dimensional motion of a body under coulombic attraction to a fixed center.

In this case the force is $F = -K/r^3$, that is, it is of magnitude $-K/r^2$ and directed radially. Newton's equations become

$$\begin{aligned} m\ddot{x} &= F_x = -\frac{Kx}{(x^2 + y^2)^{3/2}} \\ m\ddot{y} &= F_y = -\frac{Ky}{(x^2 + y^2)^{3/2}} \end{aligned} \quad (1-9)$$

Unlike our previous examples, these two equations are difficult to solve. Since the force depends, in a natural way, on the polar coordinates r and θ , it is more convenient for us to set the problem up in a polar coordinate system. Using then

$$x = r \cos \theta$$

$$y = r \sin \theta$$

and some straightforward differentiation, we get

$$\left\{ m(\ddot{r} - \dot{\theta}^2 r) + \frac{K}{r^2} \right\} \cos \theta - m(r\ddot{\theta} + 2\dot{r}\dot{\theta}) \sin \theta = 0 \quad (1-10a)$$

$$\left\{ m(\ddot{r} - \dot{\theta}^2 r) + \frac{K}{r^2} \right\} \sin \theta + m(r\ddot{\theta} + 2\dot{r}\dot{\theta}) \cos \theta = 0 \quad (1-10b)$$

By multiplying the first of these equations by $\cos \theta$ and the second by $\sin \theta$ and then adding the two, one gets

$$m(r\ddot{r} - \dot{\theta}^2 r) + \frac{K}{r^2} = 0 \quad (1-11)$$

But this is just the term in braces in Eq. (1-10), which leads us to the result that

$$m(r\ddot{\theta} + 2\dot{\theta}\dot{r}) = 0 \quad (1-12)$$

as well. Equation (1-12) can be written in the form

$$\frac{1}{r} \frac{d}{dt}(mr^2\dot{\theta}) = 0 \quad (1-13)$$

which implies that

$$mr^2\dot{\theta} = \text{constant} \quad (1-14)$$

This quantity, $mr^2\dot{\theta}$, which maintains a fixed value during the motion of the particle, is called the angular momentum of the particle and is denoted by l . The angular momentum is always conserved if the force is central, that is, directed along r (see Problem 1-10).

Equation (1-14) can be used to eliminate $\dot{\theta}$ from Eq. (1-11) to give an equation in r alone, called the radial equation:

$$m\ddot{r} - \frac{l^2}{mr^3} + \frac{K}{r^2} = 0 \quad (1-15)$$

This equation can be solved (at least numerically) to give $r(t)$, which together with Eq. (1-14) gives $\theta(t)$.

Even though the solution in this example is somewhat involved using polar coordinates, it is nevertheless much easier than if we had used Cartesian coordinates. This is just one example of many possibilities, which show that it is advantageous to recognize the symmetry of the problem by using the appropriate coordinate system.

This example was introduced, however, to illustrate another important point. Notice that Eq. (1-15) for r can be written as a Newtonian equation (i.e., in the form $F = ma$)

$$m\ddot{r} = -\frac{K}{r^2} + \frac{l^2}{mr^3}$$

if we interpret the term l^2/mr^3 as a force. This force is the well-known centrifugal force and must be introduced into the equation for $m\ddot{r}$.

This constitutes the main disadvantage of the Newtonian approach. The form of the equation $m\ddot{\eta} = F_\eta$ (where η is some general coordinate) is useful only in Cartesian systems, unless we are prepared to define additional forces, such as the centrifugal force in the above example. At times these necessary additional forces are fairly obscure.

There exist more convenient formulations of classical mechanics which are not tied to any one coordinate system. The two formulations that we are about to introduce are, in fact, independent of the coordinate system employed. These are the Lagrangian and the Hamiltonian formulations.

6 INTRODUCTION AND REVIEW

LAGRANGIAN APPROACH

Let K be the kinetic energy of a particle. In Cartesian coordinates

$$K(\dot{x}, \dot{y}, \dot{z}) = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$$

Let the potential energy be U . In many problems U is a function of position only, and so we write $U(x, y, z)$. Newton's equations are

$$m\ddot{x} = -\frac{\partial U}{\partial x}$$

with similar equations for y and z . Now introduce a new function

$$L(x, y, z, \dot{x}, \dot{y}, \dot{z}) \equiv K(\dot{x}, \dot{y}, \dot{z}) - U(x, y, z)$$

This function is called the Lagrangian of the system. In terms of L , we have

$$\frac{\partial L}{\partial \dot{x}} = \frac{\partial K}{\partial \dot{x}} = m\ddot{x}$$

$$\frac{\partial L}{\partial x} = -\frac{\partial U}{\partial x}$$

...

and we can write Newton's equations in the form

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = \frac{\partial L}{\partial x} \quad (1-16)$$

with similar equations for y and z . These are Lagrange's equations of motion in Cartesian coordinates. The remarkable and useful property of Lagrange's equations is that they have the same form in any coordinate system. If the x, y, z are transformed into any other system, say q_1, q_2, q_3 , Lagrange's equations take the form

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) = \frac{\partial L}{\partial q_j} \quad j = 1, 2, 3 \quad (1-17)$$

This can be proved by writing $x = x(q_1, q_2, q_3)$, $y = y(q_1, q_2, q_3)$, and $z = z(q_1, q_2, q_3)$ and then transforming Eq. (1-16) into Eq. (1-17). (See Problem 1-13.)

Lagrange's equations are more useful than Newton's equations in many problems because it is usually much easier to write down an expression for the potential energy in some appropriate coordinate system than it is to recognize all the various forces. The Lagrangian formalism is based on the potential energy of the system, whereas the Newtonian approach is based on the forces acting on the system.

To illustrate the utility of the Lagrangian approach, we shall redo Example 3, the two-dimensional motion of a particle in a coulombic force field.

Example 3'. The kinetic energy is

$$K = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2)$$

and the potential energy is $U = -K/r$. The Lagrangian, then, is

$$L(r, \theta, \dot{r}, \dot{\theta}) = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2) + \frac{K}{r} \quad (1-18)$$

The two Lagrangian equations of motion are

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

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

or using Eq. (1-18) for L ,

$$\frac{d}{dt}(m\dot{r}) = mr\dot{\theta}^2 - \frac{K}{r^2}$$

$$\frac{d}{dt}(mr^2\dot{\theta}) = 0$$

These two equations are just Eqs. (1-11) and (1-13). Note, however, that they were obtained in a much more straightforward manner than were Eqs. (1-11) and (1-13). Problems 1-10 through 1-12 further illustrate the utility of the Lagrangian formulation. Other problems involve the motion of one and two particles in central force fields.

Equations (1-17) are three second-order ordinary differential equations. To completely specify the solutions, we need three initial velocities $\dot{q}_1(0)$, $\dot{q}_2(0)$, $\dot{q}_3(0)$, and three initial positions $q_1(0)$, $q_2(0)$, $q_3(0)$. These six initial conditions along with Lagrange's equations completely determine the future (and past) trajectory of the system. If there were N particles in the system, there would be $3N$ Lagrange equations and $6N$ initial conditions.

There is another formulation of classical mechanics that involves $6N$ first-order differential equations. Although this formulation is not as convenient as Lagrange's for solving problems, it is more convenient from a theoretical point of view, particularly in quantum mechanics and statistical mechanics. This is the Hamiltonian formulation.

HAMILTONIAN APPROACH

We define a generalized momentum by

$$p_j = \frac{\partial L}{\partial \dot{q}_j} \quad j = 1, 2, \dots, 3N \quad (1-19)$$

This generalized momentum is said to be *conjugate* to q_j . Note that Eq. (1-19) is simply $p_x = m\dot{x}$, and so on, in Cartesian coordinates.

We now define the Hamiltonian function for a system containing just one particle (for simplicity) by

$$H(p_1, p_2, p_3, q_1, q_2, q_3) = \sum_{j=1}^3 p_j \dot{q}_j - L(\dot{q}_1, \dot{q}_2, \dot{q}_3, q_1, q_2, q_3) \quad (1-20)$$

It is understood here that the \dot{q}_j 's have been eliminated in favor of the p_j 's by means of Eq. (1-19).

An important difference between the Lagrangian approach and the Hamiltonian approach is that the Lagrangian is considered to be a function of the generalized velocities \dot{q}_j and the generalized coordinates q_j , whereas the Hamiltonian is considered to be a function of the generalized momenta p_j and the conjugate generalized coordinates q_j . This may appear to be a fine distinction at this point, but it will turn

8 INTRODUCTION AND REVIEW

out to be important later on. It also may seem, at this time, that the definition Eq. (1-20) is rather obscure, but we shall give a motivation for its form in Section 1-4. (See Problem 1-38.)

For the kinds of systems that we shall treat in this book, the kinetic energy is of the form

$$K = \sum_{j=1}^{3N} a_j(q_1, q_2, \dots, q_{3N}) \dot{q}_j^2 \quad (1-21)$$

that is, a quadratic function of the generalized velocities. The coefficients a_j are, in general, functions of generalized coordinates but not an explicit function of time. If, furthermore, the potential energy is a function only of the generalized coordinates, then the p_j occurring in Eq. (1-20) are given by

$$p_j = \frac{\partial L}{\partial \dot{q}_j} = \frac{\partial K}{\partial \dot{q}_j} = 2a_j \dot{q}_j$$

where the last equality comes from Eq. (1-21). Substituting this into Eq. (1-20) gives the important result

$$H = K + U = \text{total energy} \quad (1-22)$$

We shall now show that if the Lagrangian is not an explicit function of time, then $dH/dt = 0$. We begin with the definition of H , that is, Eq. (1-20).

$$dH = \sum_j \dot{q}_j dp_j + \sum_j p_j d\dot{q}_j - \sum_j \frac{\partial L}{\partial \dot{q}_j} d\dot{q}_j - \sum_j \frac{\partial L}{\partial q_j} dq_j$$

But if we use Eqs. (1-17) and (1-19), we see that

$$dH = \sum \dot{q}_j dp_j - \sum \dot{p}_j dq_j \quad (1-23)$$

The total derivative of H is (assuming no explicit dependence on time)

$$dH = \sum \left(\frac{\partial H}{\partial p_j} \right) dp_j + \sum \left(\frac{\partial H}{\partial q_j} \right) dq_j \quad (1-24)$$

Comparing Eqs. (1-23) and (1-24), we get Hamilton's equations of motion:

$$\frac{\partial H}{\partial p_j} = \dot{q}_j \quad \frac{\partial H}{\partial q_j} = -\dot{p}_j \quad j = 1, 2, \dots, 3N \quad (1-25)$$

Hamilton's equations are $6N$ first-order differential equations. It is easy to show from Eqs. (1-24) and (1-25) that $dH/dt = 0$. (See Problem 1-14.) This along with Eq. (1-22) says that energy is conserved in such systems.

Since the Hamiltonian is so closely related to the energy, and it is the total energy which is usually the prime quantity in quantum and statistical mechanics, the Hamiltonian formalism will turn out to be the most useful from a conceptual point of view. Fortunately, however, we shall never have to solve the equations of motion for macroscopic systems. The role of statistical mechanics is to avoid doing just that.

