

PART
B

STATISTICAL
MECHANICS

CHAPTER

6

CLASSICAL STATISTICAL
MECHANICS6.1 THE POSTULATE OF CLASSICAL
STATISTICAL MECHANICS

Statistical mechanics is concerned with the properties of matter in equilibrium in the empirical sense used in thermodynamics.

The aim of statistical mechanics is to derive all the equilibrium properties of a macroscopic molecular system from the laws of molecular dynamics. Thus it aims to derive not only the general laws of thermodynamics but also the specific thermodynamic functions of a given system. Statistical mechanics, however, does not describe how a system approaches equilibrium, nor does it determine whether a system can ever be found to be in equilibrium. It merely states what the equilibrium situation is for a given system.

We recall that in the kinetic theory of gases the process of the approach to equilibrium is rather complicated, but the equilibrium situation, the Maxwell-Boltzmann distribution, is simple. Furthermore, the Maxwell-Boltzmann distribution can be derived in a simple way, independent of the details of molecular interactions. We might suspect that a slight generalization of the method used—the method of the most probable distribution—would enable us to discuss the equilibrium situation of not only a dilute gas but also any macroscopic system. This indeed is true. The generalization is classical statistical mechanics.

We consider a classical system composed of a large number N of molecules occupying a large volume V . Typical magnitudes of N and V are

$$\begin{aligned} N &\approx 10^{23} && \text{molecules} \\ V &\approx 10^{23} && \text{molecular volumes} \end{aligned}$$

Since these are enormous numbers, it is convenient to consider the limiting case

$$\begin{aligned} \frac{N}{V} &\rightarrow \frac{m}{v} \\ \frac{V}{N} &= v \end{aligned} \quad (6.1)$$

where the specific volume v is a given finite number.

The system will be regarded as isolated in the sense that the energy is a constant of the motion. This is clearly an idealization, for we never deal with truly isolated systems in the laboratory. The very fact that measurements can be performed on the system necessitates some interaction between the system and the external world. If the interactions with the external world, however, are sufficiently weak, so that the energy of the system remains approximately constant, we shall consider the system isolated. The walls of the container containing the system (if present) will be idealized as perfectly reflecting walls.

A state of the system is completely and uniquely defined by $3N$ canonical coordinates q_1, q_2, \dots, q_{3N} and $3N$ canonical momenta p_1, p_2, \dots, p_{3N} . These $6N$ variables are denoted collectively by the abbreviation (p, q) . The dynamics of the system is determined by the Hamiltonian $\mathcal{H}(p, q)$, from which we may obtain the canonical equations of motion

$$\begin{aligned} \frac{\partial \mathcal{H}(p, q)}{\partial p_i} &= \dot{q}_i \\ \frac{\partial \mathcal{H}(p, q)}{\partial q_i} &= -\dot{p}_i \end{aligned} \quad (6.2)$$

It is convenient to introduce, as we did in Chapter 3, the $6N$ -dimensional Γ space, or phase space, of the system, in which each point represents a state of the system, and vice versa. The locus of all points in Γ space satisfying the condition $\mathcal{H}(p, q) = E$ defines a surface called the energy surface of energy E . As the state of the system evolves in time according to (6.2) the representative point traces out a path in Γ space. This path always stays on the same energy surface because by definition energy is conserved.

For a macroscopic system, we have no means, nor desire, to ascertain the state at every instant. We are interested only in a few macroscopic properties of the system. Specifically, we only require that the system has N particles, a volume V , and an energy lying between the values E and $E + \Delta$. An infinite number of states satisfy these conditions. Therefore we think not of a single system, but of an infinite number of mental copies of the same system, existing in all possible states satisfying the given conditions. Any one of these system can be the system we are dealing with. The mental picture of such a collection of systems is the Gibbsian ensemble we introduced in Chapter 3. It is represented by a distribution of points in Γ space characterized by a density function $\rho(p, q, t)$, defined in such a way that

$$\rho(p, q, t) d^{3N}p d^{3N}q = \text{no. of representative points contained in the volume element } d^{3N}p d^{3N}q \text{ located at } (p, q) \text{ in } \Gamma \text{ space at the instant } t \quad (6.3)$$

We recall Liouville's theorem:

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^{3N} \left(\frac{\partial \rho}{\partial q_i} \frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial \mathcal{H}}{\partial q_i} \frac{\partial \rho}{\partial p_i} \right) = 0 \quad (6.4)$$

In geometrical language it states that the distribution of points in Γ space moves like an incompressible fluid. Since we are interested in the equilibrium situation, we restrict our considerations to ensembles whose density function does not depend explicitly on the time and depends on (p, q) only through the Hamiltonian. That is,

$$\rho(p, q) = \rho'(\mathcal{H}(p, q)) \quad (6.5)$$

where $\rho'(\mathcal{H})$ is a given function of \mathcal{H} . It follows immediately that the second term on the left side of (6.4) is identically zero. Therefore

$$\frac{\partial}{\partial t} \rho(p, q) = 0 \quad (6.6)$$

Hence the ensemble described by $\rho(p, q)$ is the same for all times.

Classical statistical mechanics is founded on the following postulate.

Postulate of Equal a Priori Probability When a macroscopic system is in thermodynamic equilibrium, its state is equally likely to be any state satisfying the macroscopic conditions of the system.

This postulate implies that in thermodynamic equilibrium the system under consideration is a member of an ensemble, called the microcanonical ensemble, with the density function

$$\rho(p, q) = \begin{cases} \text{Const.} & \text{if } E < \mathcal{H}(p, q) < E + \Delta \\ 0 & \text{otherwise} \end{cases} \quad (6.7)$$

It is understood that all members of the ensemble have the same number of particles N and the same volume V .

Suppose $f(p, q)$ is a measurable property of the system, such as energy or momentum. When the system is in equilibrium, the observed value of $f(p, q)$ must be the result obtained by averaging $f(p, q)$ over the microcanonical ensemble in some manner. If the postulate of equal a priori probability is to be useful, all manners of averaging must yield essentially the same answer.

Two kinds of average values are commonly introduced: the most probable value and the ensemble average. The most probable value of $f(p, q)$ is the value of $f(p, q)$ that is possessed by the largest number of systems in the ensemble. The ensemble average of $f(p, q)$ is defined by

$$\langle f \rangle \equiv \frac{\int d^{3N}p d^{3N}q f(p, q) \rho(p, q)}{\int d^{3N}p d^{3N}q \rho(p, q)} \quad (6.8)$$

The ensemble average and the most probable value are nearly equal if the mean square fluctuation is small, i.e., if

$$\frac{\langle f^2 \rangle - \langle f \rangle^2}{\langle f \rangle^2} \ll 1 \quad (6.9)$$

If this condition is not satisfied, there is no unique way to determine how the observed value of f may be calculated. When it is not, we should question the validity of statistical mechanics. In all physical cases we shall find that mean square fluctuations are of the order of $1/N$. Thus in the limit as $N \rightarrow \infty$ the ensemble average and the most probable value became identical.

Strictly speaking, systems in nature do not obey classical mechanics. They obey quantum mechanics, which contains classical mechanics as a special limiting case. Logically we should start with quantum statistical mechanics and then arrive at classical statistical mechanics as a special case. We do this later. It is only for pedagogical reasons that we begin with classical statistical mechanics.

From a purely logical point of view there is no room for an independent postulate of classical statistical mechanics. It would not be logically satisfactory even if we could show that the postulate introduced here follows from the equations of motion (6.2), for, since the world is quantum mechanical, the foundation of statistical mechanics lies not in classical mechanics but in quantum mechanics. At present we take this postulate to be a working hypothesis whose justification lies in the agreement between results derived from it and experimental facts.

6.2 MICROCANONICAL ENSEMBLE

In the microcanonical ensemble every system has N molecules, a volume V , and an energy between E and $E + \Delta$. It is clear that the average total momentum of the system is zero. We show that it is possible to define quantities that correspond to thermodynamic quantities.

The fundamental quantity that furnishes the connection between the microcanonical ensemble and thermodynamics is the entropy. It is the main task of this section to define the entropy and to show that it possesses all the properties attributed to it in thermodynamics.

Let $\Gamma(E)$ denote the volume in Γ space occupied by the microcanonical ensemble:

$$\Gamma(E) \equiv \int_{E < \mathcal{H}(p, q) < E + \Delta} d^{3N}p d^{3N}q \quad (6.10)$$

The dependence of $\Gamma(E)$ on N , V , and A is understood. Let $\Sigma(E)$ denote the volume in Γ space enclosed by the energy surface of energy E :

$$\Sigma(E) = \int_{\mathcal{H}(p, q) < E} d^{3N}p d^{3N}q \quad (6.11)$$

Then

$$\Gamma(E) = \Sigma(E + \Delta) - \Sigma(E) \quad (6.12)$$

If Δ is so chosen that $\Delta \ll E$, then

$$\Gamma(E) = \omega(E) \Delta \quad (6.13)$$

where $\omega(E)$ is called the density of states of the system at the energy E and is defined by

$$\omega(E) = \frac{\partial \Sigma(E)}{\partial E} \quad (6.14)$$

The entropy is defined by

$$S(E, V) \equiv k \log \Gamma(E) \quad (6.15)$$

where k is a universal constant eventually shown to be Boltzmann's constant. To justify this definition we show that (6.15) possesses all the properties of the entropy function in thermodynamics, namely,

- (a) S is an extensive quantity: If a system is composed of two subsystems whose entropies are, respectively, S_1 and S_2 , the entropy of the total system is $S_1 + S_2$, when the subsystems are sufficiently large.
- (b) S satisfies the properties of the entropy as required by the second law of the thermodynamics.

To show the extensive property, let the system be divided into two subsystems which have N_1 and N_2 particles and the volumes V_1 and V_2 , respectively.* The energy of molecular interaction between the two subsystems is negligible compared to the total energy of each subsystem, if the intermolecular potential has a finite range, and if the surface-to-volume ratio of each subsystem is negligibly small. The total Hamiltonian of the composite system accordingly may be taken to be the sum of the Hamiltonians of the two subsystems:

$$\mathcal{H}(p, q) = \mathcal{H}_1(p_1, q_1) + \mathcal{H}_2(p_2, q_2) \quad (6.16)$$

where (p_1, q_1) and (p_2, q_2) denote, respectively, the coordinates and momenta of the particles contained in the two subsystems.

Let us first imagine that the two subsystems are isolated from each other and consider the microcanonical ensemble for each taken alone. Let the energy of the first subsystem lie between E_1 and $E_1 + A$ and the energy of the second subsystem lie between E_2 and $E_2 + A$. The entropies of the subsystems are, respectively,

$$S_1(E_1, V_1) = k \log \Gamma_1(E_1)$$

$$S_2(E_2, V_2) = k \log \Gamma_2(E_2)$$

where $\Gamma_1(E_1)$ and $\Gamma_2(E_2)$ are the volumes occupied by the two ensembles in their respective Γ spaces. They are schematically represented in Fig. 6.1 by the volumes of the shaded regions, which lie between successive energy surfaces that differ in energy by A .

Now consider the microcanonical ensemble of the composite system made up of the two subsystems, and let the total energy lie between E and $E + 2A$.

*For simplicity we assume that the same N_1, N_2 particles are always confined respectively to the volumes V_1, V_2 . The proof is therefore invalid for a gas, for which S has to be modified (See Section 6.6).

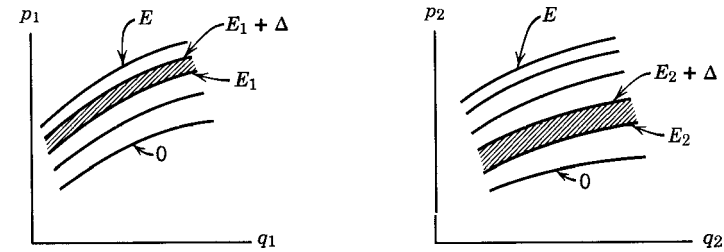


Fig. 6.1 The microcanonical ensemble of the two subsystems.

We choose A such that $A \ll E$. This ensemble contains all copies of the composite system for which

- (a) the N_1 particles whose momenta and coordinates are (p_1, q_1) are contained in the volume V_1 ,
- (b) the N_2 particles whose momenta and coordinates are (p_2, q_2) are contained in the volume V_2 ,
- (c) the energies E_1, E_2 of the subsystems have values satisfying the condition

$$E < (E_1 + E_2) < E + 2\Delta \quad (6.17)$$

Obviously, the volume of the region of Γ space that corresponds to conditions (a) and (b) with a total energy lying between $E_1 + E_2$ and $E_1 + E_2 + 2A$ is

$$\Gamma_1(E_1)\Gamma_2(E_2)$$

To obtain the total volume of the ensemble specified by (a), (b), and (c), we only have to take the sum of $\Gamma_1(E_1)\Gamma_2(E_2)$ over values of E_1 and E_2 consistent with (c). Since E_1 and E_2 are possible values of the Hamiltonians $\mathcal{H}_1(p_1, q_1)$ and $\mathcal{H}_2(p_2, q_2)$, their spectra of values must be bounded from below; otherwise the subsystems would not be stable. For simplicity we take the lower bounds for both spectra to be 0. If we divide each of the energy spectra E_1 and E_2 into intervals of size A , then between 0 and E there are E/A intervals in each spectrum. Thus, since $A \ll E$, we can write

$$\Gamma(E) = \sum_{i=1}^{E/\Delta} \Gamma_1(E_i)\Gamma_2(E - E_i) \quad (6.18)$$

where E_i is the energy lying in the center of each energy interval.

The entropy of the composite system of N particles and of volume V , with

$$N = N_1 + N_2$$

$$V = V_1 + V_2$$

is given by

$$S(E, V) = k \log \sum_{i=1}^{E/\Delta} \Gamma_1(E_i)\Gamma_2(E - E_i) \quad (6.19)$$

It will now be shown that when $N_1 \rightarrow \infty$ and $N_2 \rightarrow \infty$ a single term in the sum of (6.18) dominates the sum. The sum in (6.18) is a sum of E/A positive terms. Let the largest term in the sum be $\Gamma_1(\bar{E}_1)\Gamma_2(\bar{E}_2)$, where

$$\bar{E}_1 + \bar{E}_2 = E \tag{6.20}$$

Then it is obvious that

$$\Gamma_1(\bar{E}_1)\Gamma_2(\bar{E}_2) \leq \Gamma(E) \leq \frac{E}{\Delta} \Gamma_1(\bar{E}_1)\Gamma_2(\bar{E}_2)$$

$$k \log [\Gamma_1(\bar{E}_1)\Gamma_2(\bar{E}_2)] \approx S(E, V) \leq k \log [\Gamma_1(\bar{E}_1)\Gamma_2(\bar{E}_2)] + k \log \frac{E}{A} \tag{6.21}$$

If the subsystems are molecular systems with N_1 and N_2 particles, respectively, we expect that as $N_1 \rightarrow \infty$ and $N_2 \rightarrow \infty$,

$$\begin{aligned} \log \Gamma_1 &\propto N_1 \\ \log \Gamma_2 &\propto N_2 \\ E &\propto N_1 + N_2 \end{aligned} \tag{6.22}$$

Thus the term $\log(E/\Delta)$ in (6.21) may be neglected because A is a constant independent of N . Therefore

$$S(E, V) = S_1(\bar{E}_1, V_1) + S_2(\bar{E}_2, V_2) + O(\log N) \tag{6.23}$$

which proves the extensive property of the entropy. ■

We have actually proved more than the extensive property of the entropy because (6.23) also implies that the energies of subsystems have the definite values \bar{E}_1 and \bar{E}_2 , respectively. They are the values of E_1 and E_2 that maximize the function $\Gamma_1(E_1)\Gamma_2(E_2)$ under the restriction $E_1 + E_2 = E$. That is,

$$\delta [\Gamma_1(E_1)\Gamma_2(E_2)] = 0, \quad \delta E_1 + \delta E_2 = 0$$

This leads to the condition

$$\begin{aligned} \left[\frac{\partial}{\partial E_1} \log \Gamma_1(E_1) \right]_{E_1=\bar{E}_1} &= \left[\frac{\partial}{\partial E_2} \log \Gamma_2(E_2) \right]_{E_2=\bar{E}_2} \\ \left[\frac{\partial S_1(E_1)}{\partial E_1} \right]_{E_1=\bar{E}_1} &= \left[\frac{\partial S_2(E_2)}{\partial E_2} \right]_{E_2=\bar{E}_2} \end{aligned} \tag{6.24}$$

We define the temperature of any system by

$$\frac{\partial S(E, V)}{\partial E} \equiv \frac{1}{T} \tag{6.25}$$

Then \bar{E}_1 and \bar{E}_2 are such that the two subsystems have the same temperature:

$$T_1 = T_2 \tag{6.26}$$

The temperature defined by (6.25) is precisely the absolute temperature in thermodynamics. Not only is it a parameter associated with the condition for equilibrium, it is also related to the entropy by (6.25), which is one of the Maxwell relations in thermodynamics. Choosing the standard temperature interval to be the conventional Centigrade degree defines the constant k in (6.15) to be Boltzmann's constant. Thus the proof of the extensive property of the entropy also reveals the meaning of the temperature for an isolated system: *The temperature of an isolated system is the parameter governing the equilibrium between one part of the system and another.*

Although the condition (6.17) allows a range of values of (E_1, E_2) to occur among members of the microcanonical ensemble, the result (6.21) shows that as the number of particles becomes very large almost all members of the ensemble have the values (\bar{E}_1, \bar{E}_2) . This fact is fundamental to the success of statistical mechanics as a theory of matter.

