CHAPTER 11

Quantum Statistics

The applications of quantum theory are not limited to extremely small systems such as isolated atoms or molecules. You already know that the quantum theory is *valid* for systems of large size as well as for small systems; but you might not realize that, for very large scale behavior of many systems, liquids or solids, the quantum theory makes predictions which sometimes differ in startling ways from those of classical theory.

At first glance, the task of constructing a quantum theory for a solid or liquid might seem hopeless. If we cannot even solve for the energy levels of a two-electron atom without making approximations, what can we do with a system containing 10^{23} particles? But we are encouraged by the example of classical theory; although the three-body problem of classical mechanics remains unsolved in general, classical theory has been quite successful in the use of statistical methods to predict the *average* behavior of systems containing many particles.

Classical statistical theory describes certain properties of gases quite successfully. We have seen (Chapter 1) that this theory involves the Boltzmann distribution function $n(\varepsilon)$, which is such that $\int_{\varepsilon_1}^{\varepsilon_2} n(\varepsilon)\, d\varepsilon$ is the average number of particles with energy between ε_1 and ε_2. But the Boltzmann distribution fails when it is applied to liquids or solids. As we shall see, it cannot be used to predict such a simple thing as the specific heat of a solid.

The reason for this failure lies in the indistinguishability of elementary particles. In a gas, although the atoms are identical, they are usually far enough apart so that they can be distinguished by their positions, and one could, in principle, "follow" a single atom as it wends its way through a gas. But the electrons in a solid or a liquid, even those belonging to different atoms, are often so close together that their wavefunctions overlap. The uncertainty principle thus prevents our defining their trajectories well enough to distinguish one from another. Or to put it another way, the motion is described by means of a single wavefunction for the *system*; this wavefunction can give probabilities for finding an electron in a certain region, but it cannot say *which* electron is found.

Except for the problem of indistinguishability, the classical derivation of the Boltzmann distribution does not run afoul of any principle of quantum theory. Let us therefore attempt to construct a distribution function which will be similar to the Boltzmann function but which will take proper account of the indistinguishability of particles. To see how we might do this, let us first analyze an example system containing only four particles.

11.1 THE THREE KINDS OF STATISTICS: AN EXAMPLE

Let us consider a system of four identical particles, and suppose that each particle can possess energy only in integral multiples of a quantity E; that is, the energy levels for each particle are $\varepsilon_1 = 0$, $\varepsilon_2 = E$, $\varepsilon_3 = 2E, \ldots$. (The assumption that the possible energies are discrete is made only to simplify the calculation, and has no quantum mechanical significance, because E can be made as small as we wish. The quantum theory enters only in the treatment of the fact that the particles are identical.)

For generality, we make the further assumption that there is more than one state for each energy level; that is, we say that some energy levels are *degenerate*. Again, the terminology sounds quantum mechanical, but it can be applied to classical theory as well. We wish to allow for the fact that certain energy levels are easier to populate than others; this is true classically as well as quantum mechanically. (For example, consider the Maxwell velocity distribution, $f(v) \propto v^2 e^{-mv^2/2kT}$. The probability that an individual molecule has a speed between v and $v + dv$ is proportional not only to the Boltzmann factor

$e^{-mv^2/2kT}$ but also the factor v^2, which is analogous to the degeneracy of a quantum-mechanical energy level.)

We shall use the symbol g_s to denote the degeneracy of the sth level. In our example we let $g_1 = 1$, $g_2 = 2$, $g_3 = 2$, $g_4 = 3$, $g_5 = 3$, and $g_6 = 4$. The energy levels and their degeneracies are shown graphically in Fig. 1; the states are depicted as boxes into which the particles are placed.

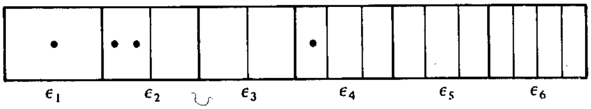

Fig. 1. *Graphic representation of states of a particle as "boxes." Heavy lines denote boundaries between different energy levels; lighter lines show boundaries between different states having the same energy. The dots indicate the particles; in the case illustrated, one particle has energy ε_1, two have energy ε_2, both occupying the same state, and a fourth particle has energy ε_4.*

The physical and chemical properties of the system are related to the average distribution of the particles among the various energy levels. This distribution, of course, depends on the temperature of the system, so let us assign a temperature by setting the total energy of the system equal to $5E$.

Before proceeding further we must pause to make sure that our definitions are clear. We define a *distribution* to be the set of numbers n_s of particles in each *energy level* ε_s, and we define an *arrangement* as a definite way of putting the particles into each of the *states*. (An arrangement is sometimes called a microscopic distribution.) Thus there can be *several* arrangements which all correspond to the *same* distribution. For example, with reference to Fig. 1, if one of the two particles in the second box were moved over to the third box (so that it still had energy ε_2, but occupied a different state), we would have a different *arrangement* but the same *distribution*: $n_1 = 1$, $n_2 = 2$, $n_3 = 0$, $n_4 = 1$, $n_5 = 0$, and $n_6 = 0$.

Now in order to find an average distribution, we must take into account all possible distributions, and combine them with the appropriate weight for each one. We determine this weighting by assuming that each *arrangement* is equally probable, because we have no reason to believe that one arrangement should be more probable than any other arrangement. If each arrangement is equally probable, the probability that a given *distribution* is present is proportional to the *number of different arrangements* which correspond to that distribution. By counting the number of particles in a given level for each distribution, and then multiplying each number by the probability of occurrence of that distribution, we can obtain the contribution of each individual distribution to the average distribution. It is this average distribution which is governed by the Boltzmann factor in the case of a gas.

At this point the indistinguishability of the particles enters the problem. In counting up the number of arrangements, we count only those which are distinct from one another. If the particles are indistinguishable, two arrangements which differ only in the exchange of two particles would be counted as only *one* arrangement. Furthermore, some arrangements are not permitted at all, if the particles obey the Pauli exclusion principle. As we have seen, electrons do obey this principle, as do all other particles of half-integer spin, but particles with integer spin (such as photons or helium-4 nuclei) do not. Thus there are two kinds of quantum statistics, depending on the kind of particles involved; particles which obey the exclusion principle are said to be governed by *Fermi–Dirac statistics*, and other particles are governed by *Bose–Einstein statistics*.

Table 1

Possible Distributions of Four Particles

	Distribution						Arrangements			Probability		
	n_1	n_2	n_3	n_4	n_5	n_6	Boltz.[a]	FD[b]	BE[c]	Boltz.	FD	BE
(a)	3	0	0	0	0	1	16	0	4	16/464	0	4/39
(b)	2	1	0	0	1	0	72	0	6	72/464	0	6/39
(c)	2	0	1	1	0	0	72	0	6	72/464	0	6/39
(d)	1	2	0	1	0	0	144	3	9	144/464	3/5	9/39
(e)	0	3	1	0	0	0	64	0	8	64/464	0	8/39
(f)	1	1	2	0	0	0	96	2	6	96/464	2/5	6/39
							464	5	39			

[a] Boltzmann.
[b] Fermi–Dirac.
[c] Bose–Einstein.

Now we are ready to solve our example. Table 1 shows all of the possible distributions which obey the given conditions, together with the number of distinct arrangements for each distribution, according to each of the three types of statistics, including Boltzmann. The probability of each distribution, found by assuming each *arrangement* to be equally likely, is also shown in the table.

To understand this table, let us see how the first line was worked out. Energy level ε_6 has energy $5E$, which is the total energy of the four particles; therefore placing one particle in that level forces the other particles to go into the zero energy level. The degeneracy of level ε_6 is $g_6 = 4$, so there are four possible arrangements for Bose–Einstein (BE) statistics, corresponding to the four different states which the particle in level ε_6 may occupy. But for Boltzmann statistics, the particles are distinguishable, so a single BE arrangement becomes four Boltzmann arrangements, one for each possible particle in level ε_6; thus there are $4 \times 4 = 16$ arrangements for Boltzmann statistics. Finally, for Fermi–Dirac (FD) statistics, there are *no* arrangements, because the exclusion principle prevents three particles from occupying the one state in the zeroth level.

Line (d) illustrates the counting of arrangements for Fermi–Dirac statistics. There is only one way to place the two particles into the two states with energy ε_2; one particle must go into each state, because both particles cannot occupy the same state. There are three states available for the particle in level ε_4, so this accounts for the presence of three possible arrangements. In contrast, there are $3 \times 3 = 9$ arrangements for this distribution in Bose–Einstein statistics, because there are three ways to place the two particles into level ε_2; one in each state, two in the first state, or two in the second state.

The "probability" column in Table 1 is found simply by dividing the number of arrangements for a given distribution by the total number of arrangements. These probabilities may then be used to find the average distribution; we simply multiply the probability of each individual distribution by each n_s in that distribution, and we add the results for a given s, to find the n_s for the average distribution. You can easily verify that the average distributions are those shown in Table 2.

The total of all the values of n_s is, of course, 4 for each distribution. It may seem a bit strange that n_2 is greater than n_1 in each distribution, while the values of n_s otherwise decrease with increasing s. The reason, of course, lies in the degeneracy; g_2 is 2, whereas g_1 is only 1. It is useful to define the quantity n_s/g_s, called the *occupation index*, which is the average number of particles

Table 2

Average Distributions Derived from Table 1

	n_1	n_2	n_3	n_4	n_5	n_6
Boltzmann:	576/464	648/464	328/464	216/464	72/464	16/464
Bose–Einstein:	51/39	54/39	26/39	15/39	6/39	4/39
Fermi–Dirac:	1	8/5	4/5	3/5	0	0

per state in the sth level. The occupation index does decrease with increasing s, or increasing energy, and we shall soon see that this index is independent of g_s for all kinds of statistics. It is the occupation index that is given by the expression $e^{-\varepsilon_s/kT}$ in Boltzmann statistics. In our example, the occupation indices are as shown in Table 3.

Table 3

Average Occupation Indices Derived from Table 1

	n_1/g_1	n_2/g_2	n_3/g_3	n_4/g_4	n_5/g_5	n_6/g_6
Boltzmann	576/464	324/464	164/464	72/464	24/464	4/464
Bose–Einstein	51/39	27/39	13/39	5/39	2/39	1/39
Fermi–Dirac	1	4/5	2/5	1/5	0	0

The difference between the Boltzmann and the Bose–Einstein distributions happens to be slight in this case, but the Fermi–Dirac distribution is significantly different from either of the others.

This example should have been helpful in clarifying the manner of application of the basic principles, so that you will be prepared for the more general derivation which follows. Obviously, the distribution of four particles should not be precisely described by the statistics of large numbers, but after we have developed the form of the three general distribution laws, you may wish to return to these example distributions to test their similarity to the general laws.

11.2 DERIVATION OF THE GENERAL FORM FOR EACH DISTRIBUTION FUNCTION

We derive the general form for each of the three types of distribution by counting arrangements of N particles with total energy E, much as we did in the four-particle example. Obviously, when there are 10^{23} particles we cannot actually enumerate the arrangements, so we must be more clever; we must find a general formula for the number of arrangements in a given distribution.

Even then, there are an enormous number of distributions to consider in computing the average distribution, but we have the advantage that it is sufficient to find the *most probable* distribution rather than the average distribution. The most probable distribution, like any other *specific* distribution of 10^{23} particles, has a very small probability of occurrence. But all of the distributions which actually occur differ from the most probable one by a negligible amount. We may understand this fact by considering the standard deviation in the number of particles to be expected in any given energy level or set of levels. When there are of the order of 10^{23} particles, the number of particles in the states in any measurably large energy range could easily be of the order of 10^{16}; the standard deviation in 10^{16} is 10^{8} (see Appendix A), a large number of ordinary standards, but only one part in 10^{8}. In other words, most of the possible distributions differ from the most probable distribution by less than one part in 10^{8}, for the number of particles in any particular energy range, and we are justified in computing the most probable distribution rather than the average distribution.

We begin the computation by finding P_s, the number of ways in which the n_s particles in the sth energy level may be put into the g_s states in that level. Then we find the *total number of arrangements* for the *whole set* of given numbers $n_1, n_2, \ldots, n_s, \ldots$ by taking the product of all the numbers $P_1, P_2, \ldots, P_s, \ldots$. This product equals the statistical *weight* $W(n_1, n_2, \ldots, n_s, \ldots)$ for the distribution $n_1, \ldots, n_s \ldots$. That is,

$$W(n_1, \ldots, n_s, \ldots) = \prod_{s=1}^{\infty} P_s$$

As mentioned above, the total number W is proportional to the probability of finding the distribution to be the set $n_1, n_2, \ldots, n_s, \ldots$ at any given time.

After finding an expression for $W(n_1, n_2, \ldots, n_s, \ldots)$, we shall find the set of numbers which maximizes W, subject to the two conditions

(1) that the total number N of particles is fixed:

$$\sum_{s=1}^{\infty} n_s = N \tag{1}$$

(2) that the total energy E is fixed:

$$\sum_{s=1}^{\infty} n_s \varepsilon_s = E \tag{2}$$

Computation of Statistical Weights. The computation of P_s and W is shown below for each of the three types of statistics.