1-3 QUANTUM MECHANICS

In the previous section we have seen that a knowledge of the initial velocities and coordinates of a particle or a system of particles was sufficient to determine the future course of the system if the equations of motion, essentially the potential field that the

on
-4.
the
21)
in
If,
es,
ves
22)
ien
23)
24)
om
22)
rgy
mil-
iew.
cro-
and
ure
the

system experiences, are known. If the state (its velocities and coordinates) of the system is known at time t_0 , then classical mechanics provides us with a method of calculating the state of the system at any other time t_1 .

By the 1920s it was realized that such a calculation was too detailed in principle. The Heisenberg uncertainty principle states that it is impossible to precisely specify both the momentum and position of a particle simultaneously. Consequently, the prescription given by classical mechanics had to be modified to include the principle of uncertainty. This modification resulted in the development of quantum mechanics.

There are a number of levels of introducing the central ideas of quantum mechanics, but for most of the material in this text, we need consider only the most elementary. A fundamental concept of quantum mechanics is the so-called wave function $\Psi(\mathbf{q}, t)$, where \mathbf{q} represents the set of coordinates necessary to describe the system. The wave function is given the physical interpretation that the probability that at time t the system is found between q_1 and $q_1 + dq_1$, q_2 and $q_2 + dq_2$, and so on, is

$$\Psi^*(\mathbf{q}, t)\Psi(\mathbf{q}, t) dq_1 dq_2 \cdots dq_{3N}$$

We shall often write $dq_1 \cdots dq_{3N}$ as $d\mathbf{q}$. The uncertainty principle dictates that $\Psi(\mathbf{q}, t)$ is the most complete description of the system that can be obtained. Since the system is sure to be somewhere, we have

$$\int \Psi^*(\mathbf{q}, t)\Psi(\mathbf{q}, t) d\mathbf{q} = 1 \quad (1-26)$$

If Eq. (1-26) is satisfied, Ψ is said to be normalized.

A central problem of quantum mechanics is the calculation of $\Psi(\mathbf{q}, t)$ for any system of interest. We denote the time-independent part of $\Psi(\mathbf{q}, t)$ by $\psi(\mathbf{q})$. The state of the system described by a particular $\psi(\mathbf{q})$ is said to be a stationary state. Throughout this book we shall deal with stationary states only.

For our purpose, the wave function ψ is given as the solution of the Schrödinger equation

$$\mathcal{H}\psi = E\psi \quad (1-27)$$

where \mathcal{H} is the Hamiltonian operator, and E is a scalar quantity corresponding to the energy of the system. The Hamiltonian operator is

$$\begin{aligned} \mathcal{H} &= -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) + U(x, y, z) \\ &= -\frac{\hbar^2}{2m} \nabla^2 + U(x, y, z) \end{aligned} \quad (1-28)$$

where \hbar is $h/2\pi$, that is, Planck's constant divided by 2π . The first term here corresponds to the kinetic energy, and the second term is the potential energy. The Hamiltonian operator, then, corresponds to the total energy. There is a quantum mechanical operator and an equation similar to Eq. (1-27) corresponding to every quantity of classical mechanics, but we shall need only the one for the energy, namely, the Schrödinger equation.

Given certain physical boundary conditions of the system, a knowledge of \mathcal{H} alone is sufficient to determine ψ and E . The wave function ψ is called an eigenfunction of the operator \mathcal{H} , and E is called an eigenvalue. There will usually be many ψ 's and E 's that satisfy Eq. (1-28), and this is indicated by labeling ψ and E with one or more subscripts. Generally, then, we have

$$\mathcal{H}\psi_j = E_j \psi_j \quad (1-29)$$

10 INTRODUCTION AND REVIEW

Equation (1-29) is a partial differential equation for ψ_j . The application of the boundary conditions often limits the values of E_j to only certain discrete values. Some simple examples are

1. a particle in a one-dimensional infinite well:

$$\mathcal{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2}$$

$$\varepsilon_n = \frac{\hbar^2 n^2}{8ma^2} \quad n = 1, 2, \dots \quad (1-30)$$

2. a simple harmonic oscillator:

$$\mathcal{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + \frac{1}{2} kx^2$$

$$\varepsilon_n = (n + \frac{1}{2})\hbar\omega \quad n = 0, 1, 2, \dots \quad (1-31)$$

where $\omega = (k/m)^{1/2}$.

3. a rigid rotor (see Problem 1-21):

$$\mathcal{H} = -\frac{\hbar^2}{2I} \left\{ \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right\}$$

$$\varepsilon_J = \frac{J(J+1)\hbar^2}{2I} \quad J = 0, 1, 2, \dots \quad (1-32)$$

Here I is the moment of inertia of the rotor (see Problem 1-15 for a treatment of the classical counterpart of this system).

The rigid rotor illustrates another important concept of quantum mechanics, namely, that of degeneracy. It happens that there may be a number of eigenfunctions or states of the system having the same eigenvalue or energy. The number of eigenfunctions having this energy is called the degeneracy of the system. For the rigid rotor, the degeneracy, ω_J , is $2J + 1$. The particle in a one-dimensional infinite well and the simple harmonic oscillator are nondegenerate, that is, the ω_n are unity. The concept of energy states and degeneracy plays an important role in statistical thermodynamics.

Consider the energy states of a particle in a three-dimensional infinite well. These are given by

$$\varepsilon_{n_x n_y n_z} = \frac{\hbar^2}{8ma^2} (n_x^2 + n_y^2 + n_z^2) \quad n_x, n_y, n_z = 1, 2, 3, \dots \quad (1-33)$$

The degeneracy is given by the number of ways that the integer $M = 8ma^2\varepsilon/\hbar^2$ can be written as the sum of the squares of three positive integers. In general, this is an erratic and discontinuous function of M (the number of ways will be zero for many values of M), but it becomes smooth for large M , and it is possible to derive a simple expression for it. Consider a three-dimensional space spanned by n_x , n_y , and n_z . There is a one-to-one correspondence between energy states given by Eq. (1-33) and the points in this n_x , n_y , n_z space with coordinates given by positive integers. Figure 1-1 shows a two-dimensional version of this space. Equation (1-33) is an equation for a sphere of radius $R = (8ma^2\varepsilon/\hbar^2)^{1/2}$ in this space

$$n_x^2 + n_y^2 + n_z^2 = \frac{8ma^2\varepsilon}{\hbar^2} = R^2$$

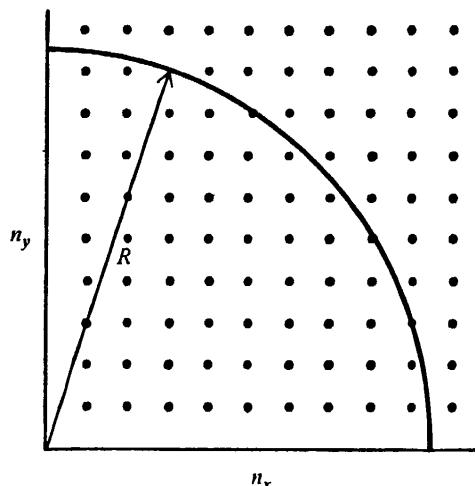


Figure 1-1. A two-dimensional version of the (n_x, n_y, n_z) space, the space with the quantum numbers n_x , n_y , and n_z as axes.

We wish to calculate the number of lattice points that are at some fixed distance from the origin in this space. In general, this is very difficult, but for large R we can proceed as follows. We treat R or ϵ as a continuous variable and ask for the number of lattice points between ϵ and $\epsilon + \Delta\epsilon$. To calculate this quantity, it is convenient to first calculate the number of lattice points consistent with an energy $\leq \epsilon$. For large ϵ , it is an excellent approximation to equate the number of lattice points consistent with an energy $\leq \epsilon$ with the volume of one octant of a sphere of radius R . We take only one octant, because n_x , n_y , and n_z are restricted to be positive integers. If we denote the number of such states by $\Phi(\epsilon)$, we can write

$$\Phi(\epsilon) = \frac{1}{8} \left(\frac{4\pi R^3}{3} \right) = \frac{\pi}{6} \left(\frac{8ma^2\epsilon}{h^2} \right)^{3/2} \quad (1-34)$$

The number of states between ϵ and $\epsilon + \Delta\epsilon$ ($\Delta\epsilon/\epsilon \ll 1$) is

$$\begin{aligned} \omega(\epsilon, \Delta\epsilon) &= \Phi(\epsilon + \Delta\epsilon) - \Phi(\epsilon) \\ &= \frac{\pi}{4} \left(\frac{8ma^2}{h^2} \right)^{3/2} \epsilon^{1/2} \Delta\epsilon + O((\Delta\epsilon)^2) \end{aligned} \quad (1-35)$$

If we take $\epsilon = 3kT/2$, $T = 300^\circ\text{K}$, $m = 10^{-22}$ g, $a = 10$ cm, and $\Delta\epsilon$ to be 0.01ϵ (in other words a 1% band around ϵ), then $\omega(\epsilon, \Delta\epsilon)$ is $O(10^{28})$.* So even for a system as simple as a particle-in-a-box, the degeneracy can be very large at room temperature.

For an N -particle system, the degeneracy is tremendously greater than $O(10^{28})$. To see this, consider a system of N noninteracting particles in a cube. The energy of this system is

$$E = \frac{\hbar^2}{8ma^2} \sum_{j=1}^N (n_{xj}^2 + n_{yj}^2 + n_{zj}^2) = \frac{\hbar^2}{8ma^2} \sum_{j=1}^{3N} s_j^2$$

where n_{xj} , n_{yj} , n_{zj} , and s_j are positive integers. The degeneracy of this system can be calculated by generalizing the above derivation for one particle. Using the volume of

* We use the notation $O(10^{28})$ say, to mean of the order of magnitude 10^{28} . This differs from standard mathematical notation, but there should be no confusion.

12 INTRODUCTION AND REVIEW

an N -dimensional sphere from Problem 1-24, the number of states with energy $\leq E$ is*

$$\Phi(E) = \frac{1}{\Gamma(N+1)\Gamma[(3N/2)+1]} \left(\frac{2\pi m a^2 E}{\hbar^2} \right)^{3N/2} \quad (1-36)$$

where $\Gamma(n)$ here is the gamma function. (See Problem 1-58.) The number of states between E and $E + \Delta E$ is

$$\Omega(E, \Delta E) = \frac{1}{\Gamma(N+1)\Gamma(3N/2)} \left(\frac{2\pi m a^2}{\hbar^2} \right)^{3N/2} E^{(3N/2-1)} \Delta E \quad (1-37)$$

In this case, $E = 3NkT/2$. If we take $T = 300^\circ\text{K}$, $m = 10^{-22}$ g, $a = 10$ cm, $N = 6.02 \times 10^{23}$, and ΔE equal to $0.01E$, we get $\Omega(E, \Delta E)$ to be $0(10^N)$ (see Problem 1-23), an extremely large number. This shows that as the number of particles in the system increases, the quantum mechanical degeneracy becomes enormous. Although we have shown this only for a system of noninteracting particles confined to a cubical box, that is, an ideal gas, the result is generally true. We shall see in the next chapter that the concept of the degeneracy of a macroscopic system is very important.