A calculation similar to that leading to (6.23) shows that the following definitions of S are equivalent, up to additive constant terms of order $\log N$ or smaller:

$$S = k \log \Gamma(E) \tag{6.27}$$

$$S = k \log \omega(E) \tag{6.28}$$

$$S = k \log \sum(E) \tag{6.29}$$

In fact, if these definitions were not equivalent, the validity of statistical mechanics would be in doubt.

To show that S possesses the properties of the entropy as required by the second law of thermodynamics, let us first state the form of the second law that is most convenient for the present purpose. The entropy in thermodynamics, just as S here, is defined only for equilibrium situations. The second law states that if an isolated system undergoes a change of thermodynamic state such that the initial and final states are equilibrium states, the entropy of the final state is not smaller than that of the initial state. For the system we are considering, the only independent macroscopic parameters are N , V , and E . By definition N and E cannot change, for the system is isolated. Thus only V can change. Now V cannot decrease without compressing the system thereby disturbing its isolation. Hence V can only increase. (An example is the free expansion of a gas when one of the containing walls is suddenly removed.) For our purpose the second law states that the entropy is a nondecreasing function of V .

Let us use the definition (6.29):

$$S(E, V) = k \log \sum(E)$$

It is obvious that $\sum(E)$ is a nondecreasing function of V , for if $V_1 > V_2$, then

the integral (6.11) for $V = V_1$ extends over a domain of integration that includes that for $V = V_2$. This shows that $S(E, V)$ is a nondecreasing function of V .

We conclude that the function $S(E, V)$, as defined by any one of the formulas (6.27)–(6.29), is the entropy of a system of volume V and internal energy E . This conclusion furnishes the connection between the microcanonical ensemble and thermodynamics.

6.3 DERIVATION OF THERMODYNAMICS

We have defined the entropy of a system and have shown that the second law of thermodynamics holds. The complete thermodynamics of a system can now be obtained.

First we discuss the analog of quasistatic thermodynamic transformations. A quasistatic thermodynamic transformation corresponds to a slow variation of E and V , induced by coupling the system to external agents. During such a transformation the ensemble is represented by a collection of representative points uniformly distributed over a slowly changing region in Γ space. The change is so slow that at every instant we have a microcanonical ensemble. Accordingly, the change in the entropy in an infinitesimal transformation is given by

$$dS(E, V) = \left(\frac{\partial S}{\partial E} \right)_V dE + \left(\frac{\partial S}{\partial V} \right)_E dV \quad (6.30)$$

The coefficient of dE has been defined earlier as the inverse absolute temperature T^{-1} . We now define the pressure of the system to be

$$P \equiv T \left(\frac{\partial S}{\partial V} \right)_E \quad (6.31)$$

Hence

$$dS = \frac{1}{T} (dE + P dV) \quad (6.32)$$

$$dE = T dS - P dV \quad (6.33)$$

This is the first law of thermodynamics.

Thus we have succeeded not only in deriving the first and second laws of thermodynamics, but also in finding means to calculate all thermodynamic functions in terms of molecular interactions. The third law of thermodynamics cannot be obtained in classical statistical mechanics, because it is quantum mechanical.

We summarize by giving a practical recipe for finding all the thermodynamic functions of a system.

RECIPE

Consider an isolated system that occupied volume V and has an energy E within a small uncertainty $\Delta \ll E$. The Hamiltonian is presumed known. To find all thermodynamic functions of the system, proceed as follows:

- (a) Calculate the density of states $\omega(E)$ of the system from the **Hamiltonian**.
- (b) Find the entropy up to an arbitrary additive constant by the formula

$$S(E, V) = k \log \omega(E)$$

where k is **Boltzmann's** constant. Alternatively we can use the formula (6.27) or (6.29).

- (c) Solve for E in terms of S and V . The resulting function is the thermodynamic internal energy of the system

$$U(S, V) \equiv E(S, V)$$

- (d) Find other thermodynamic functions from the following formulas:

$$T = \left(\frac{\partial U}{\partial S} \right)_V \quad (\text{absolute temperature})$$

$$P = - \left(\frac{\partial U}{\partial V} \right)_S \quad (\text{pressure})^*$$

$$A = U - TS \quad (\text{Helmholtz free energy})$$

$$G = U + PV - TS \quad (\text{Gibbs potential})$$

$$C_V = \left(\frac{\partial U}{\partial T} \right)_V \quad (\text{heat capacity at constant volume})$$

- (e) To study any equilibrium behavior of the system, use thermodynamics.

6.4 EQUIPARTITION THEOREM

Let x_i be either p_i or q_i ($i = 1, \dots, 3N$). We calculate the ensemble average of $x_i (\partial \mathcal{H} / \partial x_j)$, where \mathcal{H} is the Hamiltonian. Using the abbreviation $dpdq \equiv d^{3N}p d^{3N}q$, we can write

$$\left\langle x_i \frac{\partial \mathcal{H}}{\partial x_j} \right\rangle = \frac{1}{\Gamma(E)} \int_{E < \mathcal{H} < E + \Delta} dp dq x_i \frac{\partial \mathcal{H}}{\partial x_j} = \frac{\Delta}{\Gamma(E)} \frac{\partial}{\partial E} \int_{\mathcal{H} < E} dp dq x_i \frac{\partial \mathcal{H}}{\partial x_j}$$

Noting that $\partial E / \partial x_j = 0$, we may calculate the last integral in the following

*This is equivalent to (6.31) by the chain relation.

manner:

$$\begin{aligned} \int_{\mathcal{H} < E} dp dq x_i \frac{\partial \mathcal{H}}{\partial x_j} &= \int_{\mathcal{H} < E} dp dq x_i \frac{\partial}{\partial x_j} (\mathcal{H} - E) \\ &= \int_{\mathcal{H} < E} dp dq \frac{d}{dx_j} [x_i (\mathcal{H} - E)] - \delta_{ij} \int_{\mathcal{H} < E} dp dq (2 - E) \end{aligned}$$

The first integral on the right side vanishes because it reduces to a surface integral over the boundary of the region defined by $\mathcal{H} < E$, and on this boundary $\mathcal{H} - E = 0$. Substituting the latest result into the previous equation, and noting that $\Gamma(E) = \omega(E)\Delta$, we obtain

$$\begin{aligned} \left\langle x_i \frac{\partial \mathcal{H}}{\partial x_j} \right\rangle &= \frac{\delta_{ij}}{\omega(E)} \frac{\partial}{\partial E} \int_{\mathcal{H} < E} dp dq (E - \mathcal{H}) \\ &= \frac{\delta_{ij}}{\omega(E)} \int_{\mathcal{H} < E} dp dq = \frac{\delta_{ij}}{\omega(E)} \Sigma(E) \\ &= \delta_{ij} \frac{\Sigma(E)}{\partial \Sigma(E) / \partial E} = \delta_{ij} \left[\frac{\partial}{\partial E} \log \Sigma(E) \right]^{-1} = \delta_{ij} \frac{k}{\partial S / \partial E} \end{aligned}$$

that is,

$$\left\langle x_i \frac{\partial \mathcal{H}}{\partial x_j} \right\rangle = \delta_{ij} kT \tag{6.34}$$

This is the generalized equipartition theorem.

For the special case $i = j$, $x_i = p_i$, we have

$$\left\langle p_i \frac{\partial \mathcal{H}}{\partial p_i} \right\rangle = kT \tag{6.35}$$

For $i = j$ and $x_i = q_i$, we have

$$\left\langle q_i \frac{\partial \mathcal{H}}{\partial q_i} \right\rangle = kT \tag{6.36}$$

According to the canonical equations of motion, $\partial \mathcal{H} / \partial q_i = -\dot{p}_i$. Hence (6.36) leads to the statement

$$\left\langle \sum_{i=1}^{3N} q_i \dot{p}_i \right\rangle = -3NkT \tag{6.37}$$

which is known as the virial theorem because $\sum q_i \dot{p}_i$ —the sum of the i th coordinate times the i th component of the generalized force—is known in classical mechanics as the virial.

Many physical systems have **Hamiltonians** that, through a canonical transformation, can be cast in the form

$$\mathcal{H} = \sum_i A_i P_i^2 + \sum_i B_i Q_i^2 \tag{6.38}$$

where P_i, Q_i are canonically conjugate variables and A_i, B_i are constants. For such systems we have

$$\sum_i \left(P_i \frac{\partial \mathcal{H}}{\partial P_i} + Q_i \frac{\partial \mathcal{H}}{\partial Q_i} \right) = 2\mathcal{H} \tag{6.39}$$

Suppose f of the constants A_i and B_i are nonvanishing. Then (6.35) and (6.36) imply that

$$\langle \mathcal{H} \rangle = \frac{1}{2} f kT \tag{6.40}$$

That is, each harmonic term in the Hamiltonian contributes $\frac{1}{2} kT$ to the average energy of the system. This is known as the theorem of equipartition of energy. But (6.40) is the internal energy of the system. Therefore

$$\frac{C_V}{k} = \frac{f}{2} \tag{6.41}$$

Thus the heat capacity is directly related to the number of degrees of freedom of the system.

A paradox arises from the theorem of equipartition of energy. In classical physics every system must in the last analysis have an infinite number of degrees of freedom, for after we have resolved matter into atoms we must continue to resolve an atom into its constituents and the constituents of the constituents, ad infinitum. Therefore the heat capacity of any system is infinite. This is a real paradox in classical physics and is resolved by quantum mechanics. Quantum mechanics possesses the feature that the degree of freedom of a system are manifest only when there is sufficient energy to excite them, and that those degrees of freedom that are not excited can be forgotten. Thus the formula (6.41) is valid only when the temperature is sufficiently high.

6.5 CLASSICAL IDEAL GAS

To illustrate the method of calculation in the microcanonical ensemble we consider the classical ideal gas. This has been considered earlier in our discussion of the kinetic theory of gases. In that discussion we also introduced the microcanonical ensemble, but we obtained all the thermodynamic properties of the ideal gas via the distribution function. For the sake of illustration, we now derive the same results using the recipe given in Section 6.3.

The Hamiltonian is

$$\mathcal{H} = \frac{1}{2m} \sum_{i=1}^N p_i^2 \tag{6.42}$$

We first calculate

$$\sum(E) = \frac{1}{h^{3N}} \int_{\mathcal{H} < E} d^3p_1 \cdots d^3p_N d^3q_1 \cdots d^3q_N \quad (6.43)$$

where h is a constant of the dimension of momentum \times distance, introduced to make $\sum(E)$ dimensionless. The integration over q_i can be immediately carried out, giving a factor of V^N . Let

$$R = \sqrt{2mE} \quad (6.44)$$

Then

$$\sum(E) = \left(\frac{V}{h^3}\right)^N \Omega_{3N}(R) \quad (6.45)$$

where Ω_n is the volume of an n -sphere of radius R :

$$\Omega_n(R) = \int_{x_1^2 + x_2^2 + \cdots + x_n^2 < R^2} dx_1 dx_2 \cdots dx_n \quad (6.46)$$

Clearly,

$$\Omega_n(R) = C_n R^n \quad (6.47)$$

where C_n is a constant. To find C_n , consider the identity

$$\int_{-\infty}^{+\infty} dx_1 \cdots \int_{-\infty}^{+\infty} dx_n e^{-(x_1^2 + \cdots + x_n^2)} = \left(\int_{-\infty}^{+\infty} dx e^{-x^2}\right)^n = \pi^{n/2} \quad (6.48)$$

The left side of (6.48) can be re-expressed as follows. Let $S_n(R) \equiv d\Omega_n(R)/dR$ be the surface area of an n -sphere of radius R . Then

$$\begin{aligned} \int_{-\infty}^{+\infty} dx_1 \cdots \int_{-\infty}^{+\infty} dx_n e^{-(x_1^2 + \cdots + x_n^2)} &= \int_0^\infty dR S_n(R) e^{-R^2} \\ &= nC_n \int_0^\infty dR R^{n-1} e^{-R^2} \\ &= \frac{1}{2} nC_n \int_0^\infty dt t^{(n/2)-1} e^{-t} = \frac{1}{2} nC_n \Gamma(n/2) \end{aligned} \quad (6.49)$$

where $\Gamma(z)$ is the gamma function. Comparison of (6.49) and (6.48) yields

$$C_n = \frac{\pi^{n/2}}{\Gamma(n/2 + 1)} \quad (6.50)$$

$$\log C_n \xrightarrow{n \rightarrow \infty} \frac{n}{2} \log \pi - \frac{n}{2} \log \frac{n}{2} + \frac{n}{2} \quad (6.51)$$

Hence

$$\sum(E) = C_{3N} \left[\frac{V}{h^3} (2mE)^{3/2} \right]^N \quad (6.52)$$

The entropy of the ideal gas is

$$S(E, V) = k \left[\log C_{3N} + N \log \frac{V}{h^3} + \frac{3}{2} N \log(2mE) \right] \quad (6.53)$$

By (6.51), this reduces to

$$S(E, V) = Nk \log \left[V \left(\frac{4\pi m E}{3h^2 N} \right)^{3/2} \right] + \frac{3}{2} Nk \quad (6.54)$$

Solving for E in terms of S and V , and calling the resulting function $U(S, V)$ the internal energy, we obtain

$$U(S, V) = \left(\frac{3}{4\pi} \frac{h^2}{m} \right) \frac{N}{V^{2/3}} \exp \left(\frac{2}{3} \frac{S}{Nk} - 1 \right) \quad (6.55)$$

The temperature is

$$T = \left(\frac{\partial U}{\partial S} \right)_V = \frac{2}{3} \frac{U}{Nk} \quad (6.56)$$

from which follows

$$C_V = \frac{3}{2} Nk \quad (6.57)$$

Finally the equation of state is

$$P = - \left(\frac{\partial U}{\partial V} \right)_S = \frac{2}{3} \frac{U}{V} = \frac{NkT}{V} \quad (6.58)$$

This calculation shows that the microcanonical ensemble is clumsy to use. There seems little hope that we can straightforwardly carry out the recipe of the microcanonical ensemble for any system but the ideal gas. We later introduce the canonical ensemble, which gives results equivalent to those of the microcanonical ensemble but which is more convenient for practical calculations.