Boltzmann. In this case only, we assume that the particles are distinguishable, so we are concerned with *which* particles we choose to put in each state. Let us consider level 1 first. From the N particles available, we can choose n_1 particles for this *level* in $N!/n_1!(N-n_1)!$ different ways.[1] These

---

[1] There are N choices for the first particle, $N-1$ for the second, etc., so there are $N(N-1)(N-2)\ldots(N-n_1+1)$ ways to choose the n_1 particles in a *given order*. But we have the same n_1 particles regardless of the order of choosing them, so we must divide the above number by the number of possible orders of choice, which is $n_1!$. Thus the number of ways of choosing n_1 objects from N objects is $N!/n_1!(N-n_1)!$.

particles may then be distributed among the g_1 *states* of level 1 in $g_1^{n_1}$ different ways.[2] Thus there are a total of $P_1 = N!g_1^{n_1}/n_1!(N - n_1)!$ ways to put n_1 particles into level 1.

Having placed n_1 particles into level 1, we have $N - n_1$ remaining particles; we may choose n_2 of these for level 2 in $(N - n_1)!/n_2!(N - n_1 - n_2)!$ different ways. These particles may be distributed among the g_2 states of level 2 in $g_2^{n_2}$ ways, so there are a total of

$$P_2 = \frac{(N - n_1)!\, g_2^{n_2}}{n_2\,!(N - n_1 - n_2)!}$$

ways to put n_2 particles into level 2, *after* one has put n_1 particles into level 1.

By now it should be clear that, after one puts n_1 particles into level 1, n_2 into level 2, n_3 into level 3, and so on, up to n_{s-1} into level $s - 1$, there are in general

$$P_s = \frac{(N - n_1 - n_2 - \cdots - n_{s-1})!\, g_s^{n_s}}{n_s\,!(N - n_1 - n_2 - \cdots - n_s)!}$$

ways to put n_s particles into level s. The total number of ways to put N particles into the levels to produce a given distribution $n_1, n_2, \ldots, n_s \ldots$, must be the product $P_1P_2 \cdots P_s \cdots$. It is easy to see that the intermediate factors $(N - n_1)!$, $(N - n_1 - n_2)!$, etc. cancel in numerator and denominator, so the final result is

$$\begin{aligned} W(n_1, \ldots, n_s, \ldots) &= \prod_{s=1}^{\infty} P_s \\ &= N! \prod_{s=1}^{\infty} \frac{g_s^{n_s}}{n_s\,!} \qquad (3) \end{aligned}$$

Bose–Einstein. In the quantum statistics, there is no factor analogous to the factor $N!/n_1!(N - n_1)!$ which appears in the Boltzmann statistics, because that factor involves a choice of particles for a given level; when the particles are indistinguishable, this choice is irrelevant. Now we need only know the factor analogous to $g_1^{n_1}$—the factor which counts the number of ways in which the n_1 particles can go into the g_1 states of level 1. Again, since the particles are indistinguishable, we do not count the choices for a given particle; instead, we consider the whole set of particles at once. We draw a picture similar to Fig. 1, but we show only the states in level s (Fig. 2). There are n_s particles in the g_s states. We obtain a different arrangement each time we change the number of particles in one or more states, by moving particles

---

[2] There are g_1 places for each of the n_1 particles, making $g_1^{n_1}$ possibilities in all.

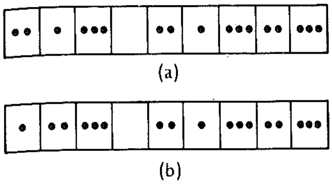

Fig. 2. (a) One arrangement of n_s particles in the g_s states of level s ($g_s = 9$). (b) A second arrangement of the same n_s particles; this arrangement is obtained from the arrangement of (a) simply by interchanging the second particle and the first partition.

from one state to another, but we obtain no new arrangement if we simply exchange two particles. We can count the arrangements by a simple trick, as follows: There are $g_s - 1$ partitions, or dividing lines, between the states, so that the total number of particles plus partitions is $n_s + g_s - 1$. We can obtain new arrangements by permuting particles and partitions; for example, the arrangement of Fig. 2b was obtained from Fig. 2a by interchanging the second particle and the first partition (between states 1 and 2). There are $(n_s + g_s - 1)!$ possible permutations of particles and partitions. However, many of these do *not* produce new arrangements; permutation of particles among themselves, or of partitions among themselves, changes nothing physically. The particles may be permuted in $n_s!$ ways, and the partitions in $(g_s - 1)!$ ways. Thus there are $n_s!(g_s - 1)!$ permutations of *each* arrangement, and the total number of arrangements must be the quotient

$$P_s = \frac{(n_s + g_s - 1)!}{n_s!(g_s - 1)!}$$

$$= \frac{\text{number of permutations}}{\text{number of permutations per arrangement}}$$

Therefore

$$W(n_1, \ldots, n_s, \ldots) = \prod_{s=1}^{\infty} P_s$$

$$= \prod_{s=1}^{\infty} \frac{(n_s + g_s - 1)!}{n_s!(g_s - 1)!} \tag{4}$$

Fermi–Dirac. Here again we need only know how many arrangements there are of the n_s particles in the g_s states of level s. But we count the arrangements in a different way, because of the condition that there is no more than one particle in each state. Because of this condition, we can simply divide the states into two groups—the n_s occupied states and the $g_s - n_s$ unoccupied states. We can choose n_s objects from a total of g_s objects in $g_s!/n_s!(g_s - n_s)!$

different ways (see the footnotes under Boltzmann statistics), so this must be the number of ways of choosing n_s *states* to be occupied when there are g_s different states. (Remember that the *states* are distinguishable, even though the *particles* are not.) Therefore

$$\begin{aligned} W(n_1, \ldots, n_s, \ldots) &= \prod_{s=1}^{\infty} P_s \\ &= \prod_{s=1}^{\infty} \frac{g_s!}{n_s!(g_s - n_s)!} \end{aligned} \tag{5}$$

Computation of n_s. We now have the three expressions for $W(n_1, \ldots, n_s, \ldots)$, the statistical weight of a given distribution $n_1, \ldots, n_s, \ldots$. In each case we can find the most probable distribution by finding the set of numbers n_s which maximizes W. We do this by maximizing the logarithm of W rather than W itself; this is equally effective and it simplifies the procedure by changing each product to a sum.

To find the maximum, we simply set the variation in ln W equal to zero as the numbers n_s are varied, just as one finds the maximum of a function of a single variable by setting the derivative equal to zero. Thus we write that the variation in ln W is

$$\delta(\ln W) = 0 \tag{6}$$

where $\delta(\ln W)$ is obtained by varying the numbers n_s by amounts δn_s and then finding the difference between the newly calculated value of W and the previous value of W. But we must remember that the variations δn_s in the numbers n_s are not arbitrary, because N and E are fixed [conditions (1) and (2)]. Therefore

$$\delta N = \sum_{s=1}^{\infty} \delta n_s = 0 \tag{7}$$

$$\delta E = \sum_{s=1}^{\infty} \varepsilon_s \, \delta n_s = 0 \tag{8}$$

Taking zero from zero leaves zero, so we may write, using Eqs. (6), (7), and (8):

$$\delta(\ln W) - \alpha \sum_{s=1}^{\infty} \delta n_s - \beta \sum_{s=1}^{\infty} \varepsilon_s \, \delta n_s = 0 \tag{9}$$

where α and β can be any numbers we choose.[3] We shall see the reason for

---

[3] This technique for maximizing a function of several variables while simultaneously satisfying certain conditions on these variables is quite widely used. The undetermined constants α and β are called Lagrange multipliers.

this step in a moment. Let us now introduce the expressions for W for the three types of statistics, and compute the occupation numbers n_s.

Boltzmann.

$$\ln W = \ln N! + \sum_{s=1}^{\infty} (n_s \ln g_s - \ln n_s!)$$

We wish to consider situations in which n_s is quite large, so we simplify this expression by using the Stirling approximation for $n!$, which for large n may be written[4]

$$\ln n! \to n(\ln n - 1) \quad \text{as} \quad n \to \infty$$

We may then write

$$\ln W = \ln N! + \sum_{s=1}^{\infty} (n_s \ln g_s - n_s \ln n_s + n_s)$$

Now, using the fact that

$$\delta(\ln W) = \sum_{s=1}^{\infty} \frac{\partial(\ln W)}{\partial n_s} \delta n_s$$

we obtain

$$\delta(\ln W) = \sum_{s=1}^{\infty} (\ln g_s - \ln n_s - 1 + 1)\, \delta n_s \tag{10}$$

Substitution of Eq. (10) into Eq. (9) yields

$$\boxed{\sum_{s=1}^{\infty} (\ln g_s - \ln n_s - \alpha - \beta\varepsilon_s)\, \delta n_s = 0} \tag{11}$$

Bose–Einstein.

$$\ln W = \sum_{s=1}^{\infty} [\ln(n_s + g_s - 1)! - \ln n_s! - \ln(g_s - 1)!]$$

Again using the Stirling approximation, we find that

$$\delta \ln W = \sum_{s=1}^{\infty} [\ln(n_s + g_s) - \ln n_s]\, \delta n_s$$

---

[4] There are more accurate forms of the Stirling approximation; see, for example, I. Sokolnikoff and R. Redheffer, "Mathematics of Physics and Modern Engineering." McGraw-Hill, New York, 1958. The expression given here is accurate to 1 percent at $n = 100$, and all expressions approach the same limit as n goes to infinity.

and substitution into Eq. (9) yields

$$\boxed{\sum_{s=1}^{\infty} [\ln(n_s + g_s) - \ln n_s - \alpha - \beta\varepsilon_s]\, \delta n_s = 0} \tag{12}$$

Fermi–Dirac.

$$\ln W = \sum_{s=1}^{\infty} [\ln g_s! - \ln n_s! - \ln(g_s - n_s)!]$$

so that the use of the Stirling approximation gives us

$$\delta \ln W = \sum_{s=1}^{\infty} [-\ln n_s + \ln(g_s - n_s)]\, \delta n_s$$

and substitution into Eq. (9) yields

$$\boxed{\sum_{s=1}^{\infty} [-\ln n_s + \ln(g_s - n_s) - \alpha - \beta\varepsilon_s]\, \delta n_s = 0} \tag{13}$$

The problem now is to eliminate the quantities δn_s from the equations. We can do this *if* the δn_s are arbitrary and can be varied *independently*, because in that case the coefficient of each δn_s must be zero in Eqs. (11), (12), and (13). Actually, all but two of the δn_s may be chosen arbitrarily, and the remaining two are then determined by conditions (1) and (2). But now the constants α and β come into play; we can *choose* these two constants so that the coefficients of the two final δn_s *are* zero. Then, since the other δn_s are all arbitrary, the coefficients of these are also zero, and we have, for *all* values of s,

Boltzmann

$$\ln \frac{n_s}{g_s} = -\alpha - \beta\varepsilon_s \qquad \boxed{\frac{n_s}{g_s} = e^{-\alpha-\beta\varepsilon_s}}$$

Bose–Einstein

$$\ln \frac{n_s}{n_s + g_s} = -\alpha - \beta\varepsilon_s \qquad \boxed{\frac{n_s}{g_s} = \frac{1}{e^{\alpha+\beta\varepsilon_s} - 1}}$$

Fermi–Dirac

$$\ln \frac{n_s}{g_s - n_s} = -\alpha - \beta\varepsilon_s \qquad \boxed{\frac{n_s}{g_s} = \frac{1}{e^{\alpha+\beta\varepsilon_s} + 1}}$$

In spite of the great differences in the assumptions used to derive them, the three distributions have a similar appearance. The Bose–Einstein and Fermi–Dirac occupation indices differ from the Boltzmann only in the presence of the -1 or the $+1$, respectively, in the denominator. But that little "1" can have enormous consequences; in order to see these consequences, we first must determine the significance of the constants α and β.

We already should know that $\beta = 1/kT$, at least in the case of Boltzmann statistics, because we have shown by other methods (Chapter 1) that the Boltzmann distribution is proportional to $e^{-\varepsilon_s/kT}$. To show that β has the same meaning in the other distributions, we consider a mixture of two kinds of particles, one obeying Boltzmann statistics, the other obeying one of the other kinds of statistics. The energy levels for the two kinds of particles may be labeled ε_s and ε'_t, respectively, with occupation numbers n_s and n'_t, respectively. The number of arrangements for a given distribution is $P(n_s, n'_t) = P(n_s) \cdot P(n'_t)$—the product of the number of arrangements of the two kinds of particle separately—and the overall distribution function for the mixture is found by maximizing $\ln P(n_s, n'_t)$, in analogy to our previous procedure. The numbers n_s and n'_t are now subject to *three* conditions rather than two; we must have

$$\sum_{s=1}^{\infty} n_s = N_1$$

$$\sum_{t=1}^{\infty} n'_t = N_2$$

$$\sum_{s=1}^{\infty} n_s \varepsilon_s + \sum_{t=1}^{\infty} n'_t \varepsilon'_t = E$$

Therefore, we introduce *three* constants, α_1, α_2, and β, into the variation equation analogous to Eq. (6):

$$\begin{aligned} &\delta[\ln W(n_1, \ldots, n_s, \ldots) + \ln W(n'_1, \ldots, n'_t, \ldots)] \\ &\qquad - \alpha_1\, \delta N_1 - \alpha_2\, \delta N_2 - \beta\, \delta E = 0 \end{aligned}$$

or

$$\begin{aligned} &\delta[\ln W(n_1, \ldots, n_s, \ldots) + \ln W(n_1, \ldots, n'_t, \ldots)] - \alpha_1 \sum_{s=1}^{\infty} \delta n_s \\ &\qquad - \alpha_2 \sum_{t=1}^{\infty} \delta n'_t - \beta \sum_{s=1}^{\infty} (\delta n_s)\varepsilon_s - \beta \sum_{t=1}^{\infty} \varepsilon'_t\, \delta n'_t = 0 \end{aligned}$$

This equation can be separated into two equations, each involving only one kind of particle, and we can then proceed to derive the same distribution functions as before. But notice that we have the *same* β for both kinds of particle. Thus if β equals $1/kT$ for Boltzmann statistics, it must equal $1/kT$ for the other kinds of statistics as well. Because only the total energy, and not the energy of each kind of particle, is absolutely fixed, we need only one β, or one temperature, to describe the system. The two kinds of particle are in thermal equilibrium, each behaving as if the other were not present.