There is another quantum mechanical result that we shall use later on. It often happens that the Hamiltonian of a many-body system can be written either exactly or approximately as a summation of one-particle or few-particle Hamiltonians, that is,

$$\mathcal{H} = \mathcal{H}_\alpha + \mathcal{H}_\beta + \mathcal{H}_\gamma + \dots \quad (1-38)$$

Let the eigenvalues of \mathcal{H}_j be ε_j , and the eigenfunctions be ψ_j , where $j = \alpha, \beta, \gamma, \dots$. To solve the many-body Schrödinger equation, we let $\psi = \psi_\alpha \psi_\beta \psi_\gamma \dots$. Then

$$\begin{aligned} \mathcal{H}\psi &= (\mathcal{H}_\alpha + \mathcal{H}_\beta + \mathcal{H}_\gamma + \dots) \psi_\alpha \psi_\beta \psi_\gamma \dots \\ &= \psi_\beta \psi_\gamma \dots \mathcal{H}_\alpha \psi_\alpha + \psi_\alpha \psi_\gamma \dots \mathcal{H}_\beta \psi_\beta + \dots \\ &= \psi_\beta \psi_\gamma \varepsilon_\alpha \psi_\alpha + \psi_\alpha \psi_\gamma \varepsilon_\beta \psi_\beta + \dots \\ &= (\varepsilon_\alpha + \varepsilon_\beta + \dots) \psi_\alpha \psi_\beta \psi_\gamma \dots = E\psi \end{aligned} \quad (1-39)$$

In other words, the energy of the entire system is the sum of the energies of the individual particles if they do not interact. This is a very important result and will allow us to reduce a many-body problem to a one-body problem if the interactions are weak enough to ignore, such as in the case of a dilute gas. We shall see a number of cases where, even though the interactions are too strong to be ignored (such as in a solid), it is possible to formally or mathematically write the Hamiltonian in the form of Eq. (1-38). This will lead to defining quasi-particles like phonons and photons.

The last quantum mechanical topic we shall discuss here is that of the symmetry of wave functions with respect to the interchange of identical particles. Consider a system of N identical particles, described by a wave function $\psi(1, 2, 3, \dots, N)$, where 1 denotes the coordinates of particle 1, and so on. If we interchange the position of any two of the particles, say particles 1 and 2, the wave function must either remain the same or change sign. (See Problem 1-26.) Thus if we let P_{12} be an operator that exchanges the two identical particles 1 and 2, then

$$\begin{aligned} P_{12}\psi(1, 2, 3, \dots, N) &= \psi(2, 1, 3, \dots, N) \\ &= \pm \psi(1, 2, 3, \dots, N) \end{aligned} \quad (1-40)$$

* The extra factor of $\Gamma(N+1)$ occurs here because of the indistinguishability of the N particles. This will be discussed fully in Chapter 4.

is*
36)
.tes
37)
:m,
em
the
igh
ical
ster
ten
/or
: is,
38)
...
-39)
ndi-
llow
weak
ases
), it
n of
ty of
te 1
any
the
hat
0)

It turns out that whether the wave function remains the same or changes sign is a function of the nature of the two identical particles that are exchanged. For particles with an integral spin (such as the He-4 nucleus, photons, ...), the wave function remains the same. In this case the wave function is called symmetric, and such particles are called bosons. For particles with half-integral spin (such as electrons, ...), the wave function is called antisymmetric, and the particles are called fermions. Chapter 4 considers the consequences of this symmetry requirement of wave functions.

1-4 THERMODYNAMICS

In this section we shall not attempt to review thermodynamics, but shall simply state the three laws and briefly discuss their consequences. Problems 1-27 through 1-36 review some of the equations and manipulations that arise in thermodynamics. Two topics that are not often treated in elementary physical chemistry are presented here, namely, Legendre transformations and Euler's theorem. Both of these topics will be used later on.

The pressure-volume work done by a system on its surroundings in going from state *A* to state *B* is

$$w = \int_A^B p \, dV$$

where *p* is the pressure exerted by the surroundings on the system. The differential quantity δw is positive if dV is positive.

The heat absorbed by the system from the surroundings during the change of the system from state *A* to state *B* is

$$q = \int_A^B \delta q$$

The first law of thermodynamics states that even though *w* and *q* depend upon the path taken from *A* to *B*, their difference does not. Their difference, then, is a function only of the two states *A* and *B*, or, namely, is a state function. This function is called the internal energy or thermodynamic energy and is denoted by *E*.

The first law of thermodynamics is

$$\Delta E = E_B - E_A = q - w$$

$$= \int_A^B \delta q - \int_A^B p \, dV \quad (1-41)$$

For simplicity, we consider only *p-V* work.

A reversible change is one in which the driving force (a difference in pressure, a difference in temperature, and so on) is infinitesimal. Any other change is called irreversible or spontaneous. Problem 1-27 asks the reader to show that for an isothermal process, $w_{\text{rev}} > w_{\text{irrev}}$ and $q_{\text{rev}} > q_{\text{irrev}}$.

The first law of thermodynamics is nothing but a statement of the law of conservation of energy. The second law is somewhat more abstract and can be stated in a number of equivalent ways. One of them is: There is a quantity *S*, called entropy, which is a state function. In an irreversible process, the entropy of the system and its surroundings increases. In a reversible process, the entropy of the system and its surroundings

14 INTRODUCTION AND REVIEW

remains constant. The entropy of the system and its surroundings never decreases. The system and its surroundings are often referred to as the universe.

The mathematical expression for the difference in entropy between states *A* and *B* of a system is given by

$$\Delta S = \int_A^B \frac{dq_{rev}}{T} \quad (1-42)$$

Note that the heat appearing here is that associated with a reversible process. To compute ΔS between two states *A* and *B*, we must take the system from *A* to *B* in a reversible manner.

Another statement of the second law is: Along any reversible path, there exists an integrating factor *T*, common to all systems such that

$$dS = \frac{dq_{rev}}{T} \quad (1-43)$$

is an exact differential, that is, that *S* is a state function. Thus

$$\Delta S = \int_A^B \frac{dq_{rev}}{T}$$

For all other processes

$$\Delta S > \int_A^B \frac{dq}{T}$$

where *T* is the temperature of the surroundings.

The third law of thermodynamics states: If the entropy of each element in some crystalline state be taken as zero at the absolute zero of temperature, every substance has a finite positive entropy, but at the absolute zero of temperature, the entropy may become zero, and does become so in the case of perfect crystalline substances.

The second law is concerned with only the difference in the entropy between two states. The third law allows us to calculate the absolute entropy of a substance by means of the expressions

$$S - S_0 = \int_0^T \frac{dq_{rev}}{T} \quad \text{and} \quad S_0 = 0 \quad (1-44)$$

Problem 1-36 asks you to calculate the absolute entropy of gaseous nitromethane at its boiling point.

For simple one-component systems, the first law can be written in the form

$$dE = T dS - p dV \quad (1-45)$$

This implies that

$$\left(\frac{\partial E}{\partial S}\right)_V = T \quad \text{and} \quad \left(\frac{\partial E}{\partial V}\right)_S = -p \quad (1-46)$$

The simplicity of these partial derivatives implies that *E* is a "natural" function of *S* and *V*. For example, if we were to consider *E* to be a function of *V* and *T* (see Problem 1-30), we would get

$$dE = \left[T \left(\frac{\partial p}{\partial T} \right)_V - p \right] dV + C_V dT$$

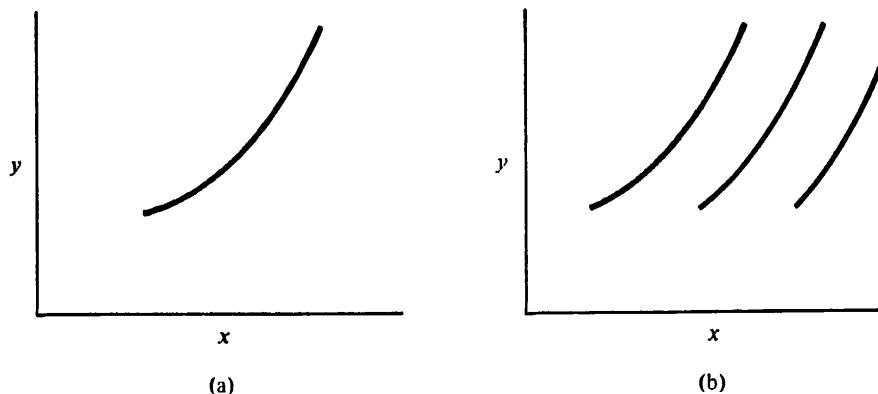


Figure 1-2. (a) shows the function $y(x)$ and (b) shows a family of functions, all of which give the same value of y for any fixed value of p .

Note that in this case the coefficients of dV and dT are not as simple as the coefficients of dV and dS obtained when E is expressed as a function of S and V . The "simplicity" of the expression $dE = T dS - p dV$ suggests that S and V are the "natural" variables for E . The quantities S and V (especially S) are difficult to control in the laboratory and consequently are not always the most desirable independent variables. A more useful pair might be (T, V) or (T, p) . An important question that arises, then, is the existence of other thermodynamic state functions whose natural variables are (T, V) or (T, p) , and so on. Furthermore, how would one find them if they do exist. This leads us to the topic of Legendre transformations.

We shall discuss a function of one variable in some detail and then simply present the generalization to a function of many variables. Consider a function $y = y(x)$, and let its slope be $p = p(x)$. We wish to describe the function $y(x)$ in terms of its slope. Figure 1-2, however, shows that the slope alone is not sufficient to completely specify $y(x)$. Figure 1-2(a) shows the curve $y(x)$, and Fig. 1-2(b) shows a family of curves, all of which give the same value of y for any one value of p . In order to uniquely describe the curve in Fig. 1-2(a), we must select one member of the family of curves in Fig. 1-2(b). We do this by specifying the intercepts of the tangent lines with the y -axis. Let the intercept be $\phi(p)$. Instead of describing the curve in Fig. 1-2(a) by y versus x , then, we can equally well represent it by specifying the slope at each point along with the intercept of the slope with the y -axis. Figure 1-3 shows these two representations. One sees that either representation can be used to describe the function. The relation between the two representations can be obtained by referring to Fig. 1-4. This figure shows that the slope p at any point is given by

$$p = \frac{y - \phi}{x - 0}$$

The result that we are after is

$$\phi(p) = y - px \quad (1-47)$$

The function $\phi(p)$ is the Legendre transformation of y . It is completely equivalent to $y(x)$, but considers p to be the independent variable instead of x . This may not be clear from the notation in Eq. (1-47), but it is understood there that y and x have been eliminated in favor of p by using the equations $y = y(x)$ and $x = x(p)$.