6.6 GIBBS PARADOX

According to (6.54), the entropy of an ideal gas is

$$S = Nk \log(Vu^{3/2}) + Ns_0 \quad (6.59)$$

where

$$\begin{aligned} u &= \frac{3}{2} kT \\ s_0 &= \frac{3k}{2} \left(1 + \log \frac{4\pi m}{3h^2} \right) \end{aligned} \quad (6.60)$$

Consider two ideal gases, with N_1 and N_2 particles, respectively, kept in two separate volumes V_1 and V_2 at the same temperature and the same density. Let us find the change in entropy of the combined system after the gases are allowed to mix in a volume $V = V_1 + V_2$. The temperature will be the same after the mixing

process. Hence u remains unchanged. From (6.59) we find that the change in entropy is

$$\frac{\Delta S}{k} = N_1 \log \frac{V}{V_1} + N_2 \log \frac{V}{V_2} > 0 \quad (6.61)$$

which is the entropy of mixing. If the two gases are different (e.g., argon and neon), this result is experimentally correct.

The Gibbs paradox presents itself if we consider the case in which the two mixing gases are of the same kind. Since the derivation of (6.61) does not depend on the identity of the gases, we would obtain the same increase of entropy (6.61). This is a disastrous result because it implies that the entropy of a gas depends on the history of the gas, and thus cannot be a function of the thermodynamic state alone. Worse, the entropy does not exist, because we can always imagine that the existing state of a gas is arrived at by pulling off any number of partitions that initially divided the gas into any number of compartments. Hence S is larger than any number.

Gibbs resolved the paradox in an empirical fashion by postulating that we have made an error in calculating $\sum(E)$, the number of states of the gas with energy less than E . Gibbs assumed that the correct answer is $N!$ times smaller than we thought it was. By this assumption we should subtract from (6.59) the term $\log N! \approx N \log N - N$ and obtain

$$S = Nk \log \left(\frac{V}{N} u^{3/2} \right) + \frac{3}{2} Nk \left(\frac{5}{3} + \log \frac{4\pi m}{3h^2} \right) \quad (6.62)$$

This formula does not affect the equation of state and other thermodynamic functions of a system, because the subtracted term is independent of T and V . For the mixing of two different gases (6.62) still predicts (6.61), because N_1 and N_2 are the same constants before and after the mixing. For the mixing of gases that are of the same kind, however, it gives no entropy of mixing because the specific volume V/N is the same before and after mixing.

The formula (6.62) has been experimentally verified as the correct entropy of an ideal gas at high temperatures, if h is numerically set equal to Planck's constant. It is known as the Sackur-Tetrode equation.

It is not possible to understand classically why we must divide $\sum(E)$ by $N!$ to obtain the correct counting of states. The reason is inherently quantum mechanical. Quantum mechanically, atoms are inherently indistinguishable in the following sense: A state of the gas is described by an N -particle wave function, which is either symmetric or antisymmetric with respect to the interchange of any two particles. A permutation of the particles can at most change the wave function by a sign, and it does not produce a new state of the system. From this fact it seems reasonable that the Γ -space volume element $dpdq$ corresponds to not one but only $dpdq/N!$ states of the system. Hence we should divide $\sum(E)$ by $N!$. This rule of counting is known as the "correct Boltzmann counting." It is something that we must append to classical mechanics to get right answers.

The foregoing discussion contains the correct reason for, but is not a derivation of, the "correct Boltzmann counting," because in classical mechanics there is no consistent way in which we can regard the particles as indistinguishable. In all classical considerations other than the counting of states we must continue to regard the particles in a gas as distinguishable.

We may derive the "correct Boltzmann counting" by showing that in the limit of high temperatures quantum statistical mechanics reduces to classical statistical mechanics with "correct Boltzmann counting." This is done in Section 9.2.

PROBLEMS

- 6.1** Show that the formulas (6.27), (6.28), and (6.29) are equivalent to one another.
- 6.2** Let the "uniform" ensemble of energy E be defined as the ensemble of all systems of the given type with energy less than E . The equivalence between (6.29) and (6.27) means that we should obtain the same thermodynamic functions from the "uniform" ensemble of energy E as from the microcanonical ensemble of energy E . In particular, the internal energy is E in both ensembles. Explain why this seemingly paradoxical result is true.
- 6.3** Consider a system of N free particles in which the energy of each particle can assume two and only two distinct values, 0 and E ($E > 0$). Denote by n_0 and n_1 the occupation numbers of the energy level 0 and E , respectively. The total energy of the system is U .
- (a) Find the entropy of such a system.
 - (b) Find the most probable values of n_0 and n_1 , and find the mean square fluctuations of these quantities.
 - (c) Find the temperature as a function of U , and show that it can be negative.
 - (d) What happens when a system of negative temperature is allowed to exchange heat with a system of positive temperature?
- Reference. N. F. Ramsey, *Phys. Rev.* 103, 20 (1956).
- 6.4** Using the corrected entropy formula (6.62), work out the entropy of mixing for the case of different gases and for the case of identical gases, thus showing explicitly that there is no Gibbs paradox.

CHAPTER

CANONICAL ENSEMBLE AND GRAND CANONICAL ENSEMBLE

7.1 CANONICAL ENSEMBLE

We wish to consider the question, "What ensemble is appropriate for the description of a system not in isolation, but in thermal equilibrium with a larger system?" To answer it we must find the probability that the system has energy E , because this probability is proportional to the density in Γ space for the ensemble we want.

We investigated a similar problem in Section 6.2, when we consider the energies of the component parts of a composite system. In the following we discuss the case in which one component part is much smaller than the other.

Consider an isolated composite system made up of two subsystems whose Hamiltonians are, respectively, $\mathcal{H}_1(p_1, q_1)$ and $\mathcal{H}_2(p_2, q_2)$, with number of particles N_1 and N_2 , respectively. We assume that $N_2 \gg N_1$ but that both N_1 and N_2 are macroscopically large. We are interested in system 1 only. Consider a microcanonical ensemble of the composite system with total energy between E and $E + 2\Delta$. The energies E_1 and E_2 of the subsystems accordingly can have any values satisfying

$$E < (E_1 + E_2) < E + 2\Delta \quad (7.1)$$

Although this includes a range of values of E_1, E_2 , the analysis of Section 6.2 shows that only one set of values, namely \bar{E}_1, \bar{E}_2 , is important. We assume that $\bar{E}_2 \gg \bar{E}_1$. Let $\Gamma_2(E_2)$ be the volume occupied by system 2 in its own Γ space. The probability of finding system 1 in a state within dp, dq_1 of (p_1, q_1) , regardless of the state of system 2, is proportional to $dp_1 dq_1 \Gamma_2(E_2)$, where $E_2 = E - E_1$. Therefore up to a proportionality constant the density in Γ space

for system 1 is

$$\rho(p_1, q_1) \propto \Gamma_2(E - E_1) \quad (7.2)$$

Since only the values near $E_1 = \bar{E}_1$ are expected to be important, and $\bar{E}_1 \ll E$, we may perform the expansion

$$\begin{aligned} k \log \Gamma_2(E - E_1) &= S_2(E - E_1) = S_2(E) - E_1 \left[\frac{\partial S_2(E_2)}{\partial E_2} \right]_{E_2=E} + \dots \\ &\approx S_2(E) - \frac{E_1}{T} \end{aligned} \quad (7.3)$$

where T is the temperature of the larger subsystem. Hence

$$\Gamma_2(E - E_1) \approx \exp \left[\frac{1}{k} S_2(E) \right] \exp \left(- \frac{E_1}{kT} \right) \quad (7.4)$$

The first factor is independent of E_1 and is thus a constant as far as the small subsystem is concerned. Owing to (7.2) and the fact that $E_1 = \mathcal{H}_1(p_1, q_1)$, we may take the ensemble density for the small subsystem to be

$$\rho(p, q) = e^{-\mathcal{H}(p, q)/kT} \quad (7.5)$$

where the subscript 1 labeling the subsystem has been omitted, since we may now forget about the larger subsystem, apart from the information that its temperature is T . The larger subsystem in fact behaves like a heat reservoir in thermodynamics. The ensemble defined by (7.5), appropriate for a system whose temperature is determined through contact with a heat reservoir, is called the canonical ensemble.

The volume in Γ space occupied by the canonical ensemble is called the partition function:

$$Q_N(V, T) \equiv \int \frac{d^{3N}p d^{3N}q}{N! h^{3N}} e^{-\beta \mathcal{H}(p, q)} \quad (7.6)$$

where $\beta = 1/kT$, and where we have introduced a constant h , of the dimension of momentum \times distance, in order to make Q_N dimensionless. The factor $1/N!$ appears, in accordance with the rule of "correct Boltzmann counting." These constants are of no importance for the equation of state.

Strictly speaking we should not integrate over the entire Γ space in (7.6), because (7.2) requires that $\rho(p_1, q_1)$ vanish if $E_1 > E$. The justification for ignoring such a restriction is that in the integral (7.6) only one value of the energy $\mathcal{H}(p, q)$ contributes to the integral and that this value will lie in the range where the approximation (7.4) is valid. We prove this contention in Section 7.2.

The thermodynamics of the system is to be obtained from the formula

$$Q_N(V, T) = e^{-\beta A(V, T)} \quad (7.7)$$

where $A(V, T)$ is the Helmholtz free energy. To justify this identification we show

that

- (a) A is an extensive quantity,
 (b) A is related to the internal energy $U \equiv (H)$ and the entropy $S \equiv -(\partial A / \partial T)_V$ by the thermodynamic relation

$$A = U - TS$$

That A is an extensive quantity follows from (7.6), because if the system is made up of two subsystems whose mutual interaction can be neglected, then Q_N is a product of two factors. To prove the relation (b), we first convert (b) into the following differential equation for A :

$$\langle \mathcal{H} \rangle = A - T \left(\frac{\partial A}{\partial T} \right)_V \quad (7.8)$$

To prove (7.8), note the identity

$$\frac{1}{N!h^{3N}} \int dp dq e^{\beta[A(V,T) - \mathcal{H}(p,q)]} = 1 \quad (7.9)$$

Differentiating with respect to β on both sides, we obtain

$$\frac{1}{N!h^{3N}} \int dp dq e^{\beta[A(V,T) - \mathcal{H}(p,q)]} \left[A(V,T) - \mathcal{H}(p,q) + \beta \left(\frac{\partial A}{\partial \beta} \right)_V \right] = 0$$

This is the same as

$$A(V,T) - U(V,T) - T \left(\frac{\partial A}{\partial T} \right)_V = 0$$

All other thermodynamic functions may be found from $A(V,T)$ by the Maxwell relations in thermodynamics:

$$P = - \left(\frac{\partial A}{\partial V} \right)_T$$

$$S = - \left(\frac{\partial A}{\partial T} \right)_V$$

$$G = A + PV$$

$$U = (H) = A + TS$$

Therefore all calculations in the canonical ensembles begin (and nearly end) with the calculation of the partition function (7.6).

7.2 ENERGY FLUCTUATIONS IN THE CANONICAL ENSEMBLE

We now show that the canonical ensemble is mathematically equivalent to the microcanonical ensemble in the sense that although the canonical ensemble contains systems of all energies the overwhelming majority of them have the same

energy. To do this we calculate the mean square fluctuation of energy in the canonical ensemble. The average energy is

$$U = \langle \mathcal{H} \rangle = \frac{\int dp dq \mathcal{H} e^{-\beta \mathcal{H}}}{\int dp dq e^{-\beta \mathcal{H}}} \quad (7.10)$$

Hence

$$\int dp dq [U - \mathcal{H}(p,q)] e^{\beta[A(V,T) - \mathcal{H}(p,q)]} = 0 \quad (7.11)$$

Differentiating both sides with respect to β , we obtain

$$\frac{\partial U}{\partial \beta} + \int dp dq e^{\beta(A - \mathcal{H})} (U - \mathcal{H}) \left(A - \mathcal{H} - T \frac{\partial A}{\partial T} \right) = 0 \quad (7.12)$$

By (7.8) this can be rewritten in the form

$$\frac{\partial U}{\partial \beta} + \langle (U - \mathcal{H})^2 \rangle = 0 \quad (7.13)$$

Therefore the mean square fluctuation of energy is

$$\langle \mathcal{H}^2 \rangle - \langle \mathcal{H} \rangle^2 = \langle (U - \mathcal{H})^2 \rangle = - \frac{\partial U}{\partial \beta} = kT^2 \frac{\partial U}{\partial T}$$

or

$$\langle \mathcal{H}^2 \rangle - \langle \mathcal{H} \rangle^2 = kT^2 C_V \quad (7.14)$$

For a macroscopic system $\langle \mathcal{H} \rangle$ a N and C_V a N . Hence (7.14) is a normal fluctuation. As $N \rightarrow \infty$, almost all systems in the ensemble have the energy $\langle \mathcal{H} \rangle$, which is the internal energy. Therefore the canonical ensemble is equivalent to the microcanonical ensemble.

It is instructive to calculate the fluctuations in another way. We begin by calculating the partition function in the following manner:

$$\begin{aligned} \frac{1}{N!h^{3N}} \int dp dq e^{-\beta \mathcal{H}(p,q)} &= \int_0^\infty dE \omega(E) e^{-\beta E} = \int_0^\infty dE e^{-\beta E + \log \omega(E)} \\ &= \int_0^\infty dE e^{\beta[TS(E) - E]} \end{aligned} \quad (7.15)$$

where S is the entropy defined in the microcanonical ensemble. Since both S and U are proportional to N , the exponent in the last integrand is enormous. We expect that as $N \rightarrow \infty$ the integral receives contribution only from the neighborhood of the maximum of the integrand. The maximum of the integrand occurs at $E = \bar{E}$, where \bar{E} satisfies the conditions

$$T \left(\frac{\partial S}{\partial E} \right)_{E=\bar{E}} = 1 \quad (7.16)$$

$$\left(\frac{\partial^2 S}{\partial E^2} \right)_{E=\bar{E}} < 0 \quad (7.17)$$

The first condition implies $\bar{E} = U$, the internal energy. Next we note that

$$\left(\frac{\partial^2 S}{\partial E^2}\right)_{E=\bar{E}} = \left(\frac{\partial}{\partial E} \frac{1}{T}\right)_{E=\bar{E}} = -\frac{1}{T^2} \left(\frac{\partial T}{\partial E}\right)_{E=\bar{E}} = -\frac{1}{T^2 C_V} \quad (7.18)$$

Thus the condition (7.17) is satisfied if $C_V > 0$, which is true for physical systems. Now let us expand the exponent in (7.15) about $E = \bar{E}$:

$$\begin{aligned} TS(E) - E &= [TS(\bar{E}) - \bar{E}] + \frac{1}{2}(E - \bar{E})^2 T \left(\frac{\partial^2 S}{\partial E^2}\right)_{E=\bar{E}} + \dots \\ &= [TS(U) - U] - \frac{1}{2TC_V}(E - U)^2 + \dots \end{aligned} \quad (7.19)$$

Hence

$$\frac{1}{N!h^{3N}} \int dp dq e^{-\beta \mathcal{H}(p, q)} \approx e^{\beta(TS - U)} \int_0^\infty dE e^{-(E - U)^2 / 2kT^2 C_V} \quad (7.20)$$

showing that in the canonical ensemble the distribution in energy is a Gaussian distribution centered about the value $E = U$ with a width equal to

$$\Delta E = \sqrt{2kT^2 C_V} \quad (7.21)$$

Since U is a N and C_V is a N , $\Delta E/U$ is negligibly small. As $N \rightarrow \infty$ the Gaussian approaches a δ -function. Finally, let us perform the integral in (7.20). It is elementary:

$$\begin{aligned} \int_0^\infty dE e^{-(E - U)^2 / 2kT^2 C_V} &= \int_{-U}^\infty dx e^{-x^2 / 2kT^2 C_V} \\ &\approx \int_{-\infty}^{+\infty} dx e^{-x^2 / 2kT^2 C_V} = \sqrt{2\pi kT^2 C_V} \end{aligned}$$

Therefore

$$\frac{1}{N!h^{3N}} \int dp dq e^{-\beta \mathcal{H}(p, q)} \approx e^{\beta(TS - U)} \sqrt{2\pi kT^2 C_V} \quad (7.22)$$

$$A \approx (U - TS) - \frac{1}{2}kT \log C_V \quad (7.23)$$

This last term is negligible when $N \rightarrow \infty$. In that limit we have exactly $A = U - TS$. Statement (7.23) shows that the entropy as defined in the canonical and microcanonical ensemble differs only by terms of the order of $\log N$.