The parameter α does not have such a simple interpretation, but it is clearly related to the number of particles in the system, as it was introduced in connection with the condition $\sum n_s = N$. The explicit evaluation of α in terms of the parameters of the system (g_s, ε_s, N, and T) depends on the type of statistics. For the Boltzmann distribution, it is easy to write an explicit expression for α, as follows:

$$N = \sum_{s=1}^{\infty} n_s$$

$$= \sum_{s=1}^{\infty} g_s e^{-\alpha} e^{-\varepsilon_s/kT}$$

therefore

$$e^{-\alpha} = \frac{N}{\sum_{s=1}^{\infty} g_s e^{-\varepsilon_s/kT}}$$

The quantity $\sum g_s e^{-\varepsilon_s/kT}$ is called the partition function of the system, and denoted by the symbol Z. Thus we may write the occupation indices for the Boltzmann distribution as

$$\frac{n_s}{g_s} = \frac{N}{Z} e^{-\varepsilon_s/kT}$$

Further inquiry into the meaning of α is best done in the course of studying applications to specific problems.

EXAMPLE PROBLEM 1. Consider a system containing only three equally-spaced energy levels, ε_1, ε_2, and ε_3, with degeneracies $g_1 = 1$, $g_2 = 2$, and $g_3 = 2$, at a temperature such that kT equals the difference between the first and second (or second and third) levels. If there are 100,000 distinguishable particles in the system, what is the most probable set of the occupation numbers n_1, n_2, and n_3?

Solution. We use Boltzmann statistics, because the particles are distinguishable. We let $\varepsilon_1 = 0$. (This choice is arbitrary.) Then $\varepsilon_2 = kT$ and $\varepsilon_3 = 2kT$. If we knew n_1, we could simply use the Boltzmann factor and the known degeneracies to find n_2 and n_3. But we are given N, so we use N and the partition function to find n_1, as follows:

$$Z = \sum g_s e^{-\varepsilon_s/kT} = 1 \cdot e^0 + 2 \cdot e^{-1} + 2 \cdot e^{-2}$$

$$= 1.000000 + 0.735759 + 0.270671$$

$$= 2.00643$$

Therefore

$$n_1 = g_1 \frac{N}{Z} e^{-\varepsilon_1/kT} = \frac{100{,}000e^0}{2.00643} = 49{,}840$$

$$n_2 = g_2 \frac{N}{Z} e^{-\varepsilon_2/kT} = \frac{200{,}000e^{-1}}{2.00643} = 36{,}670$$

$$n_3 = g_3 \frac{N}{Z} e^{-\varepsilon_3/kT} = \frac{200{,}000e^{-2}}{2.00643} = 13{,}490$$

$$n_1 + n_2 + n_3 = 100{,}000$$

Let us verify that the total number of arrangements, $W(n_1, n_2, n_3)$, is maximum for this set of the n_s relative to other possible sets. To do this, we wish to vary n_1, n_2, and n_3 in such a way that the total energy and the total N do not change. We can do this simply by making $\delta n_1 = \delta n_3 = -\delta n_2/2$. So let us set δn_1 first equal to $+1$ and then to -1, in order to generate two new sets of numbers—n_1', n_2', n_3', and n_1'', n_2'', n_3''—such that

$$n_1' = 49{,}841; \qquad n_2' = 36{,}668; \qquad n_3' = 13{,}491$$

and

$$n_1'' = 49{,}839; \qquad n_2'' = 36{,}672; \qquad n_3'' = 13{,}489$$

For the respective sets of numbers the values of W are given by

$$W(n_1, n_2, n_3) = \frac{N!}{n_1!\,n_2!\,n_3!} g_1^{n_1} g_2^{n_2} g_3^{n_3} = \frac{100{,}000!}{49840!\,36670!\,13490!} 2^{36670} 2^{13490}$$

$$W(n_1', n_2', n_3') = \frac{100{,}000!}{49841!\,36668!\,13491!} 2^{36668} 2^{13491}$$

$$W(n_1'', n_2'', n_3'') = \frac{100{,}000!}{49839!\,36672!\,13489!} 2^{36672} 2^{13489}$$

Obviously, we are not about to compute each of these W, but we can find their ratios without too much difficulty. You can easily see that

$$\frac{W(n_1, n_2, n_3)}{W(n_1', n_2', n_3')} = \frac{49841 \times 13491 \times 2}{36670 \times 36669} \approx \frac{134481}{134465} \approx 1.0001$$

and

$$\frac{W(n_1, n_2, n_3)}{W(n_1'', n_2'', n_3'')} = \frac{36672 \times 36671}{49840 \times 13490 \times 2} \approx \frac{134480}{134468} \approx 1.0001$$

Thus either increasing or decreasing n_2 leads to the same result, that the number of arrangements becomes smaller[5]; it seems that the distribution we found must indeed be the most probable one, given the assumption that each arrangement is equally likely.

11.3 APPLICATIONS OF BOSE–EINSTEIN STATISTICS

a. The Planck Black-Body Law Rederived. There is a striking similarity between the Bose–Einstein distribution

$$n_s = \frac{g_s}{e^{\alpha}e^{\varepsilon_s/kT} - 1}$$

and the Planck expression for the average energy of the oscillators in the walls of a cavity (Section 3.1):

$$\bar{\varepsilon} = \frac{h\nu}{e^{h\nu/kT} - 1} \tag{14}$$

This is somewhat surprising, because you will recall that the Planck law was derived by application of *Boltzmann* statistics to the energy levels of oscillators in the walls of the cavity. But *Bose–Einstein* statistics can just as well be applied to the energy levels of the *radiation field* in the cavity, because *photons* obey Bose–Einstein statistics.

You will recall that the radiation field in the cavity consists of standing waves of various frequencies. A standing wave of frequency ν can have energy $nh\nu$, where n is an integer; one can say that such a mode of oscillation "contains" n photons. Thus we can think of the mode as a *state* which is "occupied" by n photons, and we can apply statistics to the photons as if they were particles that are free to occupy various states. If a photon of frequency ν is absorbed by a cavity wall, and a photon of frequency ν' is emitted by the wall, the result is as if a single particle "dropped" from an energy level of $h\nu$ to an energy level of $h\nu'$, with the difference in energy, $h(\nu - \nu')$, being given to the wall of the cavity.

To describe the distribution of photons in the various energy states, or modes, we use Bose–Einstein statistics, as follows. First we note that $\alpha = 0$, because the number of photons is not fixed.[6] If N is not fixed, then $\sum \delta n_s \neq 0$,

---

[5] Notice that the three-level system is particularly well suited for this illustration, because only one n is independent when N and E are fixed. Thus we need not worry about the choice of possible variations among the n; the variation in any one of them, for example, in n_2, automatically determines the variations in the other two.

[6] If the cavity is to be at a fixed temperature, it is only necessary that *energy* be conserved. E may be conserved while N changes; for example, one photon of frequency ν may be replaced by two photons of frequency $\frac{1}{2}\nu$.

and to make the term $\alpha \sum \delta n_s$ equal to zero [to satisfy Eq. (9)], we must set α equal to zero. The number of photons of energy ε_s is therefore, according to the Bose–Einstein distribution function,

$$n_s = \frac{g_s}{e^{\varepsilon_s/kT} - 1}$$

To go from here to Eq. (14) is quite simple; n_s/g_s is the number of photons per mode of oscillation, and the energy of each photon is $\varepsilon_s = h\nu$, so the average energy per mode is $h\nu n_s/g_s$, which from the above equation is just $h\nu/(e^{h\nu/kT} - 1)$, as required.

If we wish to consider the energy to be a continuous variable, we may write an equation for the number of photons per unit energy interval as

$$n(\varepsilon) = \frac{g(\varepsilon)}{e^{\varepsilon/kT} - 1} \tag{15}$$

where $\int_{E_1}^{E_2} n(\varepsilon)\, d\varepsilon$ is the number of photons with energy between E_1 and E_2, and $\int_{E_1}^{E_2} g(\varepsilon)\, d\varepsilon$ is the number of states with energy between E_1 and E_2.

Let us complete the black-body analysis by computing the density of states $g(\varepsilon)$ for the photons. The resulting black-body spectrum is, of course, the same as that derived in Section 3.1, but the point of view is different. The derivation is repeated here in this slightly modified form because this point of view is also applicable to subsequent topics in this chapter.

We begin the computation by observing that each photon has x, y, and z components of momentum, corresponding to wavelengths in the x, y, and z directions, and each wavelength must satisfy the boundary conditions on **E** (Section 3.1). Therefore, if the cavity is a cube of side a, each wavelength is equal to $2a$ divided by an integer, and the momentum components of a given photon may be written

$$\begin{aligned} p_x &= \frac{h}{\lambda_x} = \frac{hl_x}{2a} \\ p_y &= \frac{h}{\lambda_y} = \frac{hl_y}{2a} \\ p_z &= \frac{h}{\lambda_z} = \frac{hl_z}{2a} \end{aligned} \tag{16}$$

where l_x, l_y, and l_z are positive integers. The photon energy is therefore given by

$$\begin{aligned} \varepsilon^2 &= c^2 p^2 \\ &= c^2(p_x^2 + p_y^2 + p_z^2) \\ &= \frac{c^2h^2}{4a^2}(l_x^2 + l_y^2 + l_z^2) \end{aligned} \tag{17}$$

Comparison of Eq. (17) with Eq. (4) of Chapter 3 shows that this is the same result obtained there (with $\varepsilon = h\nu$).

Following the method of Section 3.1, we see from Eq. (17) that all sets of integers for which the photon energy is ε or less obey the relation

$$l_x^2 + l_y^2 + l_z^2 \leqslant \frac{4a^2\varepsilon^2}{h^2c^2} \tag{18}$$

If we consider the numbers l_x, l_y, and l_z to be coordinates of a point in three dimensions, the points corresponding to the numbers which satisfy Eq. (18) must lie within a sphere of radius $2a\varepsilon/hc$. The number of sets of positive integers within a sphere of radius R is just $\frac{1}{8}$ of the volume of the sphere, or $\frac{1}{8}4\pi R^3/3$, and there are two modes of oscillation (two polarization directions) for each set of integers, so the number G of photon states of energy ε or less is

$$G = 2 \cdot \frac{1}{8} \cdot \frac{4\pi}{3} \left(\frac{2a\varepsilon}{hc}\right)^3$$

The density of states must then be

$$\begin{aligned} g(\varepsilon) &\equiv \frac{dG}{d\varepsilon} \\ &= \frac{8\pi V\varepsilon^2}{h^3c^3} \qquad (V = a^3) \end{aligned} \tag{19}$$

Therefore, from Eq. (15),

$$n(\varepsilon) = \frac{8\pi\, V\varepsilon^2}{h^3c^3(e^{\varepsilon/kT} - 1)}$$

The energy density dU in the cavity for radiation of frequency between ν and $\nu + d\nu$ is equal to the product of the number of photons per unit volume, $n(\varepsilon)\, d\varepsilon/V$, and the energy $\varepsilon = h\nu$ of each photon. (Remember that a *mode* has energy $nh\nu$, but a *photon* has energy $h\nu$.) Thus

$$\begin{aligned} dU &= \frac{8\pi\varepsilon^3\, d\varepsilon}{h^3c^3(e^{\varepsilon/kT} - 1)} \\ &= \frac{8\pi h\nu^3\, d\nu}{c^3(e^{h\nu/kT} - 1)} \end{aligned}$$

and we see that the black-body spectrum can be derived either by applying Boltzmann statistics to the *oscillators*, as we did in Chapter 3, or by applying Bose–Einstein statistics to the photons of the radiation field. In our next example we shall see that we can do the same thing with the vibrations of a solid; the vibrations of the lattice are analogous to the electromagnetic waves in a cavity. The energy of a specific mode of the lattice vibrations is quantized, the energy levels again differing by $h\nu$ (simply because these are the energy levels of a simple harmonic oscillator).