Let us apply this to the thermodynamic energy $E(S, V)$. We seek a function of T and V that is completely equivalent to E . Equation (1-46) shows that $T = (\partial E / \partial S)_V$,

16 INTRODUCTION AND REVIEW

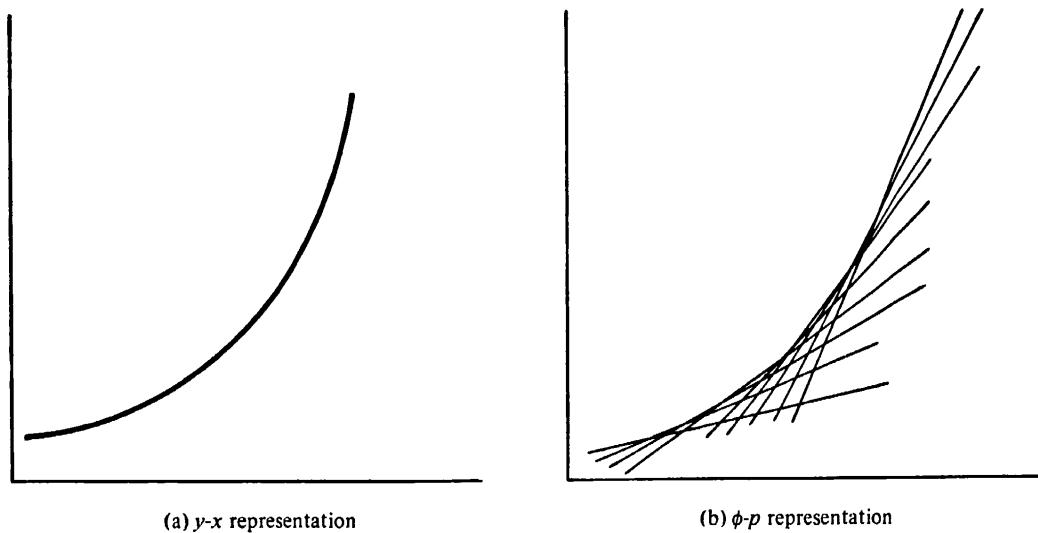


Figure 1-3. In (a), the function is represented by the locus of points, y versus x . In (b), the same function is given by the envelope of its tangent curves.

and so we are in a position to apply Eq. (1-47) directly. This can be treated as a one-variable problem, since V is held fixed throughout. Therefore the Legendre transformation of E that considers T and V to be the independent variables is $E - TS$. Of course, this is the Helmholtz free energy

$$A(T, V) = E - TS \quad (1-48)$$

whose differential form is

$$dA = -S dT - p dV \quad (1-49)$$

This shows that the natural variables of A are T and V . Another motivation for saying this is that the condition for equilibrium at constant T and V is that A assume its minimum value, or that $\Delta A \leq 0$ for a spontaneous process at constant T and V . To prove this, write

$$\begin{aligned} dA &= dE - T dS - S dT \\ &= \delta q - p dV - \delta q_{\text{rev}} - S dT \\ &= \delta q - \delta q_{\text{rev}} \end{aligned} \quad (1-50)$$

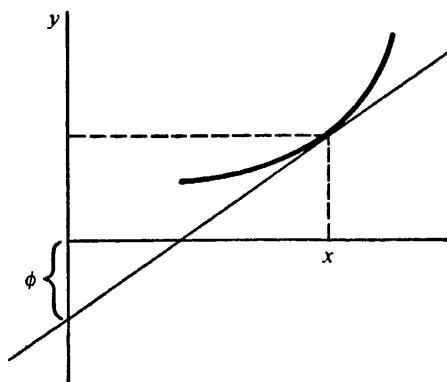


Figure 1-4. The diagram used to derive the connection between the y - x representation and the ϕ - p representation.

at constant T and V . But $\delta q \leq \delta q_{rev}$ (see Problem 1-27), and so $\Delta A \leq 0$ at constant T and V .

In elementary physical chemistry, the function $A = E - TS$ is often presented as an a priori definition. But it should be apparent now that this form is dictated by the Legendre transformation if one specifies T and V to be the independent variables.

A function, whose natural variables are S and p , can be obtained in the same manner. Equation (1-46) shows that $p = -(\partial E / \partial V)_S$, so Eq. (1-47) gives that $E + pV$ is a thermodynamic state function, whose natural variables are S and p . This function is, of course, the enthalpy.

The generalization of Eq. (1-47) to more than one variable is simply

$$\phi(p) = y - \sum_j p_j x_j \quad (1-51)$$

where the x_j 's are the independent variables of y , and $p_j = (\partial y / \partial x_j)$. We can use Eq. (1-51) to construct a thermodynamic state function, whose natural variables are T and p . Using Eqs. (1-46) and Eq. (1-51), we see that such a function is $E - TS + pV$, the Gibb's free energy. Its differential form $dG = -S dT + V dp$ and the fact that $\Delta G \leq 0$ for a spontaneous change at constant T and p suggest that T and p are the natural variables of G .

Up to this point we have considered only closed one-component systems. In general, E , H , A , and G depend upon the number of moles or molecules of each component. If we let N_j be the number of moles of component j , we have

$$dE = T dS - p dV + \sum_j \left(\frac{\partial E}{\partial N_j} \right)_{S, V, N_k, j \neq k} dN_j \quad (1-52)$$

$$= T dS - p dV + \sum_j \mu_j dN_j \quad (1-53)$$

where the second line defines μ_j . By adding $d(pV)$ to both sides of Eq. (1-53), we get

$$dH = T dS + V dp + \sum_j \mu_j dN_j \quad (1-54)$$

If we subtract $d(TS)$ from both sides of Eq. (1-53), we get

$$dA = -S dT - p dV + \sum_j \mu_j dN_j \quad (1-55)$$

Similar manipulations give

$$dG = -S dT + V dp + \sum_j \mu_j dN_j \quad (1-56)$$

Equations (1-52) through (1-56) show that

$$\mu_j = \left(\frac{\partial E}{\partial N_j} \right)_{S, V, \dots} = \left(\frac{\partial H}{\partial N_j} \right)_{S, p, \dots} = \left(\frac{\partial A}{\partial N_j} \right)_{V, T, \dots} = \left(\frac{\partial G}{\partial N_j} \right)_{p, T, \dots} \quad (1-57)$$

The quantity μ_j is called the chemical potential.

There is a mathematical theorem, called Euler's theorem, which is very useful in thermodynamics. Before discussing Euler's theorem, however, we must define extensive and intensive variables. Extensive properties are additive; their value for the whole system is equal to the sum of their values for the individual parts. Examples are the volume, mass, and entropy. Intensive properties are not additive. Examples are temperature and pressure. The temperature of any small part of a system in equilibrium is the same as the temperature of the whole system. Euler's theorem deals with

18 INTRODUCTION AND REVIEW

extensive and intensive variables. If

$$f(\lambda x_1, \lambda x_2, \dots, \lambda x_N) = \lambda^n f(x_1, x_2, \dots, x_N) \quad (1-58)$$

f is said to be a homogeneous function of order n . The functions $f(x) = 3x^2$ and $f(x, y, z) = xy^2 + z^3 - 6x^4/y$ are homogeneous functions of degree 2 and 3, respectively, whereas $f(x) = x^2 + 2x - 3$ and $f(x, y) = xy - e^{xy}$ are not homogeneous. Euler's theorem states that if $f(x_1, \dots, x_N)$ is a homogeneous function of order n , then

$$nf(x_1, \dots, x_N) = x_1 \frac{\partial f}{\partial x_1} + x_2 \frac{\partial f}{\partial x_2} + \dots + x_N \frac{\partial f}{\partial x_N} \quad (1-59)$$

The proof of Euler's theorem is simple. Differentiate Eq. (1-58) with respect to λ :

$$\begin{aligned} n\lambda^{n-1}f(x_1, x_2, \dots, x_N) &= \left(\frac{\partial f}{\partial \lambda x_1} \right) \left(\frac{\partial \lambda x_1}{\partial \lambda} \right) + \left(\frac{\partial f}{\partial \lambda x_2} \right) \left(\frac{\partial \lambda x_2}{\partial \lambda} \right) + \dots + \left(\frac{\partial f}{\partial \lambda x_N} \right) \left(\frac{\partial \lambda x_N}{\partial \lambda} \right) \\ &= x_1 \left(\frac{\partial f}{\partial \lambda x_1} \right) + x_2 \left(\frac{\partial f}{\partial \lambda x_2} \right) + \dots + x_N \left(\frac{\partial f}{\partial \lambda x_N} \right) \end{aligned}$$

Euler's theorem is proved by letting $\lambda = 1$.

Extensive thermodynamic variables are homogeneous of degree 1. Let us apply Euler's theorem to the Gibb's free energy.

$$G(T, p, \lambda N_1, \lambda N_2, \dots) = \lambda G(T, p, N_1, N_2, \dots)$$

The variables T and p here can be treated as constants. Equation (1-59) gives that

$$G = \sum_j N_j \left(\frac{\partial G}{\partial N_j} \right)_{T, p, \dots} = \sum_j N_j \mu_j \quad (1-60)$$

Taking the derivative of this at constant T and p ,

$$dG = \sum_j N_j d\mu_j + \sum_j \mu_j dN_j \quad (\text{constant } T \text{ and } p)$$

But using Eq. (1-56) at constant T and p , we have

$$\sum_j N_j d\mu_j = 0 \quad (\text{constant } T \text{ and } p) \quad (1-61)$$

This is called the Gibbs-Duhem equation and is very useful in the thermodynamic study of solutions. (See *Physical Chemistry*, 4th ed., by W. J. Moore, p. 235, under "Additional Reading," for a simple application of the Gibbs-Duhem equation.)

We shall conclude this section on thermodynamics with a brief discussion of the application of thermodynamics to chemical equilibria. Consider the general reaction



The capital letters represent the formulas of the compounds, and the v_j represent stoichiometric coefficients. It is more convenient to write Eq. (1-62) mathematically as

$$v_D D + v_E E + \dots - v_A A - v_B B - \dots = 0 \quad (1-63)$$

Define the extent of reaction λ , such that $dN_j = v_j d\lambda$ for all j , where the v 's for products are positive, and those for reactants are negative.

At constant T and p , we have

$$dG = \sum_j \mu_j dN_j = \left(\sum_j \mu_j v_j \right) d\lambda \quad (\text{constant } T \text{ and } p)$$

At equilibrium, G must be a minimum with respect to λ , so we write

$$58) \quad \sum_j \mu_j v_j = v_D \mu_D + v_E \mu_E + \cdots - v_A \mu_A - v_B \mu_B - \cdots = 0 \quad (1-64)$$

at equilibrium. The equilibrium between phases can be considered to be a chemical reaction of the form $A \rightleftharpoons B$, and so Eq. (1-64) gives that $\mu_A = \mu_B$ from the equilibrium condition between two pure phases.

Now consider the application of Eq. (1-64) to a chemical reaction between gases dilute enough to be considered ideal. Let the reaction be $v_A A + v_B B \rightleftharpoons v_C C + v_D D$. At constant temperature,

$$dG = V dp \quad (\text{constant } T)$$

and so

$$G - G^0 = \int_{p_0}^p V dp = \int_{p_0}^p \frac{NkT}{p} dp = NkT \ln \frac{p}{p_0} \quad (1-65)$$

In this equation G^0 is the standard free energy of the gas, the standard state being the gas at a pressure p_0 . Usually p_0 is taken to be 1 atmosphere. If we take N to be 1 mole, then G and G^0 become μ and μ^0 . Each component in the reactive gas mixture will have an equation of the form of Eq. (1-65), and so we have

$$\mu_j(T, p) = \mu_j^0(T) + RT \ln \frac{p_j}{p_{0j}} \quad (1-66)$$

The total free energy change is

$$\begin{aligned} \Delta\mu &= v_C \mu_C + v_D \mu_D - v_A \mu_A - v_B \mu_B \\ &= \Delta\mu^0 + RT \ln \frac{(p_C')^{v_C} (p_D')^{v_D}}{(p_A')^{v_A} (p_B')^{v_B}} \end{aligned} \quad (1-67)$$

In this equation the (p') 's are p/p_0 , that is, they are the pressures relative to the standard states. These (p') 's are unitless. The argument of the logarithm here has the form of an equilibrium constant, but is not equal to the equilibrium constant unless the pressures are those which exist at chemical equilibrium. Equation (1-67) gives the change in free energy of the conversion of reactants at arbitrary pressures to products at arbitrary pressures.