We have shown that almost all systems in the canonical ensemble have the same energy—namely, the energy that is equal to the internal energy of a system at the given temperature T . The reason for this is easy to see, both mathematically and physically.

In the canonical ensemble we distribute systems in Γ space according to the density function $\rho(p, q) = \exp[-\beta \mathcal{H}(p, q)]$, which is represented in Fig. 7.1. The density of points falls off exponentially as we go away from the origin of Γ space. The distribution in energy is obtained by "counting" the number of points

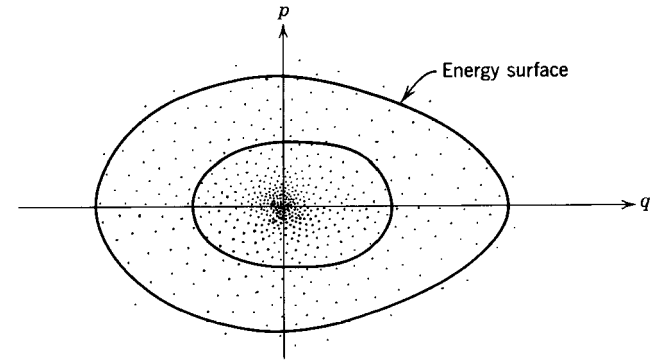


Fig. 7.1 Distribution of representative points in Γ space for the canonical ensemble.

on energy surfaces. As we go away from the origin, the energy increases and the area of the energy surface increases. This is why we get a peak in the distribution in energy. The sharpness of the peak is due to the rapidity with which the area of the energy surface increases as E increases. For an N -body system this area increases like e^E , where E is a N .

From a physical point of view, a microcanonical ensemble must be equivalent to a canonical ensemble, otherwise we would seriously doubt the utility of either. A macroscopic substance has the extensive property, i.e., any part of the substance has the same thermodynamic property as the whole substance. Now consider a piece of substance isolated from everything. Any part of the substance must still be in equilibrium with the rest of the substance, which serves as a heat reservoir that defines a temperature for the part on which we focus our attention. Therefore the whole substance must have a well-defined temperature.

We have seen earlier that in the microcanonical ensemble it matters little whether we take the entropy to be k times the logarithm of the density of states at the energy E , the number of states with energies between E , $E + A$, or all the states with energy below E . In all these cases we arrive at the same thermodynamic behavior. Now we see that it matters little whether we specify the energy of the system or the temperature of the system, for specifying one fixes the other, and we find the same thermodynamic behavior in both cases. All these examples illustrate the insensitivity of thermodynamic results to methods of derivation. The reasons behind this insensitivity are, in all cases, the facts that

- (a) density of states is e^E
- (b) $E \propto N$
- (c) $N \rightarrow \infty$

On these facts depends the validity of statistical mechanics.

7.3 GRAND CANONICAL ENSEMBLE

Although the canonical and the microcanonical ensemble give equivalent results, it may be argued that conceptually the canonical ensemble corresponds more closely to physical situations. In experiments we never deal with a completely isolated system, nor do we ever directly measure the total energy of a macroscopic system. We usually deal with systems with a given temperature—a parameter that we can control in experiments.

By the same token we should not have to specify exactly the number of particles of a macroscopic system, for that is never precisely known. All we can find out from experiments is the average number of particles. This is the motivation for introducing the grand canonical ensemble, in which the systems can have any number of particles, with the average number determined by conditions external to the system. This is analogous to the situation in the canonical ensemble, where the average energy of a system is determined by the temperature of the heat reservoir with which it is in contact.

The Γ space for the grand canonical ensemble is spanned by all the canonical momenta and coordinates of systems with $0, 1, 2, \dots$ number of particles. The density function describing the distribution of representative points in Γ space is denoted by $\rho(p, q, N)$, which gives the density of points representing systems with N particles with the momenta and coordinates (p, q) . To find $\rho(p, q, N)$ we consider the canonical ensemble for a system with N particles, volume V , and temperature T , but we focus our attention on a small subvolume V_1 of the system.

Suppose there are N_1 particles in V_1 and $N_2 = N - N_1$ particles in $V_2 = V - V_1$. We assume

$$\begin{aligned} V_2 &\gg V_1 \\ N_2 &\gg N_1 \end{aligned}$$

and designate the coordinates of the N_1 particle in V_1 by $\{p_1, q_1\}$, and those in V_2 by $\{p_2, q_2\}$. The interactions between particles in V_1 with those in V_2 are surface effects that can be neglected if V_1 is of macroscopic size. Thus the total Hamiltonian can be decomposed in the form

$$\mathcal{H}(p, q, N) = \mathcal{H}(p_1, q_1, N_1) + \mathcal{H}(p_2, q_2, N_2) \quad (7.24)$$

Note that the terms above involve the same function evaluated at different values of its arguments. The partition function of the total system is

$$Q_N(V, T) = \int \frac{dp dq}{h^{3N} N!} e^{-\beta \mathcal{H}(p, q, N)}$$

We shall segregate the contributions to the above from different values of N . In so doing, we do not care which particles are in V_1 as long as there are N_1 of them, and the coordinates of those that happen to be in V_1 will be designated by $\{p_1, q_1\}$. That is, in carrying out the integration over the phase space of the N -particle system, we always designate by $\{p, q\}$ the coordinates of those

particles that happen to be in V_1 , through a change in the variables of integration if necessary. Thus

$$\begin{aligned} Q_N(V, T) &= \frac{1}{h^{3N} N!} \int dp_1 dp_2 \sum_{N_1=0}^N \frac{N!}{N_1! N_2!} \\ &\quad \times \int_{v_1} dq_1 \int_{v_2} dq_2 e^{-\beta[\mathcal{H}(p_1, q_1, N_1) + \mathcal{H}(p_2, q_2, N_2)]} \\ &= \sum_{N_1=0}^N \frac{1}{h^{3N_1} N_1!} \int dp_1 \int_{v_1} dq_1 e^{-\beta \mathcal{H}(p_1, q_1, N_1)} \frac{1}{h^{3N_2} N_2!} \\ &\quad \times \int dp_2 \int_{v_2} dq_2 e^{-\beta \mathcal{H}(p_2, q_2, N_2)} \end{aligned} \quad (7.25)$$

The relative probability $\rho(p_1, q_1, N_1)$ that there are N_1 particles in V_1 with coordinates $\{p, q_1\}$ is proportional to the summand of $\int dp, dq, \sum_{N_1}$. We choose its normalization such that

$$\rho(p_1, q_1, N_1) = \frac{1}{Q_N(V, T)} \frac{e^{-\beta \mathcal{H}(p_1, q_1, N_1)}}{h^{3N_1} N_1!} \frac{1}{h^{3N_2} N_2!} \int dp_2 \int_{v_2} dq_2 e^{-\beta \mathcal{H}(p_2, q_2, N_2)} \quad (7.26)$$

The first factor above is chosen so that

$$\sum_{N_1=0}^N \int dp_1 dq_1 \rho(p_1, q_1, N_1) = 1 \quad (7.27)$$

which is obvious if we rewrite the last expression in (7.25) in terms of ρ as defined in (7.26).

We can rewrite (7.26) in the form

$$\rho(p_1, q_1, N_1) = \frac{Q_{N_2}(V_2, T)}{Q_N(V, T)} \frac{e^{-\beta \mathcal{H}(p_1, q_1, N_1)}}{h^{3N_1} N_1!} \quad (7.28)$$

Using (7.7) we write

$$\frac{Q_{N_2}(V_2, T)}{Q_N(V, T)} = \exp\{-\beta[A(N - N_1, V - V_1, T) - A(N, V, T)]\} \quad (7.29)$$

where $A(N, V, T)$ is the Helmholtz free energy. Since $N \gg N_1$ and $V \gg V_1$, we may use the approximation

$$A(N - N_1, V - V_1, T) - A(N, V, T) \approx -N_1 \mu + V_1 P \quad (7.30)$$

where μ and P are, respectively, the chemical potential and the pressure of the

part of the system external to the small volume V_1 :

$$\begin{aligned}\mu &= \left[\frac{\partial A(N_2, V, T)}{\partial N_2} \right]_{N_2=N} \\ P &= - \left[\frac{\partial A(N, V_2, T)}{\partial V_2} \right]_{V_2=V}\end{aligned}\quad (7.31)$$

We now introduce the *fugacity*:

$$z = e^{\beta\mu} \quad (7.32)$$

Substituting (7.32) and (7.30) into (7.29), and then substituting (7.29) into (7.28), we obtain

$$\rho(p, q, N) = \frac{z^N}{N! h^{3N}} e^{-\beta PV - \beta \mathcal{H}(p, q)} \quad (7.33)$$

where the subscript 1 identifying the volume under consideration has been omitted because the system external to the volume can now be forgotten, apart from the information that it has the temperature T , pressure P , and chemical potential μ . We now allow the system external to the volume under consideration to become infinite in size. Then the range of N in (7.33) becomes

$$0 \leq N < \infty$$

The thermodynamic functions for the volume under consideration may be found as follows. First, the internal energy shall be the ensemble average of $\mathcal{H}(p, q)$. Second, the temperature, pressure, and chemical potential shall be respectively equal to T , P , μ . To show that this is a correct recipe, it suffices to remind ourselves that thermodynamics has been derived from the canonical ensemble. It is an elementary thermodynamic exercise to show that if a system is in equilibrium any part of the system must have the same T , P , μ as any other part; but this is the desired result.

To obtain a convenient formal recipe for finding all the thermodynamic functions we define the grand partition function as follows:

$$\mathcal{Q}(z, V, T) \equiv \sum_{N=0}^{\infty} z^N Q_N(V, T) \quad (7.34)$$

which in principle can be calculated from a knowledge of the Hamiltonian. Integrating both sides of (7.33) over all (p, q) for a given N , and then summing N from 0 to ∞ , we find that

$$\frac{PV}{kT} = \log \mathcal{Q}(z, V, T) \quad (7.35)$$

Thus the grand partition function directly gives the pressure as a function of z , V , and T . The average number \bar{N} of particles in the volume V is by definition the

ensemble average

$$\bar{N} \equiv \langle N \rangle = \frac{\sum_{N=0}^{\infty} N z^N Q_N(V, T)}{\sum_{N=0}^{\infty} z^N Q_N(V, T)} = z \frac{\partial}{\partial z} \log \mathcal{Q}(z, V, T) \quad (7.36)$$

The equation of state, which is the equation expressing P as a function of N , V , and T , is obtained by eliminating z between (7.35) and (7.36).

All other thermodynamic functions may be obtained from the internal energy:

$$U = - \frac{\partial}{\partial \beta} \log \mathcal{Q}(z, V, T) \quad (7.37)$$

After eliminating z with the help of (7.36), U becomes a function of N , V , and T . We can then use the formulas

$$\begin{aligned}C_V &= \left(\frac{\partial U}{\partial T} \right)_V \\ S &= \int_0^T dT \frac{C_V}{T} \\ A &= U - TS\end{aligned}$$

7.4 DENSITY FLUCTUATIONS IN THE GRAND CANONICAL ENSEMBLE

We now calculate the density fluctuations in the grand canonical ensemble. By differentiating (7.36) with respect to z , one can easily show

$$\langle N^2 \rangle - \langle N \rangle^2 = z \frac{\partial}{\partial z} z \frac{\partial}{\partial z} \log \mathcal{Q}(z, V, T) = kTV \frac{\partial^2 P}{\partial \mu^2} \quad (7.38)$$

where the last equality is obtained through the use of (7.34) and (7.36). To express the above in terms of conveniently measurable quantities, assume that the Helmholtz free energy of the system, being an extensive quantity, can be written in the form

$$A(N, V, T) = Na(v), \quad v \equiv V/N \quad (7.39)$$

where the temperature dependence of $a(v)$ has been suppressed for brevity. Then the two equations in (7.31) can be rewritten as

$$\begin{aligned}\mu &= a(v) - v \frac{\partial a(v)}{\partial v} \\ P &= - \frac{\partial a(v)}{\partial v}\end{aligned}\quad (7.40)$$

Regarding both μ and P as functions of v and T , we obtain from the above

$$\begin{aligned}\frac{\partial \mu}{\partial v} &= -v \frac{\partial^2 a(v)}{\partial v^2} \\ \frac{\partial P}{\partial \mu} &= \frac{\partial P / \partial v}{\partial \mu / \partial v} = \frac{1}{v}\end{aligned}\quad (7.41)$$

Hence

$$\frac{\partial^2 P}{\partial \mu^2} = -\frac{1}{v^2} \frac{\partial v}{\partial \mu} = \frac{1}{v^3} \frac{\partial^2 a / \partial v^2}{\partial v} = -\frac{1}{v^3} \frac{\partial P / \partial v}{\partial v} \quad (7.42)$$

Substituting this relation into (7.38), we finally obtain, after some minor re-writing,

$$\langle N^2 \rangle - \langle N \rangle^2 = \bar{N} k T \kappa_T / v, \quad \kappa_T = \frac{1}{v(-\partial P / \partial v)} \quad (7.43)$$

This shows that the density fluctuations are vanishing small in the thermodynamic limit, provided the isothermal compressibility κ_T is finite (i.e., not infinite). This is true except in the transition region of a first-order phase transition, including the critical point.

The relation (7.43) is similar to (7.14), whereby a fluctuation is related to an appropriate "susceptibility." In the case of energy fluctuations the relevant susceptibility is the specific heat at constant volume, and in the case of density fluctuations it is the isothermal compressibility. These are special cases of a more general rule known as the fluctuation-dissipation theorem, the historically earliest form of which is the Einstein relation (2.61) pertaining to Brownian motion.

The probability that a system in the grand canonical ensemble has N particles is proportional to

$$W(N) \equiv z^N Q_N(V, T) = \exp \beta [\mu N - A(N, V, T)] \quad (7.44)$$

where A is the Helmholtz free energy calculated from the canonical ensemble with N particles. When the density fluctuations are small, $W(N)$ is strongly peaked about $N = \bar{N}$, with a width of the order of $\sqrt{\bar{N}}$, and we may obtain the Helmholtz free energy directly from the grand partition function through the formula

$$A(\bar{N}, V, T) = kT \bar{N} \log z - kT \log \mathcal{Q}(z, V, T) \quad (7.45)$$

where z is to be eliminated through (7.36).