So we can say that a mode of frequency ν, with energy $nh\nu$, is a state containing n particles called *phonons*, and we can treat the phonons just as we treated the photons in the cavity. A pho*n*on is exactly analogous to a pho*t*on; a phonon carries the energy and momentum of a lattice vibration—a sound wave—through a solid, just as a photon carries the energy and momentum of a light wave. The concept of a phonon is helpful in understanding our second application of Bose–Einstein statistics, which is the specific heat of a solid.

b. The Specific Heat of a Solid. Einstein was the first to point out, in 1907,[7] that the Planck quantization of the energy of oscillators in a cavity should be applicable to the atoms in any solid. This quantization should have an observable effect on the specific heat. It was already known that the classical specific heat law (law of Dulong and Petit) was not obeyed by all solids. According to this law, solids have a constant (temperature-independent) specific heat of $3R$ per mole, where R is the gas constant. This is to be expected classically, because one mole of a solid, with N_A atoms, has $3N_A$ modes of vibration of the lattice (because each atom is free to oscillate in three independent directions), and if each mode has an average energy of kT, the total energy is $3N_AkT$, making the specific heat $3N_Ak$, or $3R$.[8] But, as Einstein pointed out, the specific heat is less than this for many solids at room temperature, and the specific heat becomes smaller as the temperature is reduced. Einstein saw the similarity of this behavior to the behavior of black-body radiation: as T becomes smaller, the spacing of energy levels becomes larger relative to kT, so that quantization is more effective in reducing the average energy per mode of oscillation below kT. As the average energy is reduced, the specific heat goes down. Thus the failure of the classical specific heat law as T is *reduced* is analogous to the failure of the Rayleigh–Jeans law as the frequency is *increased*; both laws fail when $h\nu/kT$ becomes appreciable.

Einstein did not intend to develop a complete theory of solids. He merely wanted to show the similarity of this quantum effect to the effect seen in black-body radiation. So he simply assumed that the $3N_A$ modes of oscillation of the solid all have the same frequency ν, and he replaced the classical average energy kT by the average energy given by the Planck formula (14):

$$\bar{\varepsilon} = \frac{h\nu}{e^{h\nu/kT} - 1}$$

---

[7] A. Einstein, *Ann. Phys.* (*Leipzig*) **22**, 180 (1907); **34**, 170 (1911).

[8] You may wonder why the electrons' contribution to the specific heat is negligible. We shall answer that question when we discuss applications of Fermi–Dirac statistics.

The total energy of N_A oscillating atoms should then be $3N_A h\nu/(e^{h\nu/kT} - 1)$ which, of course, reduces to $3N_A kT$ if $h\nu \ll kT$. Thus the fact that most solids do obey the Dulong–Petit law at room temperature could be explained by saying that for these solids, the characteristic frequency ν must be much smaller than kT/h. The solids for which the law was not obeyed were assumed to differ from the others only in the value of ν.

If Einstein's assumptions were correct, one should be able to fit the specific heat data for any solid simply by choosing ν correctly for that solid; the specific heat should be the same function of T/ν for *all* solids. This function can be easily computed from the above formula for the total energy; it has the interesting feature that it goes to zero as T goes to zero.

At about the same time that Einstein's paper appeared, Nernst began a series of measurements of specific heats of solids at low temperatures, to test his belief that molar specific heats of all solids should approach the same value (not necessarily zero) as T approaches zero. When he later compared his results with the curve predicted by Einstein's model, he found serious deviations at low temperatures; although the specific heat did approach zero as T went to zero, the temperature dependence was not that of Einstein's formula.

Nernst and Lindemann tried to fit the data by assuming that there are several characteristic frequencies instead of just one for a given solid, but such an empirical approach could not be very convincing. Obviously, if we have a few curves to fit, we can do if it we use a sufficient number of independent parameters, but no fundamental understanding is gained in this way. In 1912, P. Debye found a more general approach which was quite successful in explaining the data.[9] Einstein had assumed that the atoms were vibrating independently, but Debye considered the possibility of collective motions. He recognized that the normal modes of oscillation of the atoms in a crystal lattice can have *many* different characteristic frequencies, and that the problem was to decide, on other than empirical grounds, what the frequency *distribution* should be for a given solid. He then made the reasonable assumption that the problem is similar to the problem of electromagnetic oscillations in a cavity, so that the number of modes per unit frequency range is given by $dN/d\nu = A\nu^2$, where A is a constant. This result follows from Eq. (19), if one assumes that the pho*t*on gas in a cavity has the same modes of vibration as the pho*n*on gas in a solid. But the phonon gas does differ from the photon gas, in that it has a limited number—$3N_A$—of modes of oscillation, and consequently there must be an upper limit on the frequency of a mode. We can write the constant A in terms of this upper limit ν_m as follows:

$$3N_A = \int_{\nu=0}^{\nu=\nu_m} dN = \int_0^{\nu_m} A\nu^2 \, d\nu = \frac{A\nu_m^3}{3}$$

---

[9] P. Debye, *Ann. Phys. (Leipzig)* **39**, 789 (1912). Debye later (1936) won the Nobel Prize in chemistry for his work on molecular structure.

therefore

$$A = \frac{9N_A}{\nu_m^3}$$

The energy of the dN modes whose frequency lies between ν and $\nu + d\nu$ is therefore $dN\,\bar{\varepsilon}$, where $\bar{\varepsilon}$, the average energy of a mode whose frequency is ν, is given by Eq. (14). And the total energy[10] is simply

$$\begin{aligned} E &= \int_{\nu=0}^{\nu=\nu_m} \bar{\varepsilon}\, dN \\ &= \int_0^{\nu_m} \frac{h\nu}{e^{h\nu/kT} - 1} \frac{9N_A}{\nu_m^3} \nu^2\, d\nu \end{aligned}$$

We can simplify this expression by substituting the dimensionless variable $x = h\nu/kT$ into the integral, which then becomes

$$\begin{aligned} E &= \frac{9N_A k^4 T^4}{h^3 \nu_m^3} \int_0^{h\nu_m/kT} \frac{x^3}{e^x - 1}\, dx \\ &= 9N_A kT \left(\frac{T}{\Theta_D}\right)^3 \int_0^{\Theta_D/T} \frac{x^3\, dx}{e^x - 1} \qquad (20) \end{aligned}$$

where $\Theta_D \equiv h\nu_m/k$ is called the *Debye temperature* of a solid.

When the temperature T is much greater than the Debye temperature, the variable x is much smaller than 1 over the range of integration. In that case the integrand becomes approximately $x^3/(1 + x - 1) = x^2$, and the energy of the solid becomes $E = 3N_A kT$, which is just the classical expression leading to a specific heat of $3R$ per mole. So the solids which obey the Dulong–Petit law are those for which Θ_D is small, compared to the temperature at which the specific heat is measured.

At low temperatures, that is, when $T \ll \Theta_D$, the upper limit in the integral of Eq. (20) is very large, and the integral approaches the limit

$$\int_0^\infty \frac{x^3}{e^x - 1}\, dx = \frac{\pi^4}{15}$$

At these temperatures, E must then be proportional to T^4, and the specific heat $\partial E/\partial T$, should be proportional to T^3. Experiments on a wide variety of solids has verified this temperature dependence for the specific heat at low temperatures, and Fig. 3 shows that the Debye theory also agrees with experiment over a wide range of temperatures for several solids. In each case,

---

[10] We are neglecting the ground state, or "zero point" energy of $h\nu/2$ for each mode (see Problem 4). This energy, of course, makes no contribution to the specific heat.

Θ_D is determined from the experimental data, but it is impressive that this is the only parameter needed to fit all the points. The value of Θ_D obtained in this way is also in agreement with values of the maximum vibration frequency ν_m obtained by independent methods, such as measurement of elastic constants. But the theory is not perfect; the weakness in the theory is in its assumption that the crystal is a continuous medium, like the vacuum in a cavity. As expected, the simple theory breaks down for anisotropic materials, where the spectrum of frequencies is much more complicated than the simple ν^2 dependence; more careful measurements have shown that it also breaks down in other materials. The assumption of a ν^2 dependence is a great oversimplification, and a great deal of effort has been devoted to the determination of the actual "phonon spectrum" in many solids and to the development of "modified Debye theories" which are based on different sets of normal modes. Nevertheless, the simple Debye theory is a remarkably accurate and useful first approximation.

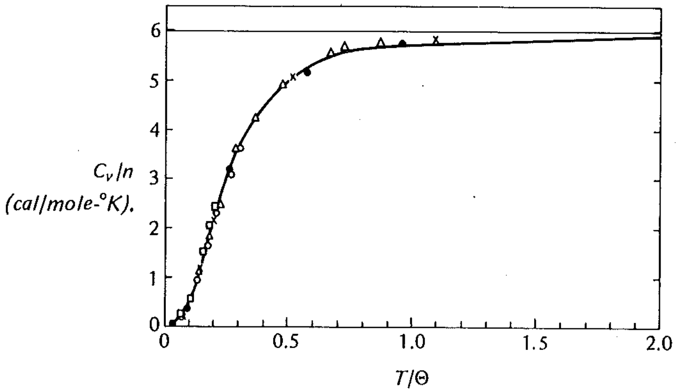

Fig. 3. Comparison of the Debye specific heat curve and the observed specific heats of a number of simple substances. (●) Ag; (△) Al; (□) *C(graphite)*; (○) Al_2O_3; (×) KCl. [*From "The Modern Theory of Solids" by F. Seitz. Copyright McGraw-Hill, New York, 1940. Used by permission of McGraw-Hill Book Company.*]

c. Liquid Helium and Superfluidity. Our first two examples have shown us how to use Bose–Einstein statistics as the basis for a calculation; however, they did not bring out some distinctive features of these statistics, because they are somewhat special applications in which the number of particles is not constant. In fact, it is not even necessary to use Bose–Einstein statistics to solve those problems, because we can apply classical Boltzmann statistics to the oscillators instead of using Bose–Einstein statistics for the photons or phonons; we have already (Section 3.1) seen how this is done for black-body radiation.

But application of Bose–Einstein statistics to a system of material particles, in which N is constant, reveals a startling new possibility, which had been suggested by Einstein in 1924. Einstein pointed out that at low temperatures, a gas of Bose–Einstein particles would undergo a "condensation" which is totally different from the ordinary gas–liquid condensation. In this condensation, a large fraction of particles would occupy the lowest energy state. In a large scale system in which the quantum number is zero, the correspondence principle would no longer apply as it does to most large systems, and thus the *mechanical* behavior of the system would no longer be correctly described by classical mechanics. Quantum effects might then be visible, and would not have to be deduced from indirect evidence.

To see why this condensation should be peculiar to Bose–Einstein statistics, consider a system containing N particles, with a set of nondegenerate energy levels ($g_s = 1$ for all s). The number of possible distributions becomes enormous for large N, but compare just two possibilities:

(a) There is one particle in each of the lowest N levels.

(b) There are $N - 1$ particles in the lowest level, and one particle in a higher level, with energy equal to that of the $N - 1$ excited particles in (a).

In Bose–Einstein statistics, both distributions are equally likely, each containing one arrangement. But classically, distribution (a) contains $N!$ arrangements while (b) contains only N arrangements, so that (a) occurs $(N - 1)!$ times as often as (b). Thus a distribution like (b), with a large fraction of the particles in the ground state, is always highly improbable in classical statistics; but we can see the possibility that, at low temperatures, such distributions may begin to make their presence felt in a Bose–Einstein gas.

Einstein's idea was intriguing, but apparently unrealistic, because at the very low temperature (about 3°K) at which such a condensation should be expected, no known substance remains in gaseous form. The condensation had been deduced from the properties of an *ideal gas*, in which one assumes no interaction between the particles of the gas; but clearly, when a substance is held together in the liquid or solid state, the interaction between particles must be considerable. In 1938, however, Fritz London pointed out[11] that many peculiar properties of liquid ^{4}He could be explained if it is treated as a Bose–Einstein "gas," even though it is in the liquid state.

---

[11] F. London, *Phys. Rev.* **54**, 947 (1938). See also London's "Superfluids," Vol. II. Wiley, New York, 1954.

^{4}He atoms, having total angular momentum of zero, obey Bose–Einstein statistics. The normal condensation into a liquid occurs at 4.2°K, at atmospheric pressure. The temperature can be reduced below that point by reducing the pressure. As the pressure is reduced, the liquid boils until the temperature reaches about 2.2°K. At this point, called the lambda (λ) point,[12] *boiling* suddenly ceases, although *evaporation* continues. Of course, boiling occurs in the first place because the liquid is not at a uniform temperature throughout, and bubbles form at the "hot" points. The cessation of boiling indicates a sharp increase in thermal conductivity; in fact, the thermal conductivity appears infinite, for all practical purposes, so that the liquid is always at a perfectly uniform temperature throughout.

Below the λ point, liquid ^{4}He has another remarkable property: it can penetrate through the tiniest capillary, as if the viscosity were zero. But when the viscosity is measured by means of a torsion pendulum, by observing the drag on a set of parallel plates moving through the fluid, the viscosity is *not* zero, and it shows no discontinuity as a function of temperature at the λ point.

To account for these properties, London suggested that liquid ^{4}He below the λ point (called liquid He II) consists of two interpenetrating fluids—a "normal" fluid and a "superfluid." The superfluid is that part of the liquid whose atoms are in the ground state; the normal fluid is the rest. The normal fluid causes the drag on a torsion pendulum; the superfluid seeps through capillaries. It appears that the atoms of the superfluid, being in a state of almost perfectly defined momentum, cannot be localized in space. Heating the liquid destroys superfluid; but the superfluid, which cannot be confined to one region of the liquid, is destroyed uniformly throughout the entire volume of the liquid whenever heat is applied to any point. Destruction of superfluid is equivalent to a rise in temperature; since this rise takes place everywhere (almost) simultaneously, the fluid appears to have infinite thermal conductivity.