At equilibrium, $\Delta\mu = 0$, and we have

$$\begin{aligned} \Delta\mu^0 &= -RT \ln \left[\frac{(p_C')^{v_C} (p_D')^{v_D}}{(p_A')^{v_A} (p_B')^{v_B}} \right]_{\text{equilibrium}} \\ &= -RT \ln K_p \end{aligned} \quad (1-68)$$

There are extensive tabulations of μ^0 's, and so $\Delta\mu^0$ is a simple matter to calculate. We see that if $\Delta\mu^0 < 0$, then $K_p > 1$, that is, the conversion of reactants in their standard states to products in their standard states proceeds spontaneously. On the other hand, if $\Delta\mu^0 > 0$, then $K_p < 1$, and we conclude that the reaction does not proceed spontaneously. It should be understood, however, that this applies only to reactants and products in their standard states. In general, it is $\Delta\mu$ along with Eq. (1-67) that determines the extent of a chemical reaction. (See Problem 1-34.)

20 INTRODUCTION AND REVIEW

1-5 MATHEMATICS

In this section we shall discuss several mathematical techniques or results that are repeatedly used in statistical thermodynamics. The topics we shall discuss here are random variables and distribution functions, Stirling's approximation, the binomial and multinomial coefficients, the Lagrange method of undetermined multipliers, and the behavior of binomial and multinomial coefficients for large numbers.

PROBABILITY DISTRIBUTIONS

Let u be a variable which can assume the M discrete values u_1, u_2, \dots, u_M with corresponding probabilities $p(u_1), p(u_2), \dots, p(u_M)$. The variable u is said to be a discrete random variable, and $p(u)$ is said to be a discrete distribution. The mean, or average, value of u is

$$\bar{u} = \frac{\sum_{j=1}^M u_j p(u_j)}{\sum_{j=1}^M p(u_j)}$$

Since $p(u_j)$ is a probability, $p(u_j)$ must be normalized, that is, the summation in the denominator must equal unity. The mean of any function of u , $f(u)$, is given by

$$\overline{f(u)} = \sum_{j=1}^M f(u_j) p(u_j) \quad (1-69)$$

If $f(u) = u^m$, $\overline{f(u)}$ is called the m th moment of the distribution $p(u)$. If $f(u) = (u - \bar{u})^m$, $\overline{f(u)}$ is called the m th central moment, that is, the m th moment about the mean. In particular, the mean of $(u - \bar{u})^2$ is called the variance, and is a measure of the spread of the distribution. The square root of the variance is the standard deviation.

A very commonly occurring and useful discrete distribution is the Poisson distribution:

$$P(m) = \frac{a^m e^{-a}}{m!} \quad m = 0, 1, 2, \dots \quad (1-70)$$

This distribution has been applied to shot noise in electron tubes, the distribution of galaxies in space, aerial search, and many others.* (See Problem 1-42.)

If the random variable U is continuous rather than discrete, then we interpret $p(u) du$ as the probability that the random variable U lies between the values u and $u + du$. The mean of any function of U is

$$\overline{f(u)} = \int f(u) p(u) du \quad (1-71)$$

The limits of the integral are over the entire range of U .

The most important continuous probability distribution is the Gaussian distribution:

$$p(x) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{-\frac{(x - \bar{x})^2}{2\sigma^2}\right\} \quad -\infty \leq x \leq \infty \quad (1-72)$$

The quantity σ^2 , which is the variance, controls the width of the Gaussian distribution. The smaller the σ , the narrower the Gaussian distribution becomes. In the limit $\sigma \rightarrow 0$, Eq. (1-72) becomes a delta function (this is one representation of a delta function of Appendix B). Problems 1-43 through 1-45 involve some important results based on Eq. (1-72).

* See *Modern Probability Theory and Its Applications* by E. Parzen (New York: Wiley, 1960).

STIRLING'S APPROXIMATION

In statistical thermodynamics we often encounter factorials of very large numbers, such as Avogadro's number. The calculation and mathematical manipulation of factorials become awkward for large N . Therefore it is desirable to find an approximation for $N!$ for large N . Problems of this sort occur often in mathematics and are called asymptotic approximations, that is, an approximation to a function which improves as the argument of that function increases. Since $N!$ is actually a product, it is convenient to deal with $\ln N!$ because this is a sum. The asymptotic approximation to $\ln N!$ is called Stirling's approximation, which we now derive.

Since $N! = N(N - 1)(N - 2) \cdots (2)(1)$, $\ln N!$ is

$$\ln N! = \sum_{m=1}^N \ln m \quad (1-73)$$

Figure 1-5 shows $\ln x$ plotted versus x . The sum of the areas under these rectangles up to N is $\ln N!$. Figure 1-5 also shows the continuous curve $\ln x$ plotted on the same graph. Thus $\ln x$ is seen to form an envelope to the rectangles, and this envelope becomes a steadily smoother approximation to the rectangles as x increases. We can approximate the area under these rectangles by the integral of $\ln x$. The area under $\ln x$ will poorly approximate the rectangles only in the beginning. If N is large enough (we are deriving an asymptotic expansion), this area will make a negligible contribution to the total area. We may write, then,

$$\ln N! = \sum_{m=1}^N \ln m \approx \int_1^N \ln x \, dx = N \ln N - N \quad (N \text{ large}) \quad (1-74)$$

which is Stirling's approximation to $\ln N!$. The lower limit could just as well have been taken as 0 in Eq. (1-74), since N is large. (Remember that $x \ln x \rightarrow 0$ as $x \rightarrow 0$.)

A more refined derivation of Stirling's approximation gives $\ln N! \approx N \ln N - N + \ln(2\pi N)^{1/2}$, but this additional term is seldom necessary. (See Problem 1-59.)

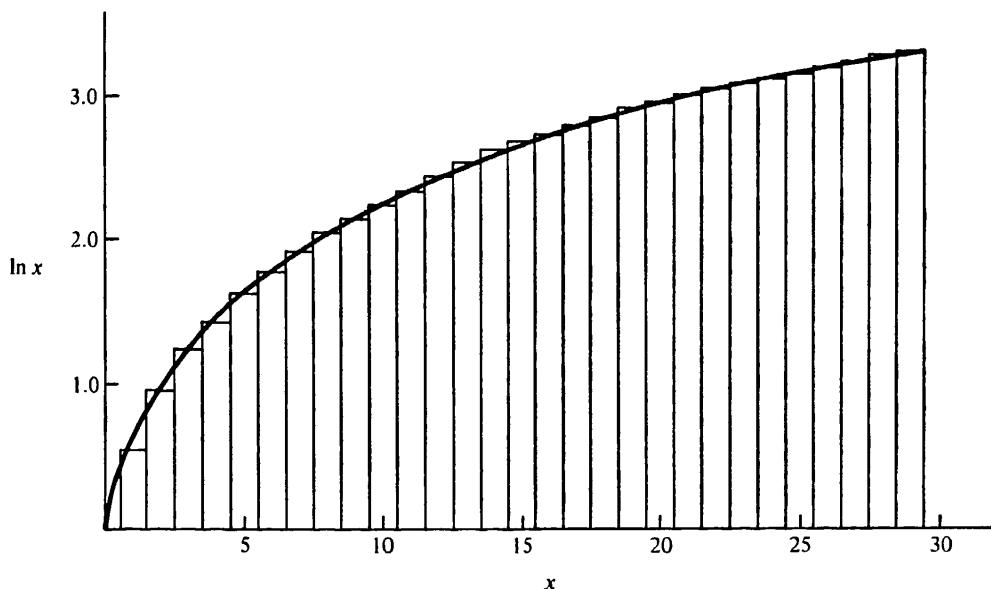


Figure 1-5. A plot of $\ln x$ versus x , showing how the summation of $\ln m$ can be approximated by the integral of $\ln x$.

22 INTRODUCTION AND REVIEW

BINOMIAL AND MULTINOMIAL DISTRIBUTION

During the course of our discussion of the canonical ensemble, we shall encounter the problem of determining how many ways it is possible to divide N distinguishable systems into groups such that there are n_1 systems in the first group, n_2 systems in the second group, and so on, and such that $n_1 + n_2 + \dots = N$, that is, all the systems are accounted for. This is actually one of the easiest problems in combinatorial analysis. To solve this, we first calculate the number of permutations of N distinguishable objects, that is, the number of possible different arrangements or ways to order N distinguishable objects. Let us choose one of the N objects and place it in the first position, one of the $N - 1$ remaining objects and place it in the second position, and so on, until all N objects are ordered. Clearly there are N choices for the first position, $N - 1$ choices for the second position, and so on, until finally there is only one object left for the N th position. The total number of ways of doing this is then the product of all the choices,

$$N(N - 1)(N - 2) \cdots (2)(1) \equiv N! \quad (\text{distinguishable objects})$$

Next we calculate the number of ways of dividing N distinguishable objects into two groups, one group containing N_1 objects, say, and the other containing the remaining $N - N_1$. There are $N(N - 1) \cdots (N - N_1 + 1)$ ways to form the first group, and $N_2! = (N - N_1)!$ ways to form the second group. The total number is, then, the product

$$N(N - 1) \cdots (N - N_1 + 1) \times (N - N_1)! = \frac{N!}{(N - N_1)!} \times (N - N_1)! = N!$$

But this has overcounted the situation drastically, since the order in which we place N_1 members in the first group and N_2 in the second group is immaterial to the problem as stated. All $N_1!$ orders of the first group and $N_2!$ orders of the second group correspond to just one division of N objects into N_1 objects and N_2 objects. Therefore the desired result is

$$\frac{N!}{N_1!(N - N_1)!} = \frac{N!}{N_1!N_2!} \quad (1-75)$$

Since the combination of factorials in Eq. (1-75) occurs in the binomial expansion,

$$(x + y)^N = \sum_{N_1=0}^N \frac{N! x^{N-N_1} y^{N_1}}{N_1!(N - N_1)!} = \sum_{N_1N_2}^* \frac{N! x^{N_1} y^{N_2}}{N_1!N_2!} \quad (1-76)$$

$N!/N_1!(N - N_1)!$ is called a binomial coefficient. The asterisk on the second summation in Eq. (1-76) signifies the restriction $N_1 + N_2 = N$.

The generalization of Eq. (1-75) to the division of N into r groups, the first containing N_1 , and so on, is easily seen to be

$$\frac{N!}{N_1!N_2! \cdots N_r!} = \frac{N!}{\prod_{j=1}^r N_j!} \quad (1-77)$$

where $N_1 + N_2 + \cdots + N_r = N$. This is known as a multinomial coefficient, since it occurs in the expansion

$$(x_1 + x_2 + \cdots + x_r)^N = \sum_{N_1=0}^N \sum_{N_2=0}^N \cdots \sum_{N_r=0}^N \frac{N! x_1^{N_1} \cdots x_r^{N_r}}{\prod_{j=1}^r N_j!} \quad (1-78)$$

where this time the asterisk signifies the restriction $N_1 + N_2 + \cdots + N_r = N$.