When $\partial P / \partial v = 0$, as happens at the critical point, the density fluctuations become very large, as is borne out experimentally by the phenomenon of critical opalescence. However, even in this case, (7.45) is still valid. To show this requires a more detailed analysis, which we postpone until the end of this chapter.

7.5 THE CHEMICAL POTENTIAL

Thermodynamics

The chemical potential μ is defined such that the Helmholtz free energy A changes by μdN , when the number of particles change by dN , at constant T and V . Hence

$$dA = -P dV - S dT + \mu dN \quad (7.46)$$

from which we can deduce a more general form of the first law of thermodynamics:

$$dU = -P dV + T dS + \mu dN \quad (7.47)$$

When μ is positive, it tends to drive N to smaller values, in order to lower the energy. Hence the name chemical potential.*

From (7.46) we can also deduce the change of the Gibbs free energy:

$$dG = -V dP - S dT + \mu dN \quad (7.48)$$

Thus we have the equivalent Maxwell relations

$$\mu = \left(\frac{\partial \hat{N}}{\partial N} \right)_{V, T} = \left(\frac{\partial \hat{G}}{\partial N} \right)_{P, T}$$

A useful result is the chemical potential of an ideal gas, which can be easily calculated from the partition function for an ideal gas:

$$\begin{aligned}Q_N &= \frac{1}{h^{3N} N!} \int dp dq \exp \left[\beta \sum_{i=1}^N p_i^2 / 2m \right] = \frac{1}{N!} \left(\frac{V}{\lambda^3} \right)^N \\ A &= \sqrt{2\pi \hbar^2 / mkT}\end{aligned}\quad (7.49)$$

Hence

$$\begin{aligned}A &= kT \log Q_N = -kTN \left[\log \left(\frac{V}{N\lambda^3} \right) + 1 \right] \\ \mu &= \partial A / \partial N = kT \log (\lambda^3 n)\end{aligned}\quad (7.50)$$

where n is the density.

Conservation of Particle Number

For ordinary matter, it makes sense to speak of a system of N atoms, because N is an effectively conserved quantity. The chemical potential may be viewed as the Lagrange multiplier to take that into account. The conservation law has its origin in the more fundamental law of baryon conservation, which states that the number of baryons (such as protons or neutrons) minus the number of anti-baryons is conserved. This means, for example, that a proton can be created or

*The name fugacity for $\exp(\beta\mu)$ has a dictionary meaning of "the tendency to flee," or "volatility." *The fugacity of pleasure, the frugility of beauty* (Samuel Johnson).

annihilated only in conjunction with an antiproton. At low temperatures the thermal energy is not sufficient to create pairs, nor are there antiprotons present. Thus the number of protons (and neutrons) is effectively conserved.

The same thing can be said about electrons, whose number appears to be conserved at low temperatures only because there is insufficient energy to create electron-positron pairs, and there are usually no positrons present. The truly conserved quantity is the number of electrons minus the number of positrons.

A correct description of matter at high temperatures must take into account the possibility of pair creation. There will be an average number of particles and antiparticles present in equilibrium, there will also be fluctuations about the average values. It is the difference between particle and antiparticle number that remains strictly constant and is determined by the initial conditions. For example, the reaction $e^+ + e^- \rightleftharpoons \gamma$ can occur in the interior of stars, and establishes the equilibrium density of electrons, positrons, and radiation.

The detailed mechanism for pair creation and annihilation is not relevant for the equilibrium situation, and affects only the relaxation time for the establishment of equilibrium. Thus, to treat the equilibrium situation we may describe the system in the grand canonical ensemble, using the effective Hamiltonian

$$\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2 - \mu(N_1 - N_2) \quad (7.51)$$

where the subscripts 1 and 2 refer to particle and antiparticle, and μ is the Lagrange multiplier introduced to enable us to treat N_1 and N_2 as unconstrained variables. The grand partition function is

$$\begin{aligned} \mathcal{Q} &= \sum_{N_1=0}^{\infty} \sum_{N_2=0}^{\infty} Q_{N_1} Q_{N_2} e^{\beta\mu(N_1 - N_2)} \\ &= \sum_{N_1=0}^{\infty} \sum_{N_2=0}^{\infty} \exp -\beta [A_{N_1} + A_{N_2} - \mu(N_1 - N_2)] \end{aligned} \quad (7.52)$$

where Q_N is a partition function, and A , the corresponding Helmholtz free energy. In the thermodynamic limit we keep only the largest term in the summand, with N_1 and N_2 determined by the conditions

$$\begin{aligned} \partial A_{N_1} / \partial N_1 &= \mu \\ \partial A_{N_2} / \partial N_2 &= -\mu \end{aligned} \quad (7.53)$$

We should calculate the Helmholtz free energies using relativistic kinematics. For the purpose of illustration, however, we shall pretend that the energy of a particle is $E = mc^2 + p^2/2m$. The inclusion of the rest energy is important, for we are concerned with processes that can convert it into other forms of energy. Thus we take the chemical potential to be (7.50) plus the rest energy. The conditions for equilibrium then become

$$\begin{aligned} kT \log(\lambda^3 n_1) + mc^2 &= \mu \\ kT \log(\lambda^3 n_2) + mc^2 &= -\mu \end{aligned} \quad (7.54)$$

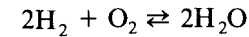
where n_1 and n_2 are the densities of particles and antiparticles. Adding the two equations gives

$$\lambda^6 n_1 n_2 = e^{-2mc^2/kT} \quad (7.55)$$

The rest energy of an electron corresponds to a temperature of 6×10^9 K. That for a proton is 2000 times higher. Thus the right side of (7.55) is essentially zero for ordinary temperatures. This means that, if either n_1 or n_2 is not essentially zero, the other must be. This shows why we can completely ignore antiparticles when $kT \ll mc^2$.

Chemical Equilibrium

Suppose we have a reaction such as



What are the fractions of each species of molecules present in an equilibrium mixture? More generally we consider a reaction in which a group of particles X_1, X_2, \dots participate in a reaction to yield a group of particles Y_1, Y_2, \dots or vice versa:

$$\nu_1 X_1 + \nu_2 X_2 + \dots \rightleftharpoons \nu'_1 Y_1 + \nu'_2 Y_2 + \dots \quad (7.56)$$

The numbers ν_i are called stoichiometric *coefficients*. The process is a generalization of our previous discussion of particle-antiparticle reaction, for which $\nu_i = \nu_2$ and $\nu'_i = 1$. For notational convenience, rewrite (7.56) as

$$\sum_{i=1}^K \nu_i X_i = 0 \quad (7.57)$$

where both X's and Y's are denoted by X , and $\nu'_i = -\nu_i$. The conservation law in this case is

$$\frac{\delta N_1}{\nu_1} = \frac{\delta N_2}{\nu_2} = \dots = \frac{\delta N_K}{\nu_K} \quad (7.58)$$

where δN_i is the increase in the number of particles of the i th type. This means $\delta N_i / \nu_i$ is independent of i :

$$\delta N_i = \nu_i \delta N \quad i = 1, \dots, K \quad (7.59)$$

To find the equilibrium condition consider the reaction proceeding at constant V and T . In equilibrium the Helmholtz energy A is at a minimum. Hence any variation of the number N_i from their equilibrium value will not change A to first order. Assume A is the sum of the component free energies. Then

$$0 = \delta A = \sum_{i=1}^K \delta A_i = \sum_{i=1}^K \frac{\partial A_i}{\partial N_i} \delta N_i = \sum_{i=1}^K \mu_i \nu_i \delta N \quad (7.60)$$

Since δN is arbitrary, we obtain as a condition for equilibrium

$$\sum_{i=1}^K \mu_i \nu_i = 0 \quad (7.61)$$

where μ_i is the chemical potential of X. If the reaction proceeds at constant P and T , one considers the Gibbs free energy instead, and arrives at the same condition.

7.6 EQUIVALENCE OF THE CANONICAL ENSEMBLE AND THE GRAND CANONICAL ENSEMBLE

We have seen that if $\partial P/\partial v < 0$ then almost all systems in the grand canonical ensemble have the same number of particles N . Then the grand canonical ensemble is trivially equivalent to the canonical ensemble for N particles.

To complete our investigation of the equivalence between the canonical and the grand canonical ensemble it is necessary to consider values of v for which $\partial P/\partial v = 0$. It will be shown that in such cases the function $W(N)$ given in (7.44) will no longer have a sharp maximum; the equation of state as given by the recipe in the grand canonical ensemble nevertheless still agrees with that given by the recipe in the canonical ensemble. In this sense the two ensembles are always equivalent.

Physically the values of v for which $\partial P/\partial v = 0$ correspond to the transition region of a first-order phase transition. In this region, (7.43) leads us to expect that the fluctuations of density in a given volume of the system will be large. This is also expected physically, for in such a region the system is composed of two or more phases of *different* densities. Therefore the number of particles in any given volume can have a whole range of values, depending on the amounts of each phase present. At the critical point of a gas-liquid system fluctuations in density are also expected to be large, because throughout the system molecules are spontaneously forming large clusters and breaking up. It is clear that under these conditions the grand canonical ensemble must continue to yield thermodynamic predictions that are in agreement with those obtained by the canonical ensemble. Otherwise the validity of either as a description of matter would be in doubt, for it is a basic experimental fact that we can obtain the same thermodynamic information whether we look at the whole system or at only a subvolume of the system.

The mathematical questions that we try to answer are as follows. Suppose $Q_N(V, T)$ is given, and we wish to calculate

$$\mathcal{Q}(z, V, T) \equiv \sum_{N=0}^{\infty} z^N Q_N(V, T) \quad (7.62)$$

for given values of z , V , and T .

(a) For a given value of z is the following true for some N ?

$$\mathcal{Q}(z, V, T) \approx z^N Q_N(V, T) \quad (7.63)$$

(b) Does there always exist a value of z for which N has any given positive value?

The answers are obviously no, if $Q_N(V, T)$ is *any* function of N, V, T . We are only interested, however, in the answers when $Q_N(V, T)$ is the partition function of a physical system. Thus we must first make some assumptions about $Q_N(V, T)$.

To incorporate the salient features of a physical system into our considerations, and yet keep the mathematics simple, we assume that we are dealing with a system

- (a) whose molecules interact through an intermolecular potential that contains a hard-sphere repulsion of finite diameter plus a finite potential of finite range, and
- (b) whose Helmholtz free energy has the form

$$A(N, V) \equiv -\frac{1}{\beta} \log Q_N(V) = -\frac{V}{\beta} f(v) \quad (7.64)$$

where $v \equiv V/N$, $\beta = 1/kT$, and $f(v)$ is finite. The temperature will be fixed throughout our discussions and will not be displayed unless necessary. The function $f(v)$ is related to the pressure $P(v)$ of the canonical ensemble by

$$f(v) = \frac{1}{v} \int_{v_0}^v dv' \beta P(v') \quad (7.65)$$

where the integration is carried out along an isotherm and v_0 is an arbitrary constant corresponding to an arbitrary additive constant in the Helmholtz free energy.

- (c) We further assume that $f(v)$ is such that

$$\frac{\partial P}{\partial v} \leq 0 \quad (7.66)$$

This immediately implies that

$$\frac{\partial^2 f(v)}{\partial (1/v)^2} \leq 0 \quad (7.67)$$

With these assumptions the grand partition function may be written in the form

$$\mathcal{Q}(z, V) = \sum_{N=0}^{\infty} e^{V\phi(V/N, z)} \quad (7.68)$$

where z is an arbitrary fixed number and

$$\phi(v, z) \equiv f(v) + \frac{1}{v} \log z \quad (7.69)$$

Using (7.65) we obtain

$$\phi(v, z) = \frac{1}{v} \log z + \frac{1}{v} \int_{v_0}^v dv' \beta P(v') \quad (7.70)$$

By (7.67), we have $\partial^2 \phi / \partial (1/v)^2 \leq 0$, or

$$\frac{\partial^2 \phi}{\partial v^2} + \frac{2}{v} \frac{\partial \phi}{\partial v} \leq 0 \quad (7.71)$$

We now calculate the grand partition function. For a fixed volume V the partition function $Q_N(V)$ vanishes whenever

$$N > N_0(V)$$

where $N_0(V)$ is the maximum number of particles that can be accommodated in the volume V , such that no two particles are separated by a distance less than the diameter of the hard sphere in the interparticle potential. Therefore $\mathcal{Q}(z, V)$ is a polynomial of degree $N_0(V)$. For large V it is clear that

$$N_0(V) = aV \quad (7.72)$$

where a is a constant. Let the largest value among the terms in this polynomial be $\exp[V\phi_0(z)]$, where

$$\phi_0(z) = \max \left[\phi \left(\frac{V}{N}, z \right) \right] \quad (N = 0, 1, 2, \dots) \quad (7.73)$$

Then the following inequality holds:

$$e^{V\phi_0(z)} \mathcal{I} \mathcal{Q}(z, v) \leq N_0(V) e^{V\phi_0(z)}$$

Using (7.72) we obtain

$$e^{V\phi_0(z)} \leq \mathcal{Q}(z, V) \leq aV e^{V\phi_0(z)}$$

or

$$\phi_0(z) \leq \frac{1}{V} \log \mathcal{Q}(z, V) \leq \phi_0(z) + \frac{\log(aV)}{V} \quad (7.74)$$

Therefore

$$\lim_{V \rightarrow \infty} \frac{1}{V} \log \mathcal{Q}(z, V) = \phi_0(z) \quad (7.75)$$

Let \bar{v} be a value of v at which $\phi(v, z)$ assumes its largest possible value. Since $\phi(v, z)$ is differentiable, \bar{v} is determined by the conditions

$$\left(\frac{\partial \phi}{\partial v} \right)_{v=\bar{v}} = 0 \quad (7.76)$$

$$\left(\frac{\partial^2 \phi}{\partial v^2} \right)_{v=\bar{v}} \leq 0 \quad (7.77)$$

By virtue of (7.71) the first condition implies the second. Therefore \bar{v} is de-

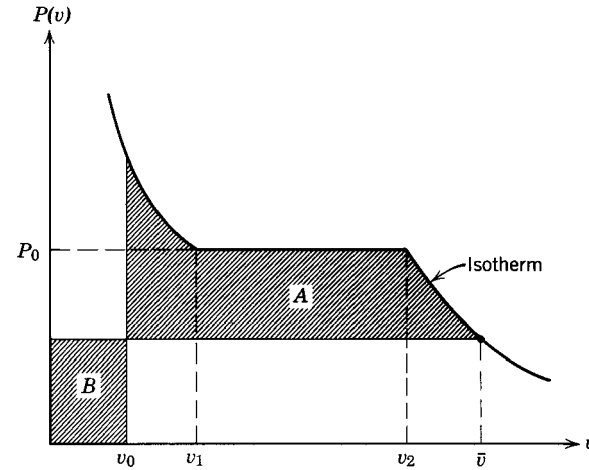


Fig. 7.2 Typical isotherm of a substance in the transition region of a first-order phase transition.

termined by (7.76) alone. By (7.69) and (7.65) we may rewrite it in the form

$$\int_{v_0}^{\bar{v}} dv' P(v') - \bar{v}P(\bar{v}) = -kT \log z$$

or

$$\left[\int_{v_0}^{\bar{v}} dv' P(v') - (v - v_0)P(\bar{v}) \right] - v_0P(\bar{v}) = -kT \log z \quad (7.78)$$

A geometrical representation of this condition is shown in Fig. 7.2. The value of \bar{v} is such that the difference between the area of the region A and that of the region B is numerically equal to $-kT \log z$. The result is shown in Fig. 7.3. It is seen that to every value of \bar{v} greater than the close-packing volume there corresponds a value of z . This answers question (b) in the affirmative.