In a moment we shall discuss further the reason for this strange behavior of atoms in the ground state. Right now let us look more carefully into the properties of an *ideal* Bose–Einstein gas at low temperatures, to see just how it is that so many atoms condense into the ground state, and why this condensation makes its presence felt at a well defined nonzero temperature. We begin with the Bose–Einstein distribution numbers

$$n_s = \frac{g_s}{e^{\alpha} e^{\varepsilon_s/kT} - 1} \tag{21}$$

---

[12] Called the λ point because the curve of specific heat versus temperature resembles the letter λ near this temperature.

The zero level of energy is arbitrary, so let us set the energy of the lowest level equal to zero; that is, we let $\varepsilon_1 = 0$. In this case, it is clear that $\alpha > 0$, for if α were zero, as in the cases when the number of particles is not fixed, the value of n_1 would be infinite, and if α were negative, n_1 would be negative, according to Eq. (21), and a negative value of n_1 makes no sense.

Now in dealing with a macroscopic sample, the usual procedure is to replace n_s by a continuous function $n(\varepsilon)$, as we did in parts a and b of this section, because in a macroscopic system the energy levels may be spaced so closely that their discrete nature is not observable. But we must be careful when we are dealing with low temperatures, because the total energy involved is so small that the discreteness of the levels may be important even in a macroscopic system. With this in mind, let us proceed to a continuum description and see what happens. We write, in place of Eq. (21),

$$n(\varepsilon) = \frac{g(\varepsilon)}{e^{\alpha}e^{\varepsilon/kT} - 1} \tag{22}$$

an equation similar to Eq. (15), except for the factor e^{α}.

We may find $g(\varepsilon)$ by the same basic method used for photons. In a cube of side a, the permissible *wavelengths* are the same for material particles as for photons, and there is the same connection between wavelength and *momentum* in both cases. Therefore the *momentum* components of a given particle are given by Eqs. (16). But if the particle has mass M, the momentum-*energy* relation at low energies is $\varepsilon = p^2/2M$ rather than $\varepsilon = pc$, and Eq. (17) is replaced by

$$\varepsilon = \frac{h^2(l_x^2 + l_y^2 + l_z^2)}{8Ma^2}$$

The integers for which the particle energy is ε or less obey the relation

$$l_x^2 + l_y^2 + l_z^2 \leqslant \frac{8Ma^2\varepsilon}{h^2}$$

and the points corresponding to these integers lie within a sphere of radius $(8Ma^2\varepsilon/h^2)^{1/2}$. The number G of states within this sphere is equal to $\frac{1}{8}$ of the volume of the sphere, or

$$G = \frac{1}{8} \cdot \frac{4\pi}{3} \left(\frac{8Ma^2\varepsilon}{h^2}\right)^{3/2} \tag{23}$$

and the density of states is

$$g(\varepsilon) = \frac{dG}{d\varepsilon} = \frac{3}{2} \cdot \frac{1}{8} \cdot \frac{4\pi}{3}\left(\frac{8Ma^2}{h^2}\right)^{3/2} \varepsilon^{1/2}$$

or, with $a^3 = V$,

$$g(\varepsilon) = 2\pi V \left(\frac{2M}{h^2}\right)^{3/2} \varepsilon^{1/2} \tag{24}$$

Now we may find a relation between N and α by using the fact that

$$N = \int_0^\infty n(\varepsilon)\, d\varepsilon = \int_0^\infty \frac{g(\varepsilon)\, d\varepsilon}{e^\alpha e^{\varepsilon/kT} - 1}$$

$$= 2\pi V \left(\frac{2M}{h^2}\right)^{3/2} \int_0^\infty \frac{\varepsilon^{1/2}\, d\varepsilon}{e^\alpha e^{\varepsilon/kT} - 1}$$

Let us make the substitutions $A = 2\pi V(2Mk/h^2)^{3/2}$ and $x = \varepsilon/kT$ to obtain

$$N = AT^{3/2} \int_0^\infty \frac{x^{1/2}\, dx}{e^{\alpha+x} - 1} \tag{25}$$

We may rewrite Eq. (25) as

$$N = AT^{3/2} \int_0^\infty \frac{x^{1/2}}{e^{\alpha+x}} (1 - e^{-\alpha-x})^{-1}\, dx$$

which may be expanded to give

$$N = AT^{3/2} \int_0^\infty x^{1/2} e^{-\alpha-x} (1 + e^{-(\alpha+x)} + e^{-2(\alpha+x)} + \cdots)\, dx$$

and integrated term by term. If one remembers the definition of the gamma function:

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

the integral may be written

$$N = AT^{3/2} \Gamma\left(\frac{3}{2}\right) \left[e^{-\alpha} + \frac{e^{-2\alpha}}{2^{3/2}} + \frac{e^{-3\alpha}}{3^{3/2}} + \cdots\right]$$

or

$$N = AT^{3/2} \Gamma\left(\frac{3}{2}\right) f(\alpha) \tag{26}$$

where

$$f(\alpha) = \sum_{p=1}^{\infty} \frac{e^{-p\alpha}}{p^{3/2}}$$

The series sum $f(\alpha)$ is plotted versus α in Fig. 4. We may use Eq. (26) to determine the value of α for a given temperature, as follows: We simply use the values of N, A, $\Gamma(\frac{3}{2})$, and T to compute the value of $f(\alpha)$, and we then read off the corresponding value of α from Fig. 4. (*Note*: $\Gamma(\frac{3}{2}) = \sqrt{\pi}/2 \approx 0.88623$.)

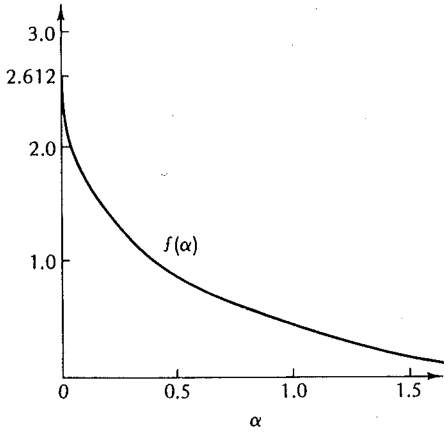

Fig. 4. The function $f(\alpha)$ appearing in Eq. (26).

But there is a serious difficulty as T approaches zero. Since N, A, and $\Gamma(\frac{3}{2})$ are all independent of T, we have, directly from Eq. (26),

$$f(\alpha) = \frac{N}{AT^{3/2}\Gamma(\frac{3}{2})} \to \infty \qquad \text{as} \quad T \to 0.$$

But $f(\alpha)$ is a mathematical function whose largest value, according to Fig. 4, is 2.612, at $\alpha = 0$,[13] so Eq. (26) *cannot possibly be correct* in the limit $T \to \infty$.

What went wrong? Remember the warning about the passage from Eq. (21) to Eq. (22). There is one obvious discrepancy between Eq. (21) and our later equations: We found that $g(\varepsilon)$ is proportional to $\sqrt{\varepsilon}$, so that $g(\varepsilon) = 0$ at $\varepsilon = 0$, but obviously a value of the degeneracy smaller than one does not make sense physically. In fact, we know that the degeneracy of the lowest level is $g_1 = 1$; for this level $l_x = l_y = l_z = 1$. Therefore this level is not represented in Eq. (22), because it has a weight $g(\varepsilon) = 0$ when we use the continuous form for the degeneracy g. We can correct this situation simply by writing n_1 separately in the expression for N:

$$N = n_1 + \int_0^\infty n(\varepsilon)\, d\varepsilon = \frac{1}{e^\alpha - 1} + \int_0^\infty \frac{g(\varepsilon)\, d\varepsilon}{e^\alpha e^{\varepsilon/kT} - 1}$$

[13] Remember that α can never be negative.

or

$$N = \frac{1}{e^\alpha - 1} + AT^{3/2}\Gamma(\tfrac{3}{2})f(\alpha) \tag{27}$$

an equation identical to Eq. (26) except for the additional term $1/(e^\alpha - 1) = n_1$ on the right-hand side. As $T \to 0$, the second term of Eq. (27) goes to zero, but the first term has no upper limit as $\alpha \to 0$, so there is always some positive value of α which satisfies the equation. Notice that α may now be expressed in terms of n_1, for if

$$n_1 = \frac{1}{e^\alpha - 1}$$

then

$$e^\alpha = 1 + \frac{1}{n_1}$$

and

$$\alpha \simeq \frac{1}{n_1} \qquad (n_1 \gg 1) \tag{28}$$

We would now like to determine the temperature T_c at which the system begins to "condense" into the ground state—that is, where n_1 becomes very large. For $T > T_c$ we assume that n_1 is negligible compared to N, because the particles are spread out over an enormous number of levels, and there is no loss of accuracy in using Eq. (26) instead of the more correct Eq. (27). We may then *define* T_c as *the temperature below which Eq.* (26) *can no longer be satisfied by any value of* α, and we may determine the value of T_c simply by setting $f(\alpha)$ in Eq. (27) equal to its maximum value, 2.612, and solving for T:

$$T = T_c = \left(\frac{N}{2.612A\Gamma(\frac{3}{2})}\right)^{2/3} \tag{29}$$

We may use Eq. (27) to study temperatures below T_c. Since $f(\alpha) = 2.612$ at $T = T_c$, we might conclude that $\alpha = 0$ at $T = T_c$, but this cannot be correct, because according to Eq. (28), n_1 becomes *infinite* at $\alpha = 0$. However, we are considering systems in which the total number of particles N is of the order of 10^{23}, so that α is certainly very close to zero when n_1 becomes at all comparable to N. For example, suppose that n_1 is "only" 10^{10}; then $\alpha \simeq 10^{-10}$, and $f(\alpha) \simeq f(0) = 2.612$ to much better than four-place accuracy. Yet in this case n_1 is certainly *negligible* compared to N, so that Eq. (26) is still valid! As n_1 increases, α becomes still smaller, so that whenever it is necessary to use Eq. (27) instead of Eq. (26) we may set $f(\alpha)$ equal to 2.612, even though α is not *quite* zero. With $f(\alpha) = 2.612$, Eq. (27) becomes

$$N = n_1 + 2.612AT^{3/2}\Gamma(\tfrac{3}{2}),$$

or, from Eq. (29),

$$N = n_1 + N\left(\frac{T}{T_c}\right)^{3/2}$$

so that

$$n_1 = N\left\{1 - \left(\frac{T}{T_c}\right)^{3/2}\right\} \tag{30}$$

Figure 5 shows the temperature dependence of n_1, as given by Eq. (30).

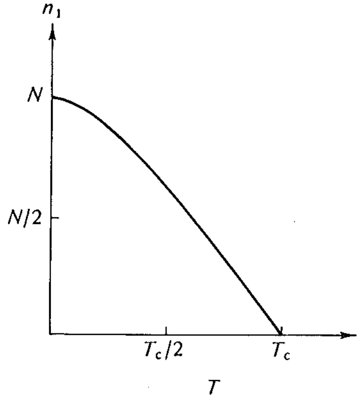

Fig. 5. *Temperature dependence of n_1 for an ideal Bose–Einstein gas. When $T > T_c$, n_1 is not necessarily zero, but it is so "small" (perhaps of order 10^{10} or less) that it does not show on this graph.*

If we use the density and particle mass of liquid ^{4}He to compute the constant A/N of Eq. (29), we obtain for the critical temperature $T_c = 3.13$ K instead of the actual value of 2.2 K. But the analysis here applies only to an *ideal* Bose-Einstein gas, in which there is no interaction between particles, so we should not expect it to be *numerically* correct for liquid helium. Interaction between the helium atoms certainly affects the energy levels, so that $g(\varepsilon)$ is not given correctly by Eq. (24). However, the derivation does show how a condensation into the ground state can occur, and it shows that Bose-Einstein statistics must have some relevance to the situation.

There have been attempts to explain the properties of liquid He (II) without reference to Bose–Einstein statistics. But convincing evidence of the role of Bose–Einstein statistics is found in the behavior of liquid ^{3}He, an isotope which one would expect to behave very much like ^{4}He, and which fulfills this expectation except in its superfluid behavior. Liquid ^{3}He is a superfluid only at temperatures below 0.0025 K. The difference results from the fact that the atoms of ^{3}He have spin $\frac{1}{2}$ rather than spin zero, so they obey Fermi–Dirac rather than Bose–Einstein statistics. This makes condensation into the ground state impossible until the temperature becomes so low that ^{3}He atoms form *pairs* whose *total* spin is integral.[14]

Thus it seems that condensation into the ground state is a necessary condition for the occurrence of superfluid behavior. But it is not a *sufficient* condition. We have not yet answered the question of why such atoms behave as a superfluid, and the answer depends on the fact that there are interactions between the particles. An *ideal* Bose–Einstein gas, even though it condenses into the ground state, does *not* become a superfluid. To understand this, we must begin by considering why ordinary fluids are not superfluids.