There are a number of other combinatorial formulas that are useful in statistical thermodynamics, but Eq. (1-77) is the most useful for our purposes. Combinatorial formulas can become rather demanding to derive. We refer to Appendix AVII of Mayer and Mayer* which contains a collection of formulas.

METHOD OF LAGRANGE MULTIPLIERS

It will be necessary, later, to maximize Eq. (1-77) with the constraint $N_1 + N_2 + \dots + N_r = \text{constant}$. This brings us to the mathematical problem of maximizing a function of several (or many) variables $f(x_1, x_2, \dots, x_r)$ when the variables are connected by other equations, say $g_1(x_1, \dots, x_r) = 0, g_2(x_1, \dots, x_r) = 0$, and so on. This type of problem is readily handled by the method of Lagrange undetermined multipliers.

If it were not for the constraints, $g_j(x_1, x_2, \dots, x_r) = 0$, the maximum of $f(x_1, \dots, x_r)$ would be given by

$$\delta f = \sum_{j=1}^r \left(\frac{\partial f}{\partial x_j} \right)_0 \delta x_j = 0 \quad (1-79)$$

where the zero subscript indicates that this equation equals zero only when the r partial derivatives are evaluated at the maximum (or minimum) of f . Denote these values of x_j by x_j^0 . If there were no constraints, each of the δx_j would be able to be varied independently and arbitrarily, and so we would conclude that $(\partial f / \partial x_j) = 0$ for every j , since δf must equal zero. This would give r equations from which the values of the $r x_j^0$ could be obtained.

On the other hand, if there is some other relation between the x 's, such as $g(x_1, x_2, \dots, x_r) = 0$, we have the additional equation

$$\delta g = \sum_{j=1}^r \left(\frac{\partial g}{\partial x_j} \right)_0 \delta x_j = 0 \quad (1-80)$$

This equation serves as a constraint that the δx_j must satisfy, thus making one of them depend upon the other $r - 1$. In the Lagrange method, one multiplies Eq. (1-80) by some parameter, say λ , and adds the result to Eq. (1-79) to get

$$\sum_{j=1}^r \left(\frac{\partial f}{\partial x_j} - \lambda \frac{\partial g}{\partial x_j} \right)_0 \delta x_j = 0 \quad (1-81)$$

The δx_j are still not independent, because of Eq. (1-80), and so they cannot be varied independently. Equation (1-80), however, can be treated as an equation giving one of the δx_j in terms of the other $r - 1$ independent ones. Pick any one of the $r \delta x_j$ as the dependent one. Let this be δx_μ .

The trick now is that we have not specified λ yet. We set it equal to $(\partial f / \partial x_\mu)_0 / (\partial g / \partial x_\mu)_0$, making the coefficient of δx_μ in Eq. (1-81) vanish. The subscript zero here indicates that $(\partial f / \partial x_\mu)$ and $(\partial g / \partial x_\mu)$ are to be evaluated at values of the x_j such that f is at its maximum (or minimum) under the constraint of Eq. (1-80). Of course, we do not know these values of x_j yet, but we can nevertheless formally define λ in this manner. This leaves a sum of terms in Eq. (1-81) involving only the independent δx_j , which can be varied independently, yielding that

$$\left(\frac{\partial f}{\partial x_j} \right)_0 - \lambda \left(\frac{\partial g}{\partial x_j} \right)_0 = 0 \quad j = 1, 2, \dots, \mu - 1, \mu + 1, \dots, r$$

* See Mayer and Mayer, *Statistical Mechanics* (New York: Wiley, 1940).

24 INTRODUCTION AND REVIEW

If we combine these $r - 1$ equations with our choice for λ , we have

$$\left(\frac{\partial f}{\partial x_j}\right)_0 - \lambda \left(\frac{\partial g}{\partial x_j}\right)_0 = 0 \quad (1-82)$$

for all j .

As we said above, the choice of λ here is certainly formal, since both $(\partial f / \partial x_\mu)_0$ and $(\partial g / \partial x_\mu)_0$ must be evaluated at these values of x_j which maximizes f , but these are known from Eq. (1-82) only in terms of λ . But this presents no difficulty, since in practice λ is determined by physical requirements. Examples of this will occur in the next two chapters.

Lagrange's method becomes no more difficult in the case in which there are several constraints. Let $g_1(x_1, \dots, x_r), g_2(x_1, \dots, x_r), \dots$ be a set of constraints. We introduce a Lagrange multiplier for each $g_i(x_1, \dots, x_r)$ and proceed as above to get

$$\frac{\partial f}{\partial x_j} - \lambda_1 \frac{\partial g_1}{\partial x_j} - \lambda_2 \frac{\partial g_2}{\partial x_j} - \dots = 0 \quad (1-83)$$

BINOMIAL DISTRIBUTION FOR LARGE NUMBERS

Lastly, there is one other mathematical observation we need here in order to facilitate the discussion in the next chapter. This observation concerns the shape of the multinomial coefficient [Eq. (1-78)] as a function of the N_j 's, as the N_j 's become very large. To simplify notation, we shall consider only the binomial coefficient, but this will not affect our conclusions. Let us first find the value of N_1 for which $f(N_1) = N! / N_1!(N - N_1)!$ reaches its maximum value. Since N_1 and N are both very large, we treat them as continuous variables. Also since $\ln x$ is a monotonic function of x , we can maximize $f(N_1)$ by maximizing $\ln f(N_1)$. This allows us to use Stirling's approximation. The maximum of $f(N_1)$ is found, then, from

$$\frac{d \ln f(N_1)}{d N_1} = 0$$

to be located at $N_1^* = N/2$. Let us now expand $\ln f(N_1)$ about this point. The Taylor expansion is

$$\ln f(N_1) = \ln f(N_1^*) + \frac{1}{2} \left(\frac{d^2 \ln f(N_1)}{d N_1^2} \right)_{N_1=N_1^*} (N_1 - N_1^*)^2 + \dots \quad (1-84)$$

The linear term in $N_1 - N_1^*$ is missing, because the first derivative of $\ln f(N)$ is zero at $N_1 = N_1^*$. The second derivative appearing in Eq. (1-84) is equal to $-4/N$. Thus if we ignore higher-order terms (see Problem 1-53), Eq. (1-84) can be written in the form of a Gaussian curve

$$f(N_1) = f(N_1^*) \exp \left\{ -\frac{2(N_1 - N_1^*)^2}{N} \right\} \quad (1-85)$$

Comparison of this with the standard form of the Gaussian function

$$f(x) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp \left\{ -\frac{(x - x^*)^2}{2\sigma^2} \right\} \quad (1-86)$$

shows that the standard deviation is of the order of $N^{1/2}$. Equation (1-85) is, therefore, a bell-shaped function, centered at $N_1^* = N/2$ and having a width of a few multiples of $N^{1/2}$. Problem 1-43 establishes the well-known fact that a Gaussian

function goes essentially to zero when x differs from x^* by a few σ 's. Since we are interested only in large values of N_1 (or N), say numbers of the order of 10^{20} , we have a bell-shaped curve that is contained between $10^{20} \pm$ a few multiples of 10^{10} , which, if plotted, would for all practical purposes look like a delta function centered at $N_1^* = N/2$. Thus we have shown that the binomial coefficient peaks very strongly at the point $N_1 = N_2 = N/2$. This same behavior occurs for a multinomial coefficient as well. If there are s N_j 's, the multinomial coefficient has a very sharp maximum at the point $N_1 = N_2 = \dots = N_s = N/s$. (See Problem 1-50.) This peak becomes sharper as the N_j 's become larger, and become a delta function in the limit $N_j \rightarrow \infty$ for all j .

MAXIMUM TERM METHOD

Another important result, which is a consequence of the large numbers encountered in statistical mechanics, is the *maximum-term method*. It says that under appropriate conditions the logarithm of a summation is essentially equal to the logarithm of the maximum term in the summation. To see how this goes, consider the sum

$$S = \sum_{N=1}^M T_N$$

where $T_N > 0$ for all N . Since all the terms are positive, the value of S must be greater than the value of the largest term, say T_{\max} , and less than the product of the number of terms and the value of the largest term. Thus we can write

$$T_{\max} \leq S \leq M T_{\max}$$

Taking logarithms gives

$$\ln T_{\max} \leq \ln S \leq \ln T_{\max} + \ln M$$

We shall see that it is often the case in statistical mechanics that T_{\max} will be $O(e^M)$. Thus we have

$$O(M) \leq \ln S \leq O(M) + \ln M$$

For large M , $\ln M$ is negligible with respect to M itself, and so we see that $\ln S$ is bounded from above and below by $\ln T_{\max}$, and so

$$\ln S = \ln T_{\max}$$

This is a rather remarkable theorem, and like a number of other theorems used in statistical mechanics, its validity results from the large numbers involved.

ADDITIONAL READING

1-2

- DAVIDSON, N. 1962. *Statistical mechanics*. New York: McGraw-Hill. Chapter 2.
 GOLDSTEIN, H. 1950. *Classical mechanics*. Reading, Mass.: Addison-Wesley.
 SYMON, K. R. 1960. *Mechanics*. Reading, Mass.: Addison-Wesley.
 TOLMAN, R. C. 1938. *Statistical mechanics*. London: Oxford University Press. Chapters 1 and 2.

1-3

- HAMEKA, H. F. 1967. *Introduction to quantum theory*. New York: Harper & Row.
 HANNA, M. W. 1969. *Quantum mechanics in chemistry*, 2nd ed. New York: Benjamin.
 HILL, T. L. 1956. *Statistical mechanics*. New York: McGraw-Hill. Section 8.
 KARPLUS, M., and PORTER, R. N. 1970. *Atoms and molecules*. New York: Benjamin.
 KESTIN, J., and DORFMAN, J. R. 1971. *A course in statistical thermodynamics*. New York: Academic. Chapter 3.
 PILAR, F. L. 1968. *Elementary quantum chemistry*. New York: McGraw-Hill.
 TOLMAN, R. C. 1938. *Statistical mechanics*. London: Oxford University Press. Chapters 7 and 8.

26 INTRODUCTION AND REVIEW

1-4

- ANDREWS, F. C. 1971. *Thermodynamics*. New York: Wiley.
 CALLEN, H. B. 1960. *Thermodynamics*. New York: Wiley.
 MAHAN, B. H. 1963. *Elementary chemical thermodynamics*. New York: Benjamin.
 MOORE, W. J. 1972. *Physical chemistry*, 4th ed. Englewood Cliffs, N.J.: Prentice-Hall.
 REISS, H. 1965. *Methods of thermodynamics*. Boston, Mass.: Blaisdell.
 WASER, J. 1966. *Basic chemical thermodynamics*. New York: Benjamin.

1-5

- ABRAMOWITZ, M., and STEGUN, I. A. "Handbook of Mathematical Functions," *Natl. Bur. Stan. Appl. Math. Series*, 55, 1964.
 ARFKEN, G., 1970. *Mathematical methods for physicists*, 2nd ed. New York: Academic.
 KESTIN, J., and DORFMAN, J. R. 1971. *A course in statistical thermodynamics*. New York: Academic. Chapter 4.
 KREYSZIG, E. 1962. *Advanced engineering mathematics*, 2nd ed. New York, Wiley.
 REIF, F. 1965. *Statistical and thermal physics*. New York: McGraw-Hill, Chapter 1.