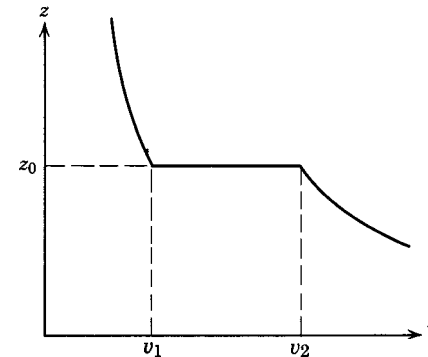


Fig. 7.3 z as a function of \bar{v} .

There is a value of z that corresponds to all the values of \bar{v} lying in the interval $v_1 \leq \bar{v} \leq v_2$. This value, denoted by z_0 , is given by

$$\log z_0 = \beta v_1 P(v_1) - \int_{v_0}^{v_1} dv' \beta P(v') \quad (7.79)$$

7.7 BEHAVIOR OF $W(N)$

In (7.44) we introduced the quantity $W(N)$, which is the (unnormalized) probability that a system in the grand canonical ensemble has N particles. Comparing (7.44) to (7.68) we see that

$$W(N) = \exp \left[V \phi \left(\frac{V}{N}, z \right) \right] \quad (7.80)$$

Hence it is of some interest to examine the function $\phi(v, z)$ in more detail. Suppose $P(v)$ has the form shown in the $P - v$ diagram of Fig. 7.2. For values of v lying in the range $v_1 \leq v \leq v_2$, P has the constant value P_0 . For this range of v we have

$$\phi(v, z) = \frac{1}{v} \left[\log z + \int_{v_0}^{v_1} dv' \beta P(v') - \beta P_0 v_1 \right] + \beta P_0$$

which is the same as

$$\phi(v, z) = \frac{1}{v} \log \left(\frac{z}{z_0} \right) + \beta P_0 \quad (v_1 \leq v \leq v_2) \quad (7.81)$$

where z_0 is defined by (7.79). Hence we can immediately make a qualitative sketch of a family of curves, one for each z , for the function $\phi(v, z)$ in the interval $v_1 \leq v \leq v_2$. The result is shown in Fig. 7.4.

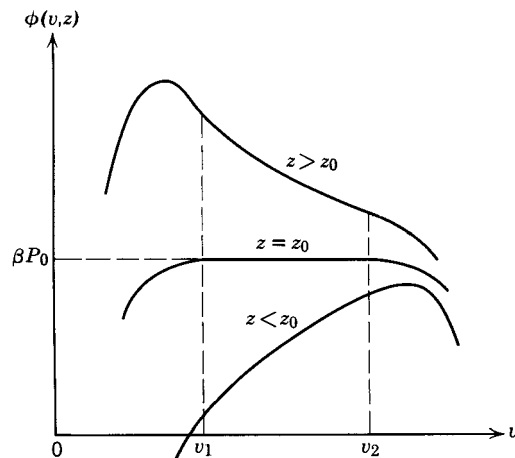


Fig. 7.4 Qualitative form of $\phi(v, z)$ for a physical substance.

To deduce the behavior of $\phi(v, z)$ outside the interval just discussed we use the following facts:

- (a) $\partial \phi / \partial v$ is everywhere continuous. This is implied by (7.70).
- (b) $\partial \phi / \partial v = 0$ implies $\partial^2 \phi / \partial v^2 \leq 0$. That is, as a function of v , ϕ cannot have a minimum. This follows from (7.71).
- (c) For $z \neq z_0$, ϕ has one and only one maximum. This follows from (b).

Guided by these facts we obtain the curves shown in Fig. 7.4.

The behavior of $W(N)$ can be immediately obtained from that of $\phi(v, z)$. It is summarized by the series of graphs in Fig. 7.5. For $z \neq z_0$, $W(N)$ has a single sharp peak at some value of N . This peak becomes infinitely sharp as $V \rightarrow \infty$. For $z = z_0$, all values of N in the interval

$$v_1 \leq \frac{V}{N} \leq v_2 \quad (7.82)$$

are equally probable. The number of N values corresponding to (7.82) is

$$\left(\frac{1}{v_1} - \frac{1}{v_2} \right) V \quad (7.83)$$

This situation corresponds to the large fluctuation of density in the transition

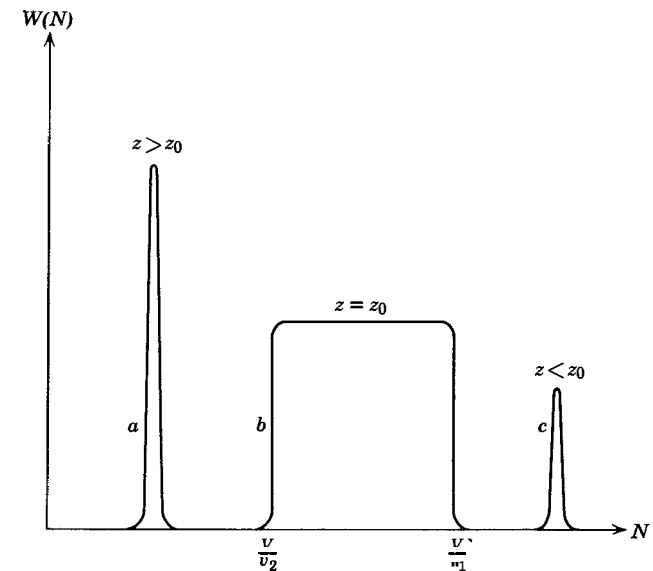


Fig. 7.5 The function $W(N)$ for three different fugacities (hence three different densities). For curves a and c the system is in a single pure phase. For curve b the system is undergoing a first-order phase transition.

region and may be stated in more physical terms as follows: The pressure is unchanged if we take any number of particles from one phase and deliver them to the other. Each time we do this, however, the total number of particles in a given volume changes, because the densities of the two phases are generally different. Let us start with the system in one pure phase and then transfer the particles one by one to the other phase, until the system exists purely in the other phase. The number of transfers we can make is proportional to V . Each transfer corresponds to a term in the grand partition function, and all these terms have the same value.

7.8 THE MEANING OF THE MAXWELL CONSTRUCTION

It has been shown that if the pressure P calculated in the canonical ensemble satisfies the condition $\partial P/\partial v \leq 0$, the pressure calculated in the grand canonical ensemble is also P . We show that the converse is also true. We then have the statement

(a) The pressure P calculated in the canonical ensemble agrees with that calculated in the grand canonical ensemble if and only if $\partial P/\partial v \leq 0$.

It will further be shown that

(b) If $\partial P/\partial v > 0$ for some v , the pressure in the grand canonical ensemble is obtainable from P by making the Maxwell construction.

Suppose the pressure calculated in the canonical ensemble is given and is denoted by $P_{\text{can}}(v)$. At a certain temperature we assume $P_{\text{can}}(v)$ to have the qualitative form shown in the $P - v$ diagram of Fig. 7.6.

The partition function of the system under consideration is

$$Q_N(V) = e^{VF(v)} \quad (7.84)$$

where

$$F(v) = \frac{1}{v} \int_{v_0}^v dv' \beta P_{\text{can}}(v') \quad (7.85)$$

It is easily seen that

$$\beta P_{\text{can}}(v) = F(v) + v \frac{\partial F(v)}{\partial v} \quad (7.86)$$

Let

$$\Phi(v, z) \equiv F(v) + \frac{1}{v} \log z \quad (7.87)$$

It is easily verified that

$$\frac{\partial^2 \Phi}{\partial v^2} + \frac{2}{v} \frac{\partial \Phi}{\partial v} = \frac{\beta}{v} \frac{\partial P_{\text{can}}}{\partial v} \begin{cases} > 0 & (a < v < b) \\ \leq 0 & (\text{otherwise}) \end{cases} \quad (7.88)$$

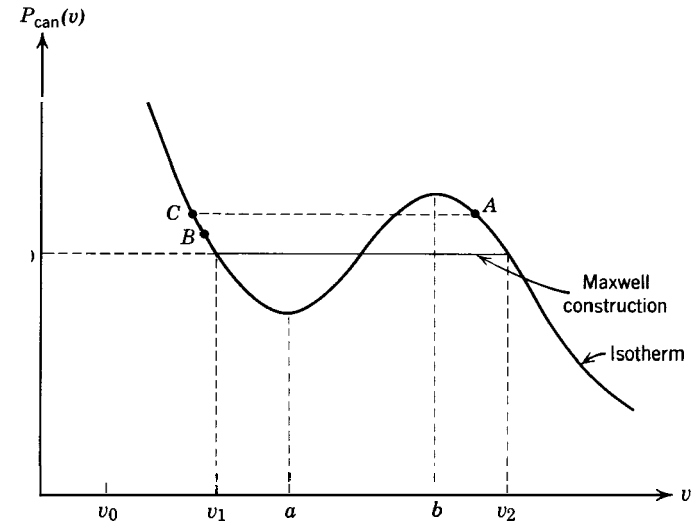


Fig. 7.6 Isotherm with $\partial P_{\text{can}}/\partial v > 0$ for v lying in the range $a < v < b$.

To calculate the grand partition function we recall that the derivation of (7.75) is independent of the sign of $\partial P/\partial v$. Hence, in analogy with (7.75), we have in the present case

$$\lim_{V \rightarrow \infty} \frac{1}{V} \log \mathcal{Q}(z, V) = \Phi(\bar{v}, z) \quad (7.89)$$

where

$$\Phi(\bar{v}, z) = \max [\Phi(v, z)] \quad (7.90)$$

This determines \bar{v} in terms of z , or vice versa. The pressure in the grand canonical ensemble, denoted by $P_{\text{gr}}(\bar{v})$, is given by

$$\beta P_{\text{gr}}(\bar{v}) = \Phi(\bar{v}, z) \quad (7.91)$$

From (7.87) and (7.85) we see that both Φ and $\partial \Phi/\partial v$ are continuous functions of v . Hence (7.90) is equivalent to the conditions

$$\begin{aligned} \left(\frac{\partial \Phi}{\partial v} \right)_{v=\bar{v}} &= 0 \\ \left(\frac{\partial^2 \Phi}{\partial v^2} \right)_{v=\bar{v}} &\leq 0 \end{aligned} \quad (7.92)$$

with the following additional rule: If (7.92) determines more than one value of \bar{v} , we must take only the value that gives the largest $\Phi(\bar{v}, z)$.

The first condition of (7.92) is the same as

$$\left(v^2 \frac{\partial F}{\partial v} \right)_{v=\bar{v}} = \log z \quad (7.93)$$

Substituting this into (7.86) we obtain

$$\beta P_{\text{can}}(\bar{v}) = F(\bar{v}) + \frac{1}{\bar{v}} \log z = \Phi(\bar{v}, z) \quad (7.94)$$

Comparing this with (7.91) we obtain

$$P_{\text{can}}(\bar{v}) = P_{\text{gr}}(\bar{v}) \quad (7.95)$$

That is, if there is a value \bar{v} that satisfies (7.92), then at this value of the specific volume the pressure is the same in the canonical and grand canonical ensemble. Therefore it only remains to investigate the possible values of \bar{v} .

It is obvious that \bar{v} can never lie between the values a and b shown in Fig. 7.6, because, as we can see from (7.88), in that region $\partial\Phi/\partial v = 0$ implies $\partial^2\Phi/\partial v^2 > 0$, in contradiction to (7.92). On the other hand, outside this region, $\partial\Phi/\partial v = 0$ implies $\partial^2\Phi/\partial v^2 \leq 0$. Hence the first condition of (7.92) alone determines \bar{v} . Using (7.85) we can write this condition in the form

$$\int_{v_0}^{\bar{v}} dv' P_{\text{can}}(v') - \bar{v} P_{\text{can}}(\bar{v}) = -kT \log z \quad (7.96)$$

There is a value of z , denoted by z_0 , at which (7.96) has two roots v_1 and v_2 for which $\Phi(v_1, z) = \Phi(v_2, z)$. The conditions for this to be so are that

$$-kT \log z_0 = \int_{v_0}^{v_1} dv' P_{\text{can}}(v') - v_1 P_{\text{can}}(v_1) = \int_{v_0}^{v_2} dv' P_{\text{can}}(v') - v_2 P_{\text{can}}(v_2) \quad (7.97)$$

$$\Phi(v_1, z_0) = \Phi(v_2, z_0)$$

The second condition is equivalent to $P_{\text{can}}(v_1) = P_{\text{can}}(v_2)$, by virtue of (7.94). Combining these conditions, we obtain

$$\int_{v_1}^{v_2} dv' P_{\text{can}}(v') = (v_2 - v_1) P_{\text{can}}(v_1) \quad (7.98)$$

which means that v_1 and v_2 are the end points of a Maxwell construction on P_{can} , as shown in Fig. 7.6.

In general we can find z as a function of \bar{v} by solving (7.96) graphically, in a manner similar to that used in the last section for (7.78). The result is qualitatively sketched in Fig. 7.7. As explained before, the interval $a < \bar{v} < b$ must be excluded. By definition of the Maxwell construction, the portions of the curves outside the interval $v_1 \leq \bar{v} \leq v_2$, shown in solid lines in Fig. 7.7, coincide with the corresponding portions in Fig. 7.3. We need to discuss further only the dashed portions of the curves.

Consider the points A and B in Fig. 7.7. Let their volumes be, respectively, v_A and v_B and let their common z value be z' . The fact that they are both solutions of (7.96) means that the function $\Phi(v, z')$ has two maxima, located respectively at $v = v_A$ and $v = v_B$. These maxima cannot be of the same height, because that would mean that v_A and v_B are, respectively, v_2 and v_1 , which they are not. To determine which maximum is higher we note that by (7.85), (7.94),

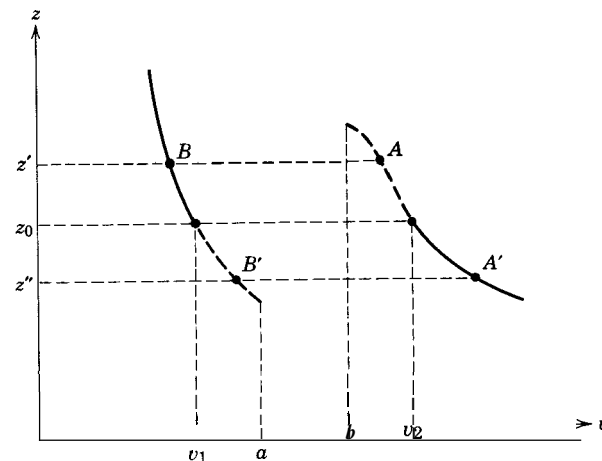


Fig. 7.7 z as a function of \bar{v} .

and the fact that z' is common to both,

$$\int_{v_B}^{v_A} dv' P_{\text{can}}(v') = v_A P_{\text{can}}(v_A) - v_B P_{\text{can}}(v_B) \quad (7.99)$$

Suppose $P_{\text{can}}(v_B) < P_{\text{can}}(v_A)$. Consider the point C indicated in Fig. 7.6. By inspection of Fig. 7.6 we see that

$$\int_{v_C}^{v_A} dv' P_{\text{can}}(v') < (v_A - v_C) P_{\text{can}}(v_A)$$

Subtracting (7.99) from this inequality, we obtain

$$\int_{v_C}^{v_B} dv' P_{\text{can}}(v') < v_B P_{\text{can}}(v_B) - v_C P_{\text{can}}(v_A)$$

which, by the original assumption, implies

$$\int_{v_C}^{v_B} dv' P_{\text{can}}(v') < (v_B - v_C) P_{\text{can}}(v_B)$$

By inspection of Fig. 7.6 we see that this is impossible. Therefore we must have $P_{\text{can}}(v_B) > P_{\text{can}}(v_A)$. By (7.94), this means that

$$\Phi(v_B, z') > \Phi(v_A, z')$$

In a similar fashion we can prove that, for the points A' and B' in Fig. 7.7,

$$\Phi(v_{A'}, z'') > \Phi(v_{B'}, z'')$$

Therefore the dashed portions of the curves in Fig. 7.7 must be discarded.