A fluid flowing through a tube is slowed down because interactions between the fluid and the wall of the tube convert the fluid's translational energy into internal energy—that is, into random motion or heat. This occurs in an apparently continuous fashion, because in a normal fluid there are many possible internal motions which require very little energy to excite them. *Superfluid* behavior requires that there be no low-energy states which can be easily excited—that there be a gap between the ground state and the lowest state which can be easily excited by friction with the walls of the tube. Then if the fluid flows sufficiently slowly through the tube, the atoms cannot acquire enough energy from the walls to enable them to cross this energy gap; they stay in the ground state, and their translational motion continues unimpeded. (We still consider the atoms moving through the tube to be in the ground state, even though they have translational kinetic energy; we can consider the atoms to be at rest in a different frame of reference, and the tube to be moving past them.) As long as the speed of the fluid is below a certain "critical velocity," superfluid flow continues. If the speed exceeds the critical velocity, the fluid–tube interaction becomes able to excite turbulent states of the fluid, and the fluid is slowed down.

Precise determination of the nature of the excited states and the magnitude of the energy gap is difficult. It has been observed that a critical velocity exists for superfluid flow of liquid helium, but it is difficult to calculate theoretically what its value should be; oversimplified calculations can lead to a value of the order of 100 times too large. However, it is not too hard to show

[14] D. D. Osheroff, N. J. Gully, R. C. Richardson, and D. M. Lee, *Phys. Rev. Letters* **28**, 885 (1972). The magnetic properties indicate that the total spin is one rather than zero. See "The Helium Liquids," J. G. Armitage & I. E. Farquhar, eds., Academic Press, N.Y., 1975.

qualitatively that an energy gap exists, and that its existence depends upon an interaction between the helium atoms.

The interaction potential between two helium atoms has been calculated by quantum mechanics to be as shown in Fig. 6. Notice the sharp rise with decreasing r near 2.5 Å, as if a helium atom has a very hard "core" which repels other helium atoms. Now consider the ground state and first excited state of a collection of N helium atoms, taking the wavefunctions to be the unperturbed functions discussed previously—single-particle states for a particle in a box. The first excited state of the *system* is one for which all but one of the *particles* is in the ground state, and one particle is in the first excited state. (Here, as elsewhere in this section, we are considering the atoms to be particles; we are not considering internal energy levels of the atoms, but rather levels of the *particles* in the "box.") We state without proof that, because of the symmetry of the wavefunction for the system, there is for the first excited state a greater probability of finding two helium atoms within 2.5 Å of each other, where the potential energy becomes very large. Thus the energy of the first excited state is raised, relative to that of the ground state, as a result of the interaction and an energy gap is thereby created.

This discussion is necessarily incomplete and somewhat vague. Because the interaction exists, there are other possible excited states in addition to the single-particle states discussed. But further analysis of the nature of these states would lead us far beyond the scope of this book, into areas which are still imperfectly understood.

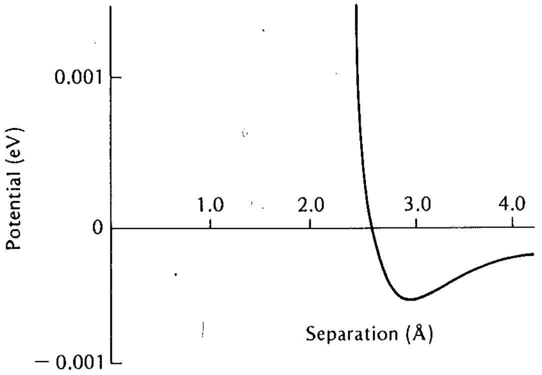

***Fig. 6.** Interaction potential between two helium atoms as a function of their separation.*

We conclude this discussion with a description of the most spectacular of the phenomena attributable to superfluidity—the fountain effect. To achieve this effect, a tube is plugged with emery powder and cotton at the bottom, with the top left open, and then it is immersed in liquid HeII, as in Fig. 7. Heat is then supplied to the helium inside the tube, so that some of the He atoms are raised from the ground state; superfluid is "destroyed" inside the tube. Destruction of superfluid is equivalent to raising the temperature. In order to maintain uniformity of temperature and of superfluid fraction throughout the volume of the fluid, superfluid rushes in toward the heater. (Remember the observation of the high thermal conductivity of liquid HeII.) If the plug were not present in the bottom of the tube, movement of superfluid toward the heater would be accompanied by movement of normal fluid away from the heater, both movements tending to transfer heat from the heater and maintain a uniform temperature throughout the fluid, without any net transfer of *mass* in either direction. But the plug prevents the normal fluid, with its nonzero viscosity, from flowing out the bottom of the tube, so there *is* a transfer of mass into the heated region, and the liquid level builds up in the tube. If the tube has a narrow top, a fountain of liquid helium eventually shoots out the top of the tube.

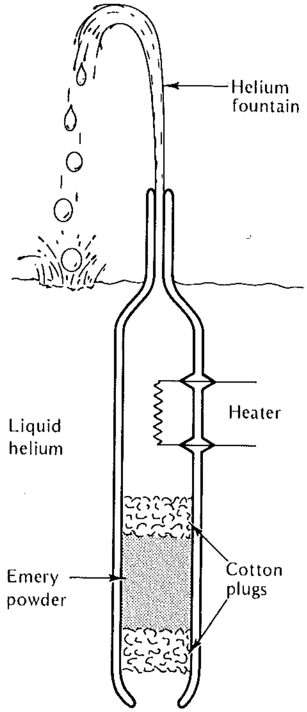

Fig. 7. Schematic illustration of the fountain effect. [From "Principles of Modern Physics" by R. B. Leighton. Copyright McGraw-Hill, New York, 1959. Used by permission of McGraw-Hill Book Company.]

11.4 APPLICATION OF FERMI–DIRAC STATISTICS: FREE ELECTRON THEORY OF METALS

The large electrical and thermal conductivity of metals indicates that many electrons—presumably the valence electrons—are not bound to individual atoms in a metal, but instead move freely throughout the volume of the metal. Apparently a valence electron can move in any direction with ease; as long as it remains inside the metal it is being pulled from all sides, and the forces in all directions tend to cancel each other. It is only when the electron reaches the surface of the metal that the positive charge of the atomic lattice pulls it back into the metal. In other words, the valence electrons in a metal behave like a gas confined within a box.

a. Electronic Specific Heat. The picture of metallic electrons as a gas is rather old, but it caused a puzzle for a long time, because this gas makes very little contribution to the specific heat of the metal. A classical gas, having an energy of $kT/2$ for each degree of freedom, would contribute a specific heat of $R/2$ per mole per degree of freedom; the electron gas, being like a monatomic gas, might therefore be expected to contribute $3R/2$ per mole. But the electronic contribution to the specific heat of a metal is known to be much smaller than this at room temperature, and it is not constant, but is proportional to the temperature.

Of course it should be no surprise to you that classical statistics fails to give the right answer, because an electron gas should obey Fermi–Dirac statistics. Using Fermi–Dirac statistics, we can easily show in a semiquantitative way that the electronic specific heat should vary linearly with temperature. To do this, we begin as we did with the other statistics, by investigating the parameter α. As before, α is a function of temperature which is determined by the condition that the total number of particles in the system is constant. In Fermi–Dirac statistics, it is convenient to write α as $-\mu/kT$, with μ another function of the temperature. This is merely a mathematical substitution; we shall see in a moment how it simplifies the analysis. The occupation index is now

$$\frac{n_s}{g_s} = \frac{1}{e^{(\varepsilon_s - \mu)/kT} + 1}$$

or, for a continuous distribution,

$$\frac{n(\varepsilon)}{g(\varepsilon)} = \frac{1}{e^{(\varepsilon - \mu)/kT} + 1} \qquad (31)$$

and μ is seen to be the energy at which the occupation index is $\frac{1}{2}$. We can evaluate μ (called the chemical potential) for a given system at any temperature by applying the condition $N = \int_0^\infty n(\varepsilon)\, d\varepsilon$ and solving the equation for μ.

Fortunately we can solve many problems without determining the precise temperature dependence of μ, simply by considering the situation near $T = 0$ K. At $T = 0$ we should expect all particles to be in their lowest possible energy states; as only one particle can occupy a given state, the occupation index should be 1 for the N lowest states and 0 for all higher states. Equation (31) bears out this expectation; at $T = 0$, the exponent becomes $+\infty$ for $\varepsilon > \mu$ and $-\infty$ for $\varepsilon < \mu$, so that

$$n(\varepsilon) = \begin{cases} 1 & (\varepsilon < \mu) \\ 0 & (\varepsilon > \mu) \end{cases}$$

and at $T = 0$ K, μ must be equal to the energy of the Nth state. This energy is called the Fermi energy ε_f, a constant for a given metal.

Figure 8 shows the effect of raising the temperature. The curve remains symmetrical about the point $\varepsilon = \mu$, where the occupation index $n(\varepsilon)/g(\varepsilon) = \frac{1}{2}$. As T increases, μ becomes smaller, but as long as kT is much smaller than ε_f, μ remains very close to ε_f. For example, if $kT = 0.01\varepsilon_f$ and $g(\varepsilon)$ is constant, it is not too hard to calculate that μ differs from ε_f by less than one part in 10^{45}. (See Problem 6 for another example.)

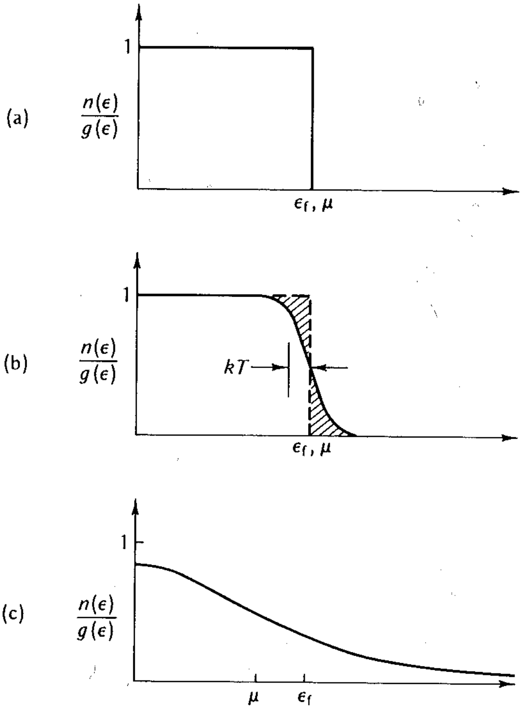

Fig. 8. The Fermi–Dirac occupation index; (a) at $T = 0°K$; (b) at small T, such that kT is of magnitude shown; (c) at very high T, where μ becomes noticeably smaller than ε_f.

If the temperature is raised so that $kT \gg \varepsilon_f$, then μ eventually becomes negative, so that only the "tail" of the Fermi–Dirac distribution appears in the positive energy levels, and the distribution merges into the Boltzmann distribution (as it should, for at such high temperatures, the density is so low that individual particles are distinguishable by their positions).

It happens that ε_f is several electron volts for metals, so that $kT \ll \varepsilon_f$ for any solid metal, and we have no further need to consider the temperature dependence of μ; we shall simply set μ equal to ε_f in the subsequent discussion.

The fact that $kT \ll \varepsilon_f$ immediately explains the linear temperature dependence of the electronic specific heat of metals. The occupation index is virtually 1 until $\varepsilon - \varepsilon_f$ is of order kT, and it falls from 1 to zero over an energy range of a few times kT. Thus only a small fraction of the electrons gain energy when the temperature is raised from 0°K to a temperature T; electrons are transferred from the shaded area below ε_f (Fig. 8) to the shaded area above ε_f. The *average* gain in energy per electron is the distance, on the energy axis, between the centroids of these areas; this distance is proportional to kT. The *number* of electrons which gain energy is proportional to the magnitude of the shaded area below ε_f; this quantity is also proportional to kT. The *total* energy gained by the electrons is proportional to the product of these two factors, that is, to $(kT)^2$. Since the specific heat is the temperature derivative of this gain in energy, the specific heat is proportional to T.

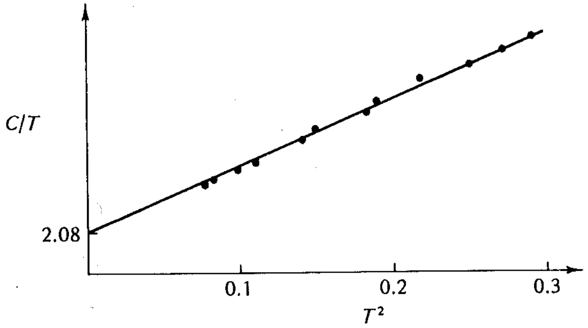

Fig. 9. Specific heat of potassium at low T, plotted as C/T versus T^2. The fit to a straight line, $C/T = 2.08 + 2.57T^2$, shows that C is given by the sum of a linear term and a cubic term [W. Lien and N. E. Phillips, Phys. Rev. 133, A1370 (1964)].

Thus the specific heat of a metal at low temperatures is the sum of an electronic contribution proportional to T, and the lattice contribution, proportional to T^3, which we have already discussed (Section 11.3b). Figure 9 shows how well this is verified by experiment.

b. Calculation of ε_f for a Metal. The general conclusions of the preceding discussion depended primarily on the fact that the electrons in a metal form a Fermi gas, and did not depend on the actual form of the potential energy for each electron. But if we wish to calculate the Fermi energy ε_f for a metal we must determine $g(\varepsilon)$; this requires that we calculate the energy levels of the system, which we can only do if we know the potential energy.

The simplest assumption we can make is that the potential energy is constant inside the metal. This is somewhat unrealistic, for we know that the electrons are strongly attracted to each atomic core in the metal, so that the potential energy should become very negative at these points. But the simple assumption of constant potential energy turns out to give very good results, and is a good starting point in any case.