PROBLEMS

1-1. Solve the equation of motion of a body of mass m dropped from a height h . Assume that there exists a viscous drag on the body that is proportional to and in the opposite direction to the velocity of the body. (Let the proportionality constant be γ .) Solve for the so-called terminal velocity, that is, the limiting velocity as $t \rightarrow \infty$.

1-2. Calculate the trajectory of a shell shot out of a cannon with velocity v_0 , assuming no aerodynamic resistance and that the cannon makes an angle θ with the horizontal axis.

1-3. Remembering that the potential energy is given by

$$V(x) = - \int_0^x F(\xi) d\xi = \frac{1}{2} k x^2$$

for a simple harmonic oscillator, derive an expression for the total energy as a function of time. Discuss how the kinetic and potential energy behave as a function of time.

1-4. Solve the equation for a harmonic oscillator of mass m and force constant k that is driven by an external force of the form $F(t) = F_0 \cos \omega_0 t$.

1-5. Show that

$$\xi(t) = A \sin \omega t + B \cos \omega t$$

can be written as

$$\xi(t) = C \sin(\omega t + \phi)$$

1-6. Show that the total linear momentum is conserved for a system of N particles with an interaction potential which depends only on the distance between particles.

1-7. When does $p = \partial L / \partial \dot{q}$ but $\neq \partial K / \partial \dot{q}$?

1-8. Consider a system of two-point particles with masses m_1 and m_2 moving in two dimensions. It is very common for their potential of interaction to depend upon their relative coordinates $(x_1 - x_2, y_1 - y_2)$ only. Thus the total energy is

$$E = \frac{m_1}{2} (\dot{x}_1^2 + \dot{y}_1^2) + \frac{m_2}{2} (\dot{x}_2^2 + \dot{y}_2^2) + U(x_1 - x_2, y_1 - y_2)$$

Now introduce four new variables

$$X = \frac{m_1 x_1 + m_2 x_2}{m_1 + m_2} \quad Y = \frac{m_1 y_1 + m_2 y_2}{m_1 + m_2}$$

$$x_{12} = x_1 - x_2 \quad y_{12} = y_1 - y_2$$

and show that this two-body problem can be reduced to two one-body problems, one involving the center of mass of the system and one involving the *relative* motion of the two particles.

Give a physical interpretation of the ratio $m_1 m_2 / (m_1 + m_2)$ that arises naturally in the relative motion. What is this quantity called? This result is easily extended to three dimensions.

1-9. Extend the development of Problem 1-8 to the case in which each particle also experiences an external potential energy, say $U(x_1, y_1, z_1)$ and $U(x_2, y_2, z_2)$. Interpret the resulting equations.

1-10. Derive Lagrange's equations for a particle moving in two dimensions under a central potential $u(r)$. Which of these equations illustrates the law of conservation of angular momentum? Is angular momentum conserved if the potential depends upon θ as well?

1-11. For a particle moving in three dimensions under the influence of a spherically symmetrical potential $U = U(r)$, write down the Lagrangian and the equations of motion in spherical coordinates (r, θ, ϕ) . Show that $H = K + V$ from

$$H = \sum p_i \dot{q}_i - L$$

for this potential.

1-12. Solve the equation of motion of two masses m_1 and m_2 connected by a harmonic spring with force constant k .

1-13. Start with Lagrange's equations in Cartesian coordinates, that is,

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

and so on. Now introduce three generalized coordinates q_1 , q_2 , and q_3 which are related to the Cartesian coordinates by $x = x(q_1, q_2, q_3)$, and so on. Show that by transforming Lagrange's equations from x , \dot{x} , y , \dot{y} , z , and \dot{z} as independent variables to q_1 , \dot{q}_1 , q_2 , \dot{q}_2 , q_3 , and \dot{q}_3 we get

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_1} \right) = \frac{\partial L}{\partial q_1}$$

and so on.

1-14. If H , the classical Hamiltonian, does not depend explicitly on time, show that $dH/dt = 0$. What does this mean physically? Is this true if H does depend explicitly upon time?

1-15. Consider the rotation of a diatomic molecule with a fixed internuclear separation l and masses m_1 and m_2 . By employing center of mass and relative coordinates, show that the rotational kinetic energy can be written in spherical coordinates as

$$\frac{1}{2} I(\dot{\theta}^2 + \dot{\phi}^2 \sin^2 \theta)$$

and from this derive the rotational Hamiltonian

$$H_{\text{rot}} = \frac{1}{2I} \left(p_\theta^2 + \frac{p_\phi^2}{\sin^2 \theta} \right)$$

In these equations, $I = \mu l^2$, where μ is the reduced mass. This Hamiltonian is useful for studying the rotation of diatomic molecules.

1-16. Show that the motion of a particle under a central force law takes place entirely in a single plane.

1-17. What is the expectation (average) value for the linear momentum p_x of a particle in a one-dimensional box p_x^2 ? Briefly discuss your results.

1-18. Show that the energy eigenvalues of a free particle confined to a cube of length a are given by

$$\epsilon = \frac{\hbar^2}{8ma^2} (n_x^2 + n_y^2 + n_z^2) \quad n_x, n_y, n_z = 1, 2, \dots$$

28 INTRODUCTION AND REVIEW

1-19. Show that the energy eigenvalues of a free particle confined to a rectangular parallelepiped of lengths a , b , and c are given by

$$\epsilon = \frac{\hbar^2}{8m} \left(\frac{n_x^2}{a^2} + \frac{n_y^2}{b^2} + \frac{n_z^2}{c^2} \right) \quad n_x, n_y, n_z = 1, 2, \dots$$

1-20. Calculate the energy eigenvalues of a particle confined to a ring of radius a .

1-21. Show that the Hamiltonian operator of a rigid rotor is given by Eq. (1-32).

1-22. Calculate the degeneracy of the first few levels of a free particle confined to a cube of length a .

1-23. Verify the calculation that follows Eq. (1-37) which shows that the quantum mechanical degeneracy of a macroscopic system is $O(10^N)$.

1-24. We need to know the volume of an N -dimensional sphere in order to derive Eq. (1-36). This can be determined by the following device. Consider the integral

$$I = \int_{-\infty}^{\infty} \cdots \int e^{-(x_1^2 + x_2^2 + \cdots + x_N^2)} dx_1 dx_2 \cdots dx_N$$

First show that $I = \pi^{N/2}$. Now one can formally transform the volume element $dx_1 dx_2 \cdots dx_N$ to N -dimensional spherical (hyperspherical) coordinates to get

$$\int_{\text{angles}} dx_1 dx_2 \cdots dx_N \rightarrow r^{N-1} S_N dr$$

where S_N is the factor that arises upon integration over the angles. Show that $S_2 = 2\pi$ and $S_3 = 4\pi$. S_N can be determined for any N by writing I in hyperspherical coordinates:

$$I = \int_0^{\infty} e^{-r^2} r^{N-1} S_N dr$$

Show that $I = S_N \Gamma(N/2)/2$, where $\Gamma(x)$ is the gamma function (see Problem 1-58). Equate these two values for I to get

$$S_N = \frac{2\pi^{N/2}}{\Gamma(N/2)}$$

Show that this reduces correctly for $N = 2$ and 3 . Lastly now, convince yourself that the volume of an N -dimensional sphere of radius a is given by

$$V_N = \int_0^a S_N r^{N-1} dr = \frac{\pi^{N/2}}{\Gamma\left(\frac{N}{2} + 1\right)} a^N$$

and show that this reduces correctly for $N = 2$ and 3 .

1-25. Derive an expression for the density of translational quantum states for a two-dimensional ideal gas.

1-26. Prove that a many-body wave function must be either symmetric or antisymmetric under the interchange of any two particles. Hint: Apply the exchange operation twice.

1-27. Show for an isothermal process that $w_{\text{rev}} > w_{\text{irrev}}$ and $q_{\text{rev}} > q_{\text{irrev}}$.

1-28. Derive the thermodynamic equation

$$C_p - C_V = \left[p + \left(\frac{\partial E}{\partial V} \right)_T \right] \left(\frac{\partial V}{\partial T} \right)_p$$

and evaluate this difference for an ideal gas and a gas that obeys the van der Waals equation.

1-29. Derive the thermodynamic equation of state

$$\left(\frac{\partial E}{\partial V} \right)_T - T \left(\frac{\partial p}{\partial T} \right)_V = -p$$

1-30. Derive the equation

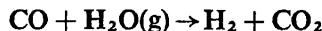
$$dE = \left[T \left(\frac{\partial p}{\partial T} \right)_V - p \right] dV + C_V dT$$

and from this show that $(\partial E / \partial V)_T = a/V^2$ for a van der Waals gas.

1-31. Show that

$$\left(\frac{\partial E}{\partial V} \right)_{\mu/T, 1/T} + \frac{1}{T} \left(\frac{\partial p}{\partial (1/T)} \right)_{\mu/T, V} = -p$$

1-32. Derive an expression for $\partial \ln K_p / \partial T$ in terms of ΔH , the heat of reaction, and in terms of C_p , the heat capacity at constant pressure.

1-33. Consider the "water-gas" reaction

where

$$K_p = \frac{P_{\text{H}_2} P_{\text{CO}_2}}{P_{\text{CO}} P_{\text{H}_2\text{O}}}$$

and given the following data:

Substance	(kcal/mole)	a	$b \times 10^3$	$c \times 10^7$	$\Delta H_{298}^0(\text{kcal/m})$
CO	-32.81	6.42	1.67	1.96	-26.4157
H ₂ O(g)	-54.64	7.26	2.30	2.83	-57.7979
CO ₂	-94.26	6.21	10.40	-35.45	-94.0518
H ₂	0.00	6.95	-0.20	4.81	0.00

where the heat capacity of the gases in cal deg⁻¹ mole⁻¹ is given by

$$C_p = a + bT + cT^2$$

Calculate K_p at 298°K and 800°K.

1-34. Calculate the free energy change at 700°C for the conversion of carbon monoxide at 10 atm and water vapor at 5 atm to carbon dioxide and hydrogen at partial pressures of 1.5 atm each. The equilibrium constant K_p for this reaction is 0.71. Is this process theoretically feasible?

1-35. It is illustrated in Chapter 17 that the speed of sound c_0 propagated through a gas is

$$c_0 = (m\rho\kappa_s)^{-1/2}$$

where κ_s is the adiabatic compressibility

$$\kappa_s = -\frac{1}{V} \left(\frac{\partial V}{\partial p} \right)_s$$

Show that this is equivalent to

$$c_0 = V \left\{ -\frac{\gamma}{M} \left(\frac{\partial p}{\partial V} \right)_T \right\}^{1/2}$$

where $\gamma = C_p/C_V$, and M is the molecular weight of the gas. Using the above result, show that

$$c_0 = \left(\gamma \frac{RT}{M} \right)^{1/2}$$

for an ideal gas.

30 INTRODUCTION AND REVIEW

1-36. Jones and Giauque obtained the following values for C_p of nitromethane.

°K	15	20	30	40	50	60	70	80	90	100
C_p	0.89	2.07	4.59	6.90	8.53	9.76	10.70	11.47	12.10	12.62
°K	120	140	160	180	200	220	240	260	280	300
C_p	13.56	14.45	15.31	16.19	17.08	17.98	18.88	25.01	25.17	25.35

The melting point is 244.7°K, heat of fusion 2319 cal/mole. The vapor pressure of the liquid at 298.1°K is 3.666 cm. The heat of vaporization at 298.0°K is 9147 cal/mole. Calculate the third-law entropy of CH_3NO_2 gas at 298.1°K and 1 atm pressure (assuming ideal gas behavior).