In Fig. 7.8, $P_{\text{gr}}(i)$ is shown as the solid curve. It is the same as $P_{\text{can}}(\bar{v})$ except that the portion between v_1 and v_2 is missing because there is no z that will give a \bar{v} lying in that interval. In other words, in the grand canonical ensemble the system cannot have a volume in that interval. We can, however, fill in a horizontal line at P_0 by the usual arguments, namely, that since the systems

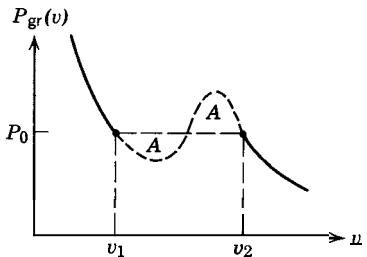


Fig. 7.8 The pressure in the grand canonical ensemble (solid lines).

at v_1 and v_2 have the same temperature, pressure, and chemical potential, a system at v_1 can coexist with a system at v_2 with any relative amount of each present.

It is an experimental fact that $\partial P/\partial v \leq 0$. It could not be otherwise, for then the system would be in the highly unstable situation in which releasing the pressure on it leads to a shrinkage. The quantity P_{can} is the result of a (generally approximate) calculation, and may or may not have this desirable property. However, the corresponding pressure in the grand canonical ensemble always satisfies the stability condition because the ensemble explicitly includes all possible density fluctuations of the system.

PROBLEMS

7.1 (a) Obtain the pressure of a classical ideal gas as a function of N , V , and T , by calculating the partition function.

(b) Obtain the same by calculating the grand partition function.

7.2 Consider a classical system of N noninteracting diatomic molecules enclosed in a box of volume V at temperature T . The Hamiltonian for a *single* molecule is taken to be

$$\mathcal{H}(\mathbf{p}_1, \mathbf{p}_2, \mathbf{r}_1, \mathbf{r}_2) = \frac{1}{2m}(p_1^2 + p_2^2) + \frac{1}{2}K|\mathbf{r}_1 - \mathbf{r}_2|^2$$

where $\mathbf{p}_1, \mathbf{p}_2, \mathbf{r}_1, \mathbf{r}_2$, are the momenta and coordinates of the two atoms in a molecule. Find

(a) the Helmholtz free energy of the system;

(b) the specific heat at constant volume;

(c) the mean square molecule diameter $\langle |\mathbf{r}_1 - \mathbf{r}_2|^2 \rangle$.

7.3 Repeat the last problem, using the Hamiltonian

$$\mathcal{H}(\mathbf{p}_1, \mathbf{p}_2, \mathbf{r}_1, \mathbf{r}_2) = \frac{1}{2m}(p_1^2 + p_2^2) + \epsilon|\mathbf{r}_{12} - r_0|$$

where ϵ and r_0 are given positive constants and $r_{12} \equiv |\mathbf{r}_1 - \mathbf{r}_2|$.

Answer.

$$\frac{C_V}{Nk} = 6 - \frac{x^2[2(x^2 - 2) + (x + 2)^2 e^{-x}]}{(x^2 + 2 - e^{-x})^2} \quad (x \equiv \epsilon r_0/kT)$$

7.4 Prove Van Leeuwen's Theorem: The phenomenon of diamagnetism does not exist in classical physics.

The following hints may be helpful:

(a) If $\mathcal{H}(\mathbf{p}_1, \dots, \mathbf{p}_N; \mathbf{q}_1, \dots, \mathbf{q}_N)$ is the Hamiltonian of a system of charged particles in the absence of an external magnetic field, then $\mathcal{H}[\mathbf{p}_1 - (e/c)\mathbf{A}_1, \dots, \mathbf{p}_N - (e/c)\mathbf{A}_N; \mathbf{q}_1, \dots, \mathbf{q}_N]$ is the Hamiltonian of the same system in the presence of an external magnetic field $\mathbf{H} = \nabla \times \mathbf{A}$, where \mathbf{A}_i is the value of \mathbf{A} at the position \mathbf{q}_i .

(b) The induced magnetization of the system along the direction of \mathbf{H} is given by

$$M = \left\langle -\frac{\partial \mathcal{H}}{\partial H} \right\rangle = kT \frac{\partial}{\partial H} \log Q_N$$

where \mathcal{H} is the Hamiltonian in the presence of \mathbf{H} , $H = |\mathbf{H}|$, and Q_N is the partition function of the system in the presence of \mathbf{H} .

7.5 Langevin's Theory of Paramagnetism. Consider a system of N atoms, each of which has an intrinsic magnetic moment of magnitude μ . The Hamiltonian in the presence of an external magnetic field \mathbf{H} is

$$\mathcal{H}(p, q) - \mu H \sum_{i=1}^N \cos \alpha_i$$

where $\mathcal{H}(p, q)$ is the Hamiltonian of the system in the absence of an external magnetic field, and α_i is the angle between \mathbf{H} and the magnetic moment of the i th atom. Show that

(a) The induced magnetic moment is

$$M = N\mu \left(\coth \theta - \frac{1}{\theta} \right) \quad (\theta \equiv \mu H/kT)$$

(b) The magnetic susceptibility per atom is

$$\chi = \frac{\mu^2}{kT} \left(\frac{1}{\theta^2} - \text{csch}^2 \theta \right)$$

(c) At high temperatures χ satisfies Curie's law, namely χ a T^{-1} . Find the proportionality constant, which is called Curie's constant.

7.6 Imperfect Gas. Consider a system of N molecules ($N \rightarrow \infty$) contained in a box of volume V ($V \rightarrow m$). The Hamiltonian of the system is

$$\mathcal{H} = \sum_{i=1}^N \frac{p_i^2}{2m} + \sum_{i < j} v_{ij}$$

$$v_{ij} = v(|\mathbf{r}_i - \mathbf{r}_j|)$$

where \mathbf{p} and \mathbf{r} are, respectively, the momentum and the position of the i th molecule. The intermolecular potential $v(r)$ has the qualitative form shown in the accompanying figure.

Let

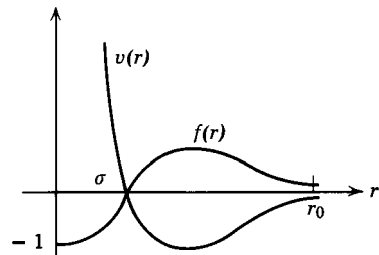
$$f_{ij} \equiv f(|\mathbf{r}_i - \mathbf{r}_j|)$$

$$f(r) \equiv e^{-\beta v(r)} - 1$$

A sketch of $f(r)$ is also shown in the same figure.

(a) Show that the equation of state of the system is

$$\frac{Pv}{kT} = 1 + v \frac{\partial Z(v, T)}{\partial v}$$



where $v \equiv V/N$ and

$$Z(v, T) \equiv \frac{1}{N} \log \left[\frac{1}{V^N} \int d^3 r_1 \cdots d^3 r_N \prod_{i < j} (1 + f_{ij}) \right]$$

(b) By expanding the product $\prod (1 + f_{ij})$, show that

$$\begin{aligned} Z(v, T) &= \frac{1}{N} \log \left[\frac{1}{V^N} \int d^3 r_1 \cdots d^3 r_N \left(1 + \sum_{i < j} f_{ij} + \cdots \right) \right] \\ &= \log \left[1 + \frac{N}{2V} \int d^3 r f(r) + \cdots \right]^{1/N} \end{aligned}$$

(c) Show that at low densities, i.e.,

$$r_0^3/v \ll 1$$

it is a good approximation to retain only the first two terms in the series appearing in the expression $Z(v, T)$. Hence the equation of state is approximately given by

$$\frac{Pv}{kT} \approx 1 - \frac{1}{2v} \int_0^\infty dr 4\pi r^2 f(r)$$

The coefficient of $1/v$ is called the *second virial coefficient*.

Note. (i) Retaining the first two terms in the series appearing in $Z(v, T)$ is a good approximation because $Z(v, T)$ is the logarithm of the N th root of the series. The approximation is certainly invalid for the series itself

(ii) If all terms in the expansion of $\prod (1 + f_{ij})$ were kept, we would have obtained a systematic expansion of Pv/kT in powers of $1/v$. Such an expansion is known as the *virial expansion*.

(iii) The complete virial expansion is difficult to obtain by the method described in this problem. It is obtained in Chapter 10 via the grand canonical ensemble. See (10.27) and (10.30).

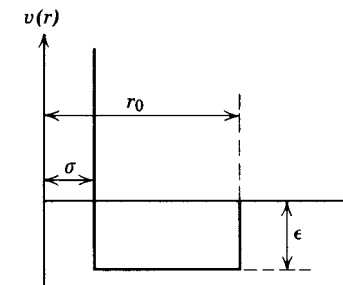
7.7 Van der Waals Equation of State

(a) Show that for low densities the Van der Waals equation of state (2.28) reduces to

$$\frac{Pv}{kT} \approx 1 + \frac{1}{v} \left(b' - \frac{a'}{kT} \right)$$

(b) Show that the imperfect gas of Problem 7.6 has an equation of state of the same form as shown in (a), with

$$\begin{aligned} b' &= \frac{2\pi}{3} \sigma^3 \\ a' &= -2\pi kT \int_0^\infty dr r^2 (1 - e^{-\beta v(r)}) \end{aligned}$$



7.8 The equation of state for an N_2 gas can be written in the form

$$PV/NkT = 1 + a_2(T)(N/V)$$

for low densities. The second virial coefficient $a_2(T)$ has been measured as a function of temperature and is given in the accompanying table. Assume that the intermolecular potential $v(r)$ between N_2 molecules has the form shown in the accompanying sketch. From the data given, determine what you consider to be the best choice for the constants a , r_0 , and ϵ .

Temperature, K	$a_2(T)$, K/atm
100	-1.80
200	-4.26×10^{-1}
300	-5.49×10^{-2}
400	1.12×10^{-1}
500	2.05×10^{-1}

7.9 A dilute mixture of H_2 and O_2 gases is kept at constant temperature T . Initially the density of H_2 was n_0 , the density of O_2 was $n_0/2$, and there was no H_2O present. After a certain time, the mixture becomes an equilibrium mixture of H_2 , O_2 , and H_2O . Find the equilibrium densities of the three components n_1, n_2, n_3 , as a function of T and n_0 .

CHAPTER

8

QUANTUM STATISTICAL
MECHANICS8.1 THE POSTULATES OF QUANTUM
STATISTICAL MECHANICS

All systems in nature obey quantum mechanics. In quantum mechanics an observable of a system is associated with a Hermitian operator, which operates on a Hilbert space. A state of the system is a vector $|\Psi\rangle$ in the same Hilbert space. If $|q\rangle$ is an eigenvector of the position operators of all the particles in the system, then $\langle q|\Psi\rangle \equiv \Psi(q)$ is the wave function of the system in the state $|\Psi\rangle$. The wave function furnishes a complete description of the state.

At any instant of time the wave function Ψ of a truly isolated system may be expressed as a linear superposition of a complete orthonormal set of stationary wave functions $\{\Phi_n\}$:

$$\Psi = \sum_n c_n \Phi_n \quad (8.1)$$

where c_n is a complex number and is a function of time. The index n stands for a set of quantum numbers, which are eigenvalues of certain chosen dynamical operators of the system. The square modulus $|c_n|^2$ is the probability that a measurement performed on the system will find it to have the quantum numbers n .

In statistical mechanics we always deal with systems that interact with the external world. Here we can regard the system plus the external world as a truly isolated system. The wave function Ψ for this whole system will depend on both the coordinates of the system under consideration and the coordinates of the external world. If $\{\Phi_n\}$ denotes a complete set of orthonormal stationary wave functions of the system, then Ψ is still formally given by (8.1), but c_n is to be interpreted as a wave function of the external world. It depends on the coordinates of the external world as well as on the time.

Suppose O is an operator corresponding to an observable of the system. According to the rules of quantum mechanics, the average result of a large number of measurements of this observable is instantaneously given by the expectation value

$$\frac{(\Psi, O\Psi)}{(\Psi, \Psi)} = \frac{\sum_n \sum_m (c_n, c_m) (\Phi_n, O\Phi_m)}{\sum_n (c_n, c_n)} \quad (8.2)$$

where (c_n, c_m) , the scalar product of the n th and the m th wave function of the external world, is a function of time. The denominator of (8.2), being identical with (Ψ, Ψ) , is independent of time, because the Hamiltonian of the system plus external world is Hermitian. When we actually measure an observable in the laboratory, we measure not its instantaneous value but a time average. Thus the directly measurable quantity is not (8.2) but the following quantity:

$$\langle O \rangle \equiv \frac{\overline{(\Psi, O\Psi)}}{(\Psi, \Psi)} = \frac{\sum_n \sum_m \overline{(c_n, c_m)} (\Phi_n, O\Phi_m)}{\sum_n \overline{(c_n, c_n)}} \quad (8.3)$$

where $\overline{(c_n, c_m)}$ is the average of (c_n, c_m) over a time interval that is short compared to the resolving time of the measuring apparatus but long compared to molecular times (e.g., collision times or periods of molecular motion). We note that $\sum_n \overline{(c_n, c_n)}$ is identical with $\sum_n (c_n, c_n)$, because the latter is independent of time.

The postulates of quantum statistical mechanics are postulates concerning the coefficients (c_n, c_m) , when (8.3) refers to a macroscopic observable of a macroscopic system in thermodynamic equilibrium.

For definiteness, we consider a macroscopic system which, although not truly isolated, interacts so weakly with the external world that its energy is approximately constant. Let the number of particles in the system be N and the volume of the system be V , and let its energy lie between E and $E + \Delta$ ($\Delta \ll E$). Let A be the Hamiltonian of the system. For such a system it is convenient (but not necessary) to choose a standard set of complete orthonormal wave functions $\{\Phi_n\}$ such that every Φ_n is a wave function for N particles contained in the volume V and is an eigenfunction of A with the eigenvalue E_n :

$$\mathcal{H}\Phi_n = E_n\Phi_n \quad (8.4)$$

The postulates of quantum statistical mechanics are the following:

Postulate of Equal a Priori Probability

$$\overline{(c_n, c_n)} = \begin{cases} 1 & (E < E_n < E + \Delta) \\ 0 & (\text{otherwise}) \end{cases} \quad (8.5)$$

Postulate of Random Phases

$$\overline{(c_n, c_m)} = 0 \quad (n \neq m) \quad (8.6)$$

As a consequence of these postulates we may *effectively* regard the wave function of the system as given by

$$\Psi = \sum_n b_n \Phi_n \quad (8.7)$$

where

$$|b_n|^2 = \begin{cases} 1 & (E < E_n < E + \Delta) \\ 0 & (\text{otherwise}) \end{cases} \quad (8.8)$$

and where the phases of the complex numbers $\{b_n\}$ are random numbers. In this manner the effect of the external world is taken into account in an average way. The observed value of an observable associated with the operator \mathcal{O} is then given by

$$\langle \mathcal{O} \rangle = \frac{\sum_n |b_n|^2 (\Phi_n, \mathcal{O} \Phi_n)}{\sum_n |b_n|^2} \quad (8.9)$$

It should be emphasized that for (8.7) and (8.8) to be valid the system must interact with the external world. Otherwise the postulate of random phases is false. By the **randomness** of the phases we mean no more and no less than the absence of interference of probability amplitudes, as expressed by (8.9). For a truly isolated system such a circumstance may be true at an instant, but it cannot be true for all times.