If the potential energy is constant, we have already solved the problem of finding $g(\varepsilon)$, because it is almost the same as in the liquid helium problem (Section 11.3c). The metal is a box within which the electrons are confined, and we require that the wavefunction be zero at the walls of the box. The possible wavelengths are therefore the same as those already found for He atoms (Sec. 11.3). To find G, the number of states below a given energy ε, we therefore use Eq. (23), replacing the helium mass M by the electron mass m_e, and multiplying by two, because there are two possible spin states for each set of wavenumbers l_x, l_y, l_z. Thus

$$G = \frac{\pi V}{3}\left(\frac{8m_e\varepsilon}{h^2}\right)^{3/2} \tag{32}$$

When G = N, the total number of valence electrons, then $\varepsilon = \varepsilon_f$, and we may write

$$N = \frac{\pi V}{3}\left(\frac{8m_e\varepsilon_f}{h^2}\right)^{3/2}$$

and solve for the Fermi energy:

$$\varepsilon_f = \frac{h^2}{8m_e}\left(\frac{3N}{\pi V}\right)^{2/3} \tag{33}$$

The appropriate electron density N/V to use here is the number of *valence* electrons per unit volume; the other electrons are bound tightly to the atomic cores and are not part of the "gas."

EXAMPLE PROBLEM 2. Compute the Fermi energy of aluminum.

Solution. The atomic mass is 27, the density is 2.7 gm/cm³, and the valence is 3, so the valence electron density is

$$\frac{N}{V} = \frac{3 \times 6.022 \times 10^{23}}{26.98\ \text{g}} \times \frac{2.702\ \text{g}}{\text{cm}^3} = 1.809 \times 10^{23}\ \text{cm}^{-3}$$

and

$$\begin{aligned}\varepsilon_f &= \frac{h^2}{8m_e}\left(\frac{3}{\pi}\frac{N}{V}\right)^{2/3}\\ &= \frac{h^2c^2}{8m_ec^2}\left(\frac{0.1809\ \text{Å}^{-3}}{1.047}\right)^{2/3}\\ &= \frac{(1.240 \times 10^4)^2\ \text{eV}^2\text{-Å}^2}{8 \times 5.11 \times 10^5\ \text{eV}} \times (0.1727)^{2/3}\ \text{Å}^{-2}\\ \varepsilon_f &= 11.66\ \text{eV}\end{aligned}$$

Aluminum *boils* at $T = 2330$ K, where kT is still only 0.19 eV. The Fermi energy is, of course, lower for monovalent metals, but it still corresponds to a very high temperature; Fermi energies for several monovalent metals are listed in Table 4, with the equivalent temperatures (ε_f/k).

Table 4

Fermi Energies and Equivalent Temperatures for Some Metals

Metal	ε_f (eV)	$\frac{\varepsilon_f}{k}$ (K)
Li	4.7	5.5×10^4
Na	3.1	3.7
K	2.1	2.4
Rb	1.8	2.1
Cs	1.5	1.8
Cu	7.0	8.2
Ag	5.5	6.4
Au	5.5	6.4

In Chapter 13, we shall see how it is possible to use positrons as a "probe" to measure the momentum distribution of electrons in a metal and thereby obtain a fairly direct test of the value of ε_f. The final result for ε_f has been found to be in agreement with Eq. (33), with a logical number of valence electrons, for every metal which has been tested. Thus, in spite of the crudeness of the underlying assumption that the potential energy is constant, this model is a useful approximation which is a good starting point for a study of electronic structure of metals.

c. Work Function and Contact Potential Difference. The Fermi energy is the *kinetic* energy of the electrons in the highest occupied states. We can relate this to the work function of the metal by a diagram like Fig. 10, showing the potential well in which the electrons reside and the filled states up to energy ε_f above the bottom of the well. If the well depth is W, the work function, being the energy needed to remove an electron from the metal, is obviously

$$e\phi = W - \varepsilon_f.$$

A diagram similar to Fig. 10 is useful in explaining the contact difference of potential between two metals. When a wire is connected between two metals, electrons can flow from one metal to the other until the Fermi *levels*[15]—the energies at the top of the filled states—are equal. Figure 11

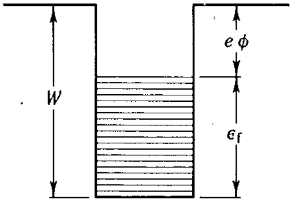

Fig. 10. *Relationship between well depth W, Fermi energy ε_f, and work function $e\phi$ for electrons in a metal. Horizontal lines indicate filled energy levels.*

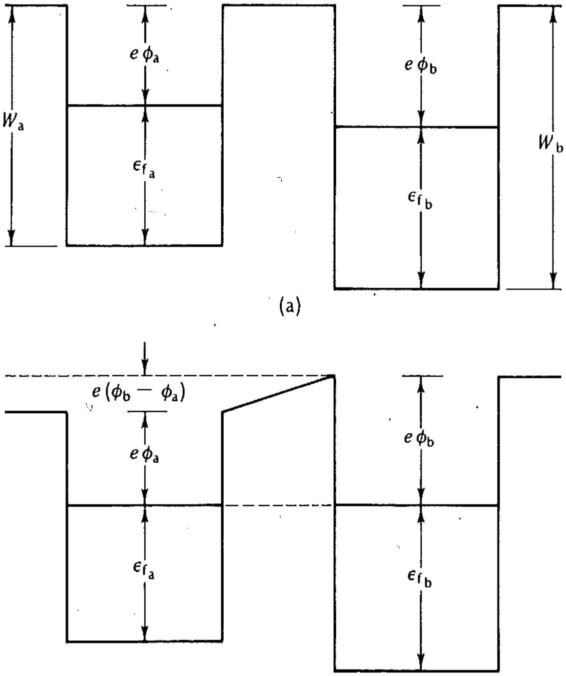

Fig. 11. *Effect on Fermi level when two metals are (a) separated (b) connected. Change in level results from change in the potential well, with negligible change in Fermi energy of either metal.*

---

[15] Note the distinction here between Fermi *level* and Fermi *energy*. The Fermi *energy* is the *kinetic* energy of each electron in the highest occupied states; the Fermi *level* is the *total* energy of each of these electrons, with respect to some outside reference energy.

shows what happens. Initially the metals are separated and uncharged, so that the potential is zero in the space between them. The metals have different well depths, different Fermi energies, and different work functions, so that there are electrons in metal A with higher *total* energy than that of any electron in metal B (at $T = 0$). When the metals are connected together, electrons flow from metal A to metal B in order to occupy the lower energy states available in B. Metal B becomes negatively charged, metal A becomes positively charged, and the potential energy is no longer zero in the region between the metals, so the potential well in B rises, relative to that in A, until the energy at the top of the filled levels in B equals that in A. The number of electrons transferred is minute[16] in comparison to the total number present, *so the values of ϕ and ε_f are unchanged in each metal.* The figure makes it clear that the difference in electrostatic potential between A and B becomes $\phi_b - \phi_a$.

Figure 10 helps us to understand the circumstances of a photoelectric experiment. If metals A and B are the two electrodes in a photocell, with $\phi_b > \phi_a$, the applied retarding potential V that is required to stop all the electrons is given by

$$eV = h\nu - e\phi_b$$

regardless of whether metal a or metal b is irradiated.[17]

It should also be noted that the contact potential difference cannot be measured with a conventional voltmeter. (Why not?) There is, however, a fairly direct way to measure it, by making a parallel-plate capacitor of the two metals, with a wire joining them, and measuring the current in the wire as the distance between the plates is varied. Changing the distance changes the capacitance, while the voltage across the capacitor remains constant (because it is the contact potential difference), so the charge on the capacitor must change, and a flow of current results. A known additional potential difference can then be inserted into the circuit, and varied until there is no longer a current flow when the capacitance is changed; at this point the voltage on the capacitor must be zero, so the additional potential difference must be just equal to the contact potential difference (and opposite in direction, of course.)

d. Emission of Electrons from a Metal. In the preceding discussion we ignored the electrons whose kinetic energy is greater than the Fermi energy. At any temperature above 0°K there are always some electrons, in the tail of the distribution, with enough kinetic energy to go over the hill and escape

---

[16] See Problem 7 for verification of this statement.

[17] For further discussion of this point, see J. Rudnick and D. S. Tannhauser, *Am. J. Phys.* **44**, 796 (1976).

from the metal completely. If one measures the energy distribution of the electrons after they have escaped, one finds that it is a Boltzmann distribution. This seems odd at first glance, because the electrons in the metal obey Fermi–Dirac statistics. What causes them to switch over to Boltzmann statistics?

If you have followed us to this point, you will realize that no switch is involved. The tail of the Fermi-Dirac distribution is a simple exponential, just like the Boltzmann distribution; the energies of the escaping electrons exceed the Fermi energy by many times kT, so that $e^{(\varepsilon-\varepsilon_{\mathrm{f}})/kT} \gg 1$, and the occupation index is

$$\frac{1}{e^{(\varepsilon-\varepsilon_{\mathrm{f}})/kT}+1} \rightarrow e^{-(\varepsilon-\varepsilon_{\mathrm{f}})/kT}$$

$$\rightarrow \text{constant} \times e^{-\varepsilon/kT}$$

which is the same as that for the Boltzmann distribution.

But in addition to measuring the energy of the escaping electrons, we can also measure the *number* of electrons which escape, as a function of temperature. The ability of Fermi–Dirac statistics to explain this temperature dependence, after classical calculations failed, was a great triumph for quantum theory. According to Boltzmann statistics, the current density of escaping electrons should be proportional to $T^{1/2}e^{-e\phi/kT}$, but experiments show that this current density is proportional to $T^2e^{-e\phi/kT}$. The derivation of the latter temperature dependence is straightforward, if Fermi–Dirac statistics are used. Details are left as a problem (Problem 8).

Another interesting situation arises when an electric field is applied at the surface of a metal. The effect of such a field is to change the potential energy for an electron near the surface. Therefore, let us abandon our approximation that the electrons are in a *square* potential well, and let us examine more closely the shape of the potential near the metallic surface. Standard electrostatic theory tells us that an electron just *outside* a metal is *not* in a region of constant (zero) potential, but rather that the electron is attracted to the metal by a force which is equal to that of an "image charge" located inside the metal at the point where the mirror image of the electron would be. If the electron is at a distance x from the surface, the distance between the electron and its image is $2x$, so the electron feels a force of $e^2/4\pi\varepsilon_0(4x)^2$ (in mks units), and the corresponding potential energy is

$$V(x) = -\frac{e^2}{16\pi\varepsilon_0 x} \tag{34}$$

This potential would go to $-\infty$ as x goes to zero, but we know that the well has a finite depth W. Therefore we cut off the potential (34) at the point where it is equal to the well depth, that is, at $x = x_c$, where

$$-W = -\frac{e^2}{16\pi\varepsilon_0 x_c}$$

so that $V(x) = -W$ for $x \leqslant x_c$, and the resulting potential is as shown in Fig. 12a. This potential is still a bit unrealistic, because the metal is not really a continuous medium, but we can use this model to make rough calculations of the effect of applying an external electric field.

When a uniform external field is applied to the metal surface, in a direction toward the surface, so that it tends to pull electrons out of the metal, the potential becomes that of Fig. 12b. The potential energy never reaches

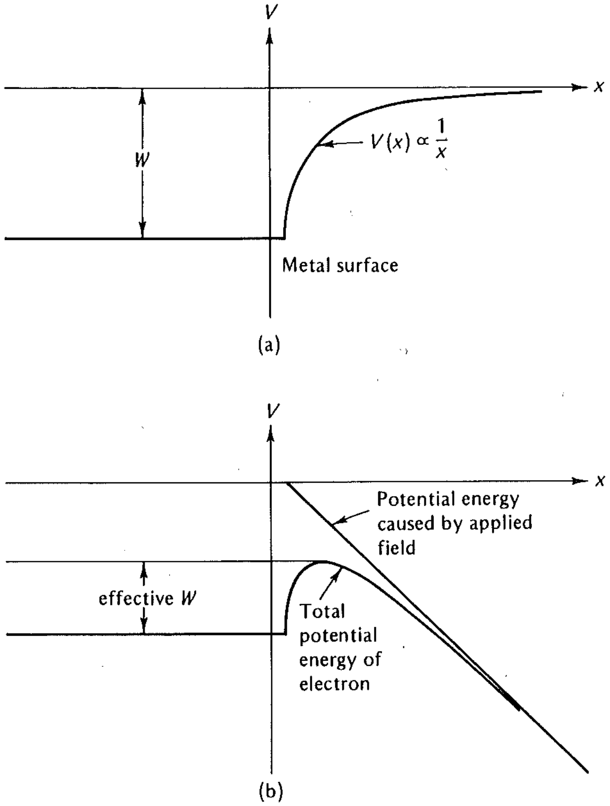

Fig. 12. *(a) Potential energy of an electron near the surface of a metal, including the effect of the image charge, but ignoring the effect of individual atoms. (b) Potential energy after a uniform electric field is applied to the surface; the linear potential of the applied field is added to the potential shown in (a).*

zero, but instead goes to a maximum and then decreases. Thus the *effective* well depth—the difference in energy between the bottom of the well and the maximum value of $V(x)$—is smaller than W, and the work function is correspondingly smaller than in the field-free situation.