1-37. Derive the Legendre transformation of E in which $S \rightarrow T$ and $N \rightarrow \mu$.

1-38. Apply a Legendre transformation to the Lagrangian $L(q_j, \dot{q}_j)$ to eliminate the generalized velocities in favor of generalized momenta, defined by $p_j \equiv \partial L / \partial \dot{q}_j$. What function does this turn out to be?

1-39. Find the natural function of V , E , and μ . Hint: Start with the natural function of V , E , and N , namely, S , and transform $N \rightarrow \mu$.

1-40. Derive the Legendre transformation of E in which $S \rightarrow T$, $N \rightarrow \mu$, and $V \rightarrow p$. What peculiar thing happens when all the extensive variables are transformed out?

1-41. Show that $\overline{(x - \bar{x})^2} = \bar{x^2} - \bar{x}^2$.

1-42. Show that the Poisson distribution $P(m) = a^m e^{-a}/m!$ is normalized. Calculate \bar{m} and the variance. What is the significance of the parameter a ?

1-43. Sketch the Gaussian distribution as σ (or even σ/\bar{x}) becomes smaller and smaller. To what type of distribution does a Gaussian go in the limit $\sigma \rightarrow 0$. Discuss the meaning of this distribution.

1-44. For the Gaussian distribution $p(x)$ show that

(a)

$$\int_{-\infty}^{\infty} p(x) dx = 1$$

(b) Calculate the n th central moment where $n = 0, 1, 2$, and 3 .

(c) In the limit $\sigma \rightarrow 0$ what kind of distribution is approached where

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x - \bar{x})^2}{2\sigma^2}\right)$$

1-45. The quantity $\overline{(x - \bar{x})^j}$ is called the j th central moment. Show that all odd central moments of a Gaussian vanish. What about the even ones? Relate the $j=2$ central moment to the parameter σ .

1-46. Let $f(x, y)$ be a joint probability density, that is, $f(x, y) dx dy$ is the probability that X lies between x and $x + dx$ and Y lies between y and $y + dy$. If X and Y are independent, then

$$f(x, y) dx dy = f_1(x)f_2(y) dx dy$$

If X and Y are independent, show that the mean and variance of their sum is equal to the sum of the means and variances, respectively, of X and Y ; that is, show that if $W = X + Y$, then

$$\bar{W} = \bar{X} + \bar{Y}$$

$$\overline{(W - \bar{W})^2} = \overline{(X - \bar{X})^2} + \overline{(Y - \bar{Y})^2}$$

1-47. Let X be a random variable on the positive numbers, $0 \leq x < \infty$, and let $p(x)$ be its probability density function. The function $\phi(s)$ defined by

$$\phi(s) = \int_0^{\infty} e^{-sx} p(x) dx$$

is called the characteristic function of $p(x)$. Find the relation between $\phi(s)$ and the moments of $p(x)$. Is knowledge of all the moments of $p(x)$ (assuming they exist) sufficient to specify $p(x)$ itself? Why or why not?

1-48. Show that the characteristic function of the density function of the sum of two independent random variables is the product of the characteristic functions of the densities of the two random variables themselves. What is the density function of $W = X + Y$?

1-49. Maximize

$$W(N_1, N_2, \dots, N_M) = \frac{N!}{\prod_{j=1}^M N_j!}$$

with respect to each N_j under the constraints that

$$\sum N_j = N = \text{a fixed constant}$$

$$\sum E_j N_j = \mathcal{E} = \text{another fixed constant}$$

Hint: Consider the N_j 's to be continuous, large enough to use Stirling's approximation of $N_j!$, and leave your answer in terms of the two undetermined multipliers.

1-50. Show that the maximum of a multinomial distribution is given when $N_1 = N_2 = \dots = N_s = N/s$.

1-51. Use the method of undetermined multipliers to show that

$$-\sum_{j=1}^N P_j \ln P_j$$

subject to the condition

$$\sum_{j=1}^N P_j = 1$$

is a maximum when $P_j = \text{constant}$.

1-52. Consider the sum

$$\sum_{N=0}^M \frac{M!x^N}{N!(M-N)!}$$

where $x = O(1)$, and M and N are $O(10^{20})$. First show that $\ln \sum = M \ln(1+x)$ exactly, and then calculate the logarithm of the maximum term. Hint: Remember the binomial expansion.

1-53. Show that the higher terms that were dropped in the expansion of $\ln f(N)$ in Eq. (1-84) are completely negligible for large values of N and M .

1-54. The Planck blackbody distribution law

$$\rho(\omega, T) d\omega = \frac{\hbar}{\pi^2 c^3} \frac{\omega^3 d\omega}{\exp(\beta \hbar \omega) - 1}$$

gives the blackbody radiation energy density between frequencies ω and $\omega + d\omega$. ($\hbar \equiv h/2\pi$, $\omega \equiv 2\pi\nu$, and $\epsilon = h\nu = \hbar\omega$.) Substitute this into

$$\frac{E}{V} = \int_0^\infty \rho(\omega, T) d\omega$$

to derive the temperature dependence of E/V . Do this by expressing your result as a group of factors multiplying a dimensionless integral. You do not need to evaluate this integral.

1-55. Show that $e^x/(1 \pm e^x)^2$ is an even function of x .

1-56. The heat capacity of the Einstein model of a crystal is given by

$$C_V = 3Nk \left(\frac{\Theta_E}{T}\right)^2 \frac{e^{\Theta_E/T}}{(e^{\Theta_E/T} - 1)^2}$$

32 INTRODUCTION AND REVIEW

where Θ_E is the "characteristic temperature" of the crystal. Determine both the high- and low-temperature limiting expressions for the heat capacity. Do the same thing for the Debye model of crystals, in which

$$C_V = 9Nk \left(\frac{T}{\Theta_D}\right)^3 \int_0^{\Theta_D/T} \frac{x^4 e^x dx}{(e^x - 1)^2}$$

where Θ_D is the Debye temperature of the crystal.

1-57. Recognizing it as a geometric series, sum the following series in closed form:

$$S = \sum_{n=0}^{\infty} e^{-an}$$

Compare this result to

$$I = \int_0^{\infty} e^{-an} dn$$

Under what conditions are these two results the same?

1-58. One often encounters the gamma function in statistical thermodynamics. It was introduced by Euler as a function of x , which is continuous for positive values of x and which reduces to $n!$ when $x = n$, an integer. The gamma function $\Gamma(x)$ is defined by

$$\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1} dt$$

First show by integrating by parts that

$$\Gamma(x+1) = x\Gamma(x)$$

Using this, show that $\Gamma(n+1) = n!$ for n an integer. Show that

$$\Gamma(\frac{1}{2}) = \sqrt{\pi}$$

Evaluate $\Gamma(\frac{3}{2})$ using the recurrence formula $\Gamma(x+1) = x\Gamma(x)$. Lastly show that

$$\begin{aligned} \Gamma\left(n + \frac{1}{2}\right) &= \frac{1 \cdot 3 \cdots (2n-1)}{2^n} \Gamma\left(\frac{1}{2}\right) \\ &= \frac{(2n)!}{2^{2n} n!} \sqrt{\pi} \end{aligned}$$

For a discussion of the gamma function, see G. Arfken, *Mathematical Methods for Physicists*, 2nd ed. (New York: Academic, 1970).

1-59. We can derive Stirling's approximation from an asymptotic approximation to the gamma function $\Gamma(x)$. From the previous problem

$$\begin{aligned} \Gamma(N+1) &= N! = \int_0^{\infty} e^{-x} x^N dx \\ &= \int_0^{\infty} e^{Ng(x)} dx \end{aligned}$$

where $g(x) = \ln x - x/N$. If $g(x)$ possesses a maximum at some point, say x_0 , then for large N , $\exp(Ng(x))$ will be extremely sharply peaked at x_0 . Under this condition, the integral for $N!$ will be dominated by the contribution of the integrand from the point x_0 . First show that $g(x)$ does, in fact, possess a maximum at the point $x_0 = N$. Expand $g(x)$ about this point, keeping terms only up to and including $(x-N)^2$ to get

$$g(x) \approx g(N) - \frac{(x-N)^2}{2N^2} + \dots$$

and
ebye

Why is there no linear term in $(x - N)$? Substitute this expression for $g(x)$ into the integral for $N!$ and derive the asymptotic formula

$$\ln N! \approx N \ln N - N + \ln(2\pi N)^{1/2}$$

1-60. Verify the energy conversion factors in Appendix A. (The one labeled "temperature" means that temperature required to give an energy equal to kT , where k is the Boltzmann constant.)

1-61. An integral that appears often in statistical mechanics and particularly in the kinetic theory of gases is

$$I_n = \int_0^\infty x^n e^{-ax^2} dx$$

This integral can be readily generated from two basic integrals. For even values of n , we first consider

$$I_0 = \int_0^\infty e^{-ax^2} dx$$

The standard trick to evaluate this integral is to square it, and then transform the variables into polar coordinates.

$$\begin{aligned} I_0^2 &= \int_0^\infty \int_0^\infty e^{-ax^2} e^{-ay^2} dx dy \\ &= \int_0^\infty \int_0^{\pi/2} e^{-ar^2} r dr d\theta \\ &= \frac{\pi}{4a} \\ I_0 &= \frac{1}{2} \left(\frac{\pi}{a}\right)^{1/2} \end{aligned}$$

Using this result, show that for even n

$$I_n = \frac{1 \cdot 3 \cdot 5 \cdots (n-1)}{2(2a)^{n/2}} \left(\frac{\pi}{a}\right)^{1/2} \quad n \text{ even}$$

For odd values of n , the basic integral I_1 is easy. Using I_1 , show that

$$I_n = \frac{\Gamma\left(\frac{n+1}{2}\right)}{2a^{(n+1)/2}} \quad n \text{ odd}$$

1-62. Show that a Gaussian distribution is extremely small beyond a few multiples of σ .

1-63. Another function that occurs frequently in statistical mechanics is the Riemann zeta function, defined by

$$\zeta(s) = \sum_{k=1}^{\infty} k^{-s}$$

First show that $\zeta(1) = \infty$, but that $\zeta(s)$ is finite for $s > 1$. Show that another definition of $\zeta(s)$ is

$$\zeta(s) = \frac{1}{\Gamma(s)} \int_0^\infty \frac{x^{s-1} dx}{e^x - 1}$$

was
which

cists,

to the

large
al for
that
point,

34 INTRODUCTION AND REVIEW

that is, show that this is identical to the first definition. In addition, show that

$$\eta(s) = \sum_{k=1}^{\infty} (-1)^{k-1} k^{-s} = (1 - 2^{1-s}) \zeta(s)$$

$$\lambda(s) = \sum_{k=0}^{\infty} (2k+1)^{-s} = (1 - 2^{-s}) \zeta(s)$$

The evaluation of $\zeta(s)$ for integral s can be done using Fourier series, and some results are $\zeta(2) = \pi^2/6$ and $\zeta(4) = \pi^4/90$.

For a discussion of the Riemann zeta function, see G. Arfken, *Mathematical Methods for Physicists*, 2nd ed. (New York: Academic, 1970).