At any given temperature the application of an electric field increases the current of electron emission, not only because of the decrease in the effective work function, but also because it is then possible for some electrons to *tunnel through* the barrier formed by the potential. This tunneling is possible because the barrier is no longer infinitely thick; as you can see from Fig. 12b, there is a region outside the metal where the potential energy of an electron is lower than it is inside. Calculation of the electric-field dependence of the electron emission current, taking account of these two effects, gives results which are in reasonable agreement with experiment.

e. Pauli Paramagnetism. Paramagnetic susceptibility is another property of metals which is not correctly accounted for by classical statistics. Consider a gas of N particles of spin $\frac{1}{2}$ and magnetic moment $\boldsymbol{\mu}$. If a magnetic field **B** is applied, a particle whose moment is parallel to **B** has a magnetic energy of $-\mu B$,[18] and a particle whose moment is antiparallel to **B** has a magnetic energy of $+\mu B$. According to classical statistics, the number of particles in each group is determined simply by the Boltzmann factor $e^{-\varepsilon/kT}$, so that if $\mu B \ll kT$, the number of particles whose moment is parallel to **B** is

$$\begin{aligned} n_1 &\simeq \frac{N}{2}\, e^{+\mu B/kT} \\ &\simeq \frac{N}{2}\left\{1 + \frac{\mu B}{kT}\right\} \end{aligned} \tag{35}$$

and the number whose moment is antiparallel to B is

$$n_2 \simeq \frac{N}{2}\, e^{-\mu B/kT} \simeq \frac{N}{2}\left\{1 - \frac{\mu B}{kT}\right\} \tag{36}$$

We may use the numbers n_1 and n_2 to compute the paramagnetic susceptibility, as follows: The difference between n_1 and n_2, multiplied by μ, is the total magnetic moment of the N particles. The paramagnetic susceptibility χ is defined by the equation[19]

$$\mathbf{M} = \chi \mathbf{H} \tag{37}$$

where **M**, the magnetization, is the total magnetic moment divided by the volume. According to Eqs. (35) and (36), the magnitude of **M** must be

---

[18] Since the z component of μ has only one possible magnitude, we denote that magnitude simply by the symbol μ. For an electron, $\mu = e\hbar/2m$ (Section 7.3). As usual, we assume **B** to be in the z direction, with $B_z \equiv B$.

[19] In mks units, where $\mathbf{B} = (\mathbf{H} + \mathbf{M})/\epsilon_0 c^2$.

$$M = \frac{\mu(n_1 - n_2)}{V}$$

$$= \frac{\mu\left(\dfrac{N\mu B}{kT}\right)}{V} = \frac{N\mu^2 B}{kTV} \qquad (38)$$

where V is the volume of the gas. We are concerned here with a situation in which χ is quite small, in which case we may write $\mathbf{B} \approx \mathbf{H}/\varepsilon_0 c^2$, and Eq. (37) becomes

$$\mathbf{M} = \chi \mathbf{B}\, \varepsilon_0 c^2$$

Comparison with Eq. (38) then yields

$$\chi = \frac{N\mu^2}{kTV\varepsilon_0 c^2} \qquad (39)$$

Equation (39) says that χ is proportional to $1/T$, a result known as the *Curie law*, which is valid for many materials. But the measured paramagnetic susceptibilities of solid metals are *independent of T*, and are only about 1 percent as large as the values predicted by Eq. (39) for room temperature. Of course, a solid metal does not behave like a classical gas, but in this case the electrons in the metal do form a *Fermi* gas. We can quickly explain the temperature independence of χ by means of Fermi–Dirac statistics, as follows: Electrons with energy much less than ε_f cannot contribute to the total magnetic moment, because there are equal numbers of states with spin parallel to **B** and with spin antiparallel to **B**, and all of the states are filled. The only electrons which can contribute to the susceptibility are those in the region where some states are filled and some are empty. The number of such electrons is proportional to kT; the T dependence of this number cancels the $1/T$ dependence expected from Eq. (39), leaving a susceptibility which is independent of T.

Pauli explained this point in 1927, and he derived the correct formula to replace Eq. (39), using the following line of reasoning: Since we expect the result to be independent of T, let us compute it at $T = 0$. We consider the electron gas to be *two* gases, one with spin parallel to **B** and density $n_2(\varepsilon)$ per unit energy interval, the other with spin antiparallel to **B** and with density $n_1(\varepsilon)$ per unit energy interval. We may plot both densities on the same graph, using the upward axis for $n_1(\varepsilon)$ and the downward axis for $n_2(\varepsilon)$ (see Fig. 13). Because the occupation index is 1 for $\varepsilon < \varepsilon_f$, each curve is the same as the curve of $g(\varepsilon)$, which is proportional to $\varepsilon^{1/2}$. Figure 13a represents the situation when $B = 0$; there are equal numbers of electrons in each gas. Figure 13b shows what happens when a field **B** is applied. The energy of each electron in gas 1 is reduced by an amount μB, and the energy of each electron in gas 2 is increased by the same amount. Electrons in the shaded part of gas 2 now find that lower energy states of opposite spin are available to them, so they flip their spins and join gas 1, filling the shaded area there. The number of elec-

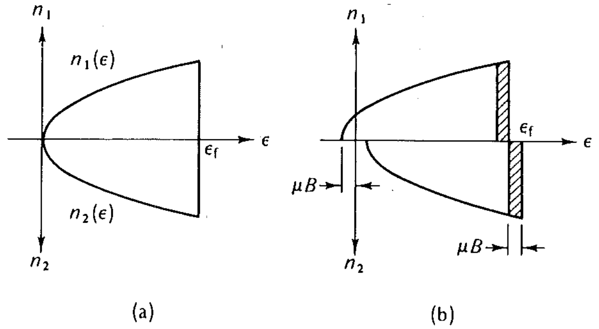

Fig. 13. *Shift in occupation density $n(\varepsilon)$ for an electron gas when a magnetic field is applied. See text for details.*

trons which flip is equal to the shaded area in either gas; if $\mu B \ll kT$, this number is

$$n_f = \tfrac{1}{2}\mu B\, n(\varepsilon_f) = \tfrac{1}{2}\mu B\, g(\varepsilon_f)$$

the factor of $\frac{1}{2}$ entering because the density of electrons in each gas is only one-half the total occupation density $n(\varepsilon)$.

Each electron's magnetic moment changes by 2μ when it flips, so the net resulting magnetic moment is $2\mu n_f = \mu^2 B\, g(\varepsilon_f)$, and the resulting magnetization is of magnitude

$$M = \frac{\mu^2 B g(\varepsilon_f)}{V}$$

The susceptibility is thus

$$\chi = \frac{\mu^2 g(\varepsilon_f)}{V\varepsilon_0 c^2}$$

The density of states is

$$g(\varepsilon_f) = \left.\frac{dG}{d\varepsilon}\right|_{\varepsilon=\varepsilon_f}$$

with G as given in Section 11.4(b), so that

$$\frac{g(\varepsilon_f)}{V} = \frac{\pi}{2}\left(\frac{8m_e}{h^2}\right)^{3/2} \varepsilon_f^{1/2}$$

and

$$\chi = \frac{\pi\mu^2}{2\varepsilon_0 c^2}\left(\frac{8m_e}{h^2}\right)^{3/2} \varepsilon_f^{1/2}$$

The susceptibility is often written in terms of the electron density N/V. Since, from Eq. (33),

$$\frac{N}{V} = \frac{\pi}{3}\left(\frac{8m_e\varepsilon_f}{h^2}\right)^{3/2}$$

we may write

$$\chi = \frac{\mu_0 \pi \mu^2}{2}\left(\frac{8m_e}{h^2}\right)^{3/2} \varepsilon_f^{1/2}\left(\frac{\dfrac{N}{V}}{\left(\dfrac{\pi}{3}\right)\left(\dfrac{8m_e\varepsilon_f}{h^2}\right)^{3/2}}\right)$$

or

$$\chi = \frac{3\mu_0 \mu^2 N}{2V\varepsilon_f} \qquad (\mu_0 \equiv 1/\varepsilon_0 c^2) \tag{40}$$

While it is a vast improvement on Eq. (39), Eq. (40) yields values of χ which differ from experimental results by almost a factor of 2 even for alkali metals, for which the agreement is best. The error must lie in the calculation of $g(\varepsilon)$; we cannot expect to obtain $g(\varepsilon)$ accurately from an assumption that the potential energy of an electron is constant inside a metal. The actual potential energy is periodic, with a minimum at each lattice point, where a positively charged atomic core resides. In an alkali metal, the Fermi energy is so small that the wavelengths of the free electrons are all much greater than the lattice spacing; thus the periodic variation of the potential has less effect than in other metals and we should expect Eq. (40) to work best in this case.

f. Electrical Conductivity. One might expect the exclusion principle to have a great effect on the electrical conductivity of a metal, because acceleration of an electron requires the transition of an electron to another state, and many states are unavailable because they are already occupied. However, the effect is not drastic, as we show with the aid of Fig. 14.

When the field is applied, electrons at the "front" of the distribution can easily be accelerated, because they can transfer to empty states of larger p_x. This means that *all* the electrons can be accelerated, because the electrons with smaller p_x simply move into the states vacated by the electrons with larger p_x. The entire distribution can thus be rigidly accelerated by the field, without running afoul of the exclusion principle.

Eventually the whole distribution is displaced by an amount mv, where v is the "terminal velocity," determined by energy losses to the lattice as electrons are scattered by atoms in the lattice. These energy losses result from the recoil of the atomic cores; since these cores are much more massive than an electron, an electron loses only a small fraction of its energy in each scattering event. If the electric field is increased, the electrons will go to a larger average velocity, until the larger energy losses from scattering just compensate for the larger energy put in by the electric field.

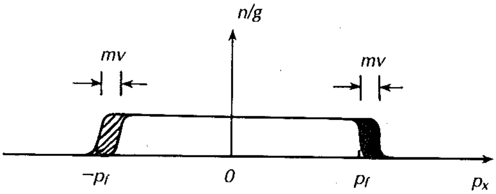

Fig. 14. Occupation index n/g as a function of x component of momentum. *In zero electric field, distribution extends from $-p_f$ to $+p_f$. When field is applied in x direction, electrons are accelerated into the blackened region, and accelerated out of the shaded region, shifting the whole distribution by the amount mv.*

When the field is turned off, these energy losses to the lattice cause the average velocity to return to zero, and the current ceases. Although the energy change in each scattering event is very small (relative to the Fermi energy), the momentum change is usually large (relative to the Fermi momentum p_f), because an electron can only be scattered if its final state, after scattering, was previously unoccupied. Thus electrons are scattered from the blackened region (Fig. 14) into the shaded region. For example, an electron with momentum $p_x = p_f + mv$ can lose energy by scattering to a state with $|p_x'| < p_x$. But the lowest-energy unoccupied state such that $|p_x'| < p_x$ is at approximately (from Fig. 14)

$$p_x' = -p_f + mv$$

Thus the *maximum* energy *loss* is

$$\Delta E_{max} = \frac{p_x^2 - p_x'^2}{2m}$$

$$= \frac{(p_f + mv)^2 - (p_f - mv)^2}{2m}$$

$$= 2p_f v$$

Since $E_f = p_f^2/2m$, we can write this result as

$$\frac{\Delta E_{max}}{E_f} = \frac{2p_f v}{p_f^2/2m} = \frac{2mv}{p_f} \tag{41}$$

In most cases, $mv << p_f$, so $\Delta E_{\text{max}} << E_f$. But since the atomic recoil energy is small even for classical collisions, this restriction on the size of ΔE, which results from the exclusion principle, has no serious effect on the relaxation of a current when the electric field is turned off, except in special cases (e.g., superconductivity).

In a real metal, the scattering process results from imperfections in the lattice; we shall see in Chapter 12 that electrons can move through a perfect, static lattice with no energy loss from scattering. But defects, impurities, or vibrations of the lattice lead to scattering processes in which energy is lost. Thus all of these tend to increase the electrical resistivity. In Chapter 12 we shall discuss the connection between these ideas and the explanation of the conduction process in semiconductors and superconductors.

SUMMARY

In this chapter, armed with nothing but elementary quantum statistics, we have been able to explain an enormous variety of phenomena which had puzzled investigators for a long time. It is remarkable that we could do this even though we used only the simplest possible models for the structure of the systems which we have studied. By now we should be convinced of the validity of the quantum statistical formulas, and we are in a position to study more realistic models. In the next chapter, we shall see how various models of solids can be analyzed by straightforward application of the elements of quantum mechanics which we have already developed in Chapters 4-11.

PROBLEMS

1. Use the general formulas (3), (4), and (5) to compute the number of arrangements for each distribution in the four-particle example of Section 11.1, and verify that your results agree with Table 1.

2. Derive the Einstein expression for the specific heat in the limits of high $T\,(T \gg h\nu/k)$ and low $T\,(T \ll h\nu/k)$. Show that your low-T formula gives a specific heat of $0.27R$ at $T = h\nu/6k$.

3. Find the value of the specific heat at $T = \Theta_D/2$, according to the Debye theory. To do this, note that Eq. (20) may be written

$$E = 9R\frac{T^4}{\Theta_D^3} f\left(\frac{\Theta_D}{T}\right)$$