The postulate of random phases implies that the state of a system in equilibrium may be regarded as an incoherent superposition of eigenstates. It is possible to think of the system as one member of an infinite collection of systems, each of which is in an eigenstate whose wave function is Φ_n . Since these systems do not interfere with one another, it is possible to form a mental picture of each system one at a time. This mental picture is the quantum mechanical generalization of the Gibbsian ensemble. The ensemble defined by the previous postulates is the microcanonical ensemble.

The postulates of quantum statistical mechanics are to be regarded as working hypotheses whose justification lies in the fact that they lead to results in agreement with **experiments**. Such a point of view is not entirely satisfactory, because these postulates cannot be independent of, and should be derivable from, the quantum mechanics of molecular systems. A rigorous derivation is at present lacking. We return to this subject very briefly at the end of this chapter.

We should recognize that the postulates of quantum statistical mechanics, even if regarded as phenomenological statements, are more fundamental than the laws of **thermodynamics**. The reason is twofold. First, the postulates of quantum statistical mechanics not only imply the laws of **thermodynamics**, they also lead to definite formulas for all the **thermodynamic** functions of a given substance.

Second, they are more directly related to molecular dynamics than are the laws of **thermodynamics**.

The concept of an ensemble is a familiar one in quantum mechanics. A trivial example is the description of an incident **beam** of particles in the theory of scattering. The incident beam of particles in a scattering experiment is composed of many particles, but in the theory of scattering we consider the particles one at a time. That is, we calculate the scattering cross section for a single incident particle and then add the cross sections for all the particles to obtain the physical cross section. Inherent in this method is the assumption that the wave functions of the particles in the incident beam bear no definite phase with respect to one another. What we have described is in fact an ensemble of particles.

A less trivial example is the description of a beam of incident electrons whose spin can be polarized. If an electron has the wave function

$$\left[A \begin{pmatrix} 1 \\ 0 \end{pmatrix} + B \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right] e^{i\mathbf{k} \cdot \mathbf{r}}$$

where A and B are definite complex numbers, the electron has a spin pointing in some definite direction. This corresponds to an incident beam of completely polarized electrons. In the cross section calculated with this wave function there will appear interference terms proportional to $A^*B + AB^*$. If we have an incident beam that is partially polarized, we first calculate the cross section with a wave function proportional to $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and then do the **same** thing for $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, adding the two cross sections with appropriate weighting factors. This is equivalent to describing the incident beam by an ensemble of electrons in which the states $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ occur with certain relative probabilities.

8.2 DENSITY MATRIX

An ensemble is an incoherent superposition of states. Its relevance to physics has been postulated in the previous section. We note that only the square moduli $|b_n|^2$ appear in (8.9). Hence it should be possible to describe an ensemble in such a way that the random phases of the states never need to be mentioned. Such a goal is achieved by introducing the density matrix.

Before we define the density matrix let us note that an operator is defined when all its matrix elements with respect to a complete set of states are defined. Its matrix elements with respect to any other complete set of states can be found by the well-known rules of transformation theory in quantum mechanics. Therefore, if all the matrix elements of an operator are defined in one representation, the operator is thereby defined in any representation.

We define the density matrix ρ_{mn} corresponding to a given ensemble by

$$\rho_{mn} \equiv (\Phi_n, \rho \Phi_m) \equiv \delta_{mn} |b_n|^2 \quad (8.10)$$

where Φ_n and b_n have the same meaning as in (8.7). In this particular representation ρ_{mn} is a diagonal matrix, but in some other representation it need not be. Equation (8.10) also defines the density operator ρ whose matrix elements are ρ_{mn} . The operator ρ operates on state vectors in the Hilbert space of the system under consideration.

In terms of the density matrix, (8.9) can be rewritten in the form

$$\langle \mathcal{O} \rangle = \frac{\sum_n (\Phi_n, \mathcal{O} \rho \Phi_n)}{\sum_n (\Phi_n, \rho \Phi_n)} = \frac{\text{Tr}(\mathcal{O} \rho)}{\text{Tr} \rho} \quad (8.11)$$

The notation $\text{Tr} A$ denotes the trace of the operator A and is the sum of all the diagonal matrix elements of A in any representation. An elementary property of the trace is that

$$\text{Tr}(AB) = \text{Tr}(BA)$$

It follows immediately that $\text{Tr} A$ is independent of the representation; if $\text{Tr} A$ is calculated in one representation, its value in another representation is

$$\text{Tr}(SAS^{-1}) = \text{Tr}(S^{-1}SA) = \text{Tr} A$$

The introduction of the density matrix merely introduces a notation. It does not introduce new physical content. The usefulness of the density matrix lies solely in the fact that with its help (8.11) is presented in a form that is manifestly independent of the choice of the basis $\{\Phi_n\}$, although this independence is a property that this expectation value always possesses.

The density operator ρ defined by (8.10) contains all the information about an ensemble. It is independent of time if it commutes with the Hamiltonian of the system and if the Hamiltonian is independent of time. This statement is an immediate consequence of the equation of motion of ρ :

$$i\hbar \frac{\partial \rho}{\partial t} = [\mathcal{H}, \rho] \quad (8.12)$$

which is the quantum mechanical version of Liouville's theorem.

Formally we can represent the density operator ρ as

$$\rho = \sum_n |\Phi_n\rangle |b_n|^2 \langle \Phi_n| \quad (8.13)$$

where $|\Phi_n\rangle$ is the state vector whose wave function is Φ_n . To prove (8.13), we verify that it has the matrix elements (8.10):

$$\rho_{mn} \equiv (\Phi_m, \rho \Phi_n) \equiv \langle \Phi_m | \rho | \Phi_n \rangle = \sum_k \langle \Phi_m | \Phi_k \rangle |b_k|^2 \langle \Phi_k | \Phi_n \rangle = \delta_{mn} |b_n|^2 \quad \blacksquare$$

The time-averaging process through which we averaged out the effect of the external world on the system under consideration may be reformulated in terms of the density matrix.

Formula (8.2) is a general formula for the expectation value of any operator \mathcal{O} with respect to an arbitrary wave function Ψ . It may be trivially rewritten in the form

$$\frac{(\Psi, \mathcal{O}, \Psi)}{(\Psi, \Psi)} = \frac{\sum_n \sum_m R_{mn} \mathcal{O}_{nm}}{\sum_n R_{nn}} = \frac{\text{Tr}(R\mathcal{O})}{\text{Tr} R}$$

where $R_{nm} \equiv (c_m, c_n) \equiv (\Phi_n, R\Phi_m)$, the last identity being a definition of the operator R , and $\mathcal{O}_{nm} \equiv (\Phi_n, \mathcal{O}\Phi_m)$. Although R may depend on the time, $\text{Tr} R$ is independent of time. The density operator is the time average of R :

$$\rho \equiv \bar{R}$$

8.3 ENSEMBLES IN QUANTUM STATISTICAL MECHANICS

Microcanonical Ensemble

The density matrix for the microcanonical ensemble in the representation in which the Hamiltonian is diagonal is

$$\rho_{mn} = \delta_{mn} |b_n|^2 \quad (8.14)$$

where

$$|b_n|^2 = \begin{cases} \text{Const.} & (E < E_n < E + \Delta) \\ 0 & (\text{otherwise}) \end{cases} \quad (8.15)$$

where $\{E_n\}$ are the eigenvalues of the Hamiltonian. The density operator is

$$\rho = \sum_{E < E_n < E + \Delta} |\Phi_n\rangle \langle \Phi_n| \quad (8.16)$$

The trace of ρ is equal to the number of states whose energy lies between E and $E + \Delta$:

$$\text{Tr} \rho = \sum_n \rho_{nn} = \Gamma(E) \quad (8.17)$$

For macroscopic systems the spectrum $\{E_n\}$ almost forms a continuum. For $\Delta \ll E$, we may take

$$\Gamma(E) = \omega(E) \Delta \quad (8.18)$$

where $\omega(E)$ is the density of states at energy E . The connection between the microcanonical ensemble and thermodynamics is established by identifying the entropy as

$$S(E, V) = k \log \Gamma(E) \quad (8.19)$$

where k is Boltzmann's constant. This definition is the same as in classical statistical mechanics, except that $\Gamma(E)$ must be calculated in quantum mechanics. From this point on all further developments become exactly the same as in classical statistical mechanics and so they need not be repeated. No Gibbs

paradox will result from (8.19) because the correct counting of states is automatically implied by the definition of $\Gamma(E)$ in (8.17).

The only new result following from (8.19) that is not obtainable in classical statistical mechanics is the third law of thermodynamics, which we discuss separately in Section 8.4.

Canonical Ensemble

The derivation of the canonical ensemble from the microcanonical ensemble given in Chapter 8 did not make essential use of classical mechanics. That derivation continues to be valid in quantum statistical mechanics, with the trivial change that the integration over Γ space is replaced by a sum over all the states of the system:

$$\frac{1}{N!h^{3N}} \int dp dq \rightarrow \sum_n \quad (8.20)$$

Thus the canonical ensemble is defined by the density matrix

$$\rho_{mn} = \delta_{mn} e^{-\beta E_n} \quad (8.21)$$

where $\beta = 1/kT$. This result states that at the temperature T the relative probability for the system to have the energy eigenvalue E_n is $e^{-\beta E_n}$, which is called the **Boltzmann factor**. The partition function is given by

$$Q_N(V, T) = \text{Tr} \rho = \sum_n e^{-\beta E_n} \quad (8.22)$$

where it must be emphasized that *the sum on the right side is a sum over states and not over energy eigenvalues*. The connection with thermodynamics is the same as in classical statistical mechanics.

The density operator ρ is

$$\rho = \sum_n |\Phi_n\rangle e^{-\beta E_n} \langle \Phi_n| = e^{-\beta \mathcal{H}} \sum_n |\Phi_n\rangle \langle \Phi_n|$$

where \mathcal{H} is the Hamiltonian operator. Now the operator $\sum_n |\Phi_n\rangle \langle \Phi_n|$ is the identity operator, by the completeness property of eigenstates. Therefore

$$\rho = e^{-\beta \mathcal{H}} \quad (8.23)$$

The partition function can be written in the form

$$Q_N(V, T) = \text{Tr} e^{-\beta \mathcal{H}} \quad (8.24)$$

where the trace is to be taken over all states of the system that has N particles in the volume V . This form, which is explicitly independent of the representation, is sometimes convenient for calculations. The ensemble average of \mathcal{O} in the canonical ensemble is

$$\langle \mathcal{O} \rangle = \frac{\text{Tr} (\mathcal{O} e^{-\beta \mathcal{H}})}{Q_N} \quad (8.25)$$

Grand Canonical Ensemble

For the grand canonical ensemble the density operator ρ operates on a Hilbert space with an indefinite number of particles. We do not display it because we do not need it. It is sufficient to state that the grand partition function is

$$\mathcal{Q}(z, V, T) = \sum_{N=0}^{\infty} z^N Q_N(V, T) \quad (8.26)$$

where Q_N is the partition function for N particles. The connection between $\log \mathcal{Q}$ and thermodynamics is the same as in classical statistical mechanics. The ensemble average of \mathcal{O} in the grand canonical ensemble is

$$\langle \mathcal{O} \rangle = \frac{1}{\mathcal{Q}} \sum_{N=0}^{\infty} z^N \langle \mathcal{O} \rangle_N \quad (8.27)$$

where $\langle \mathcal{O} \rangle_N$ is the ensemble average (8.25) in the canonical ensemble for N particles. These equations can be written more generally in the forms

$$\begin{aligned} \mathcal{Q}(z, V, T) &= \text{Tr} e^{-\beta(\mathcal{H} - \mu N)} \\ \langle \mathcal{O} \rangle &= \frac{1}{\mathcal{Q}} \text{Tr} [\mathcal{O} e^{-\beta(\mathcal{H} - \mu N)}] \end{aligned} \quad (8.28)$$

where N is an operator representing a conserved quantity (i.e., one that commutes with the Hamiltonian), and the trace is taken over all states without restriction on the eigenvalues of N . The only restrictions on the trace are boundary conditions, which specify the volume containing the system, and the symmetry property of the states under the interchange of identical particles.

8.4 THIRD LAW OF THERMODYNAMICS

The definition of entropy is given by (8.19). At the absolute zero of temperature a system is in its ground state, i.e., a state of lowest energy. For a system whose energy eigenvalues are **discrete**, (8.19) implies that at absolute zero $S = k \log G$, where G is the degeneracy of the ground state. If the ground state is unique, then $S = 0$ at absolute zero. If the ground state is not unique, but $G \leq N$, where N is the total number of molecules in the system, then at absolute zero $S \leq k \log N$. In both of these cases the third law of thermodynamics holds, because the entropy per molecule at absolute zero is of order $(\log N)/N$.

The energy eigenvalues for most macroscopic systems, however, essentially form a continuous spectrum. For these systems the previous argument only shows that the entropy per molecule approaches zero when the temperature T is so low that

$$kT \ll AE$$

where AE is the energy difference between the first excited state and the ground

state. As an estimate let us put

$$\Delta E \approx \frac{\hbar^2}{mV^{2/3}}$$

where m is the mass of a nucleon, $V = 1 \text{ cm}^3$. Then we find that $T \approx 5 \times 10^{-15} \text{ K}$. Clearly this phenomenon has nothing to do with the third law of thermodynamics, which is a phenomenological statement based on experiments performed above 1 K.

To verify the third law of thermodynamics for systems having an almost continuous energy spectrum we must study the behavior of the density of states $\omega(E)$ near $E = 0$. Most of the substances known to us become crystalline solids near absolute zero. For these substances all thermodynamic functions near absolute zero may be obtained through Debye's theory, which is discussed in Section 12.2. It is shown there that the third law of thermodynamics is fulfilled.

The only known substance that remains a liquid at absolute zero is helium, which is discussed in Chapter 13. There it is shown that near absolute zero the density of states for liquid helium is qualitatively the same as that for a crystalline solid. Therefore the third law of thermodynamics is also fulfilled for liquid helium.

Apart from these specific examples, which include all known substances, we cannot give a more universal proof of the third law of thermodynamics. But this is perhaps sufficient; after all, the third law of thermodynamics is a summary of empirical data gathered from known substances.

8.5 THE IDEAL GASES: MICROCANONICAL ENSEMBLE

The simplest system of N identical particles is that composed of N noninteracting members. The Hamiltonian is

$$\mathcal{H} = \sum_{i=1}^N \frac{p_i^2}{2m} \tag{8.29}$$

where $p_i^2 = \mathbf{p}_i \cdot \mathbf{p}_i$, and \mathbf{p}_i is the momentum operator of the i th particle. The Hamiltonian is independent of the positions of the particles or any other coordinates, e.g., spin, if any.

In nature a system of N identical particles is one of two types: A *Bose system* or a *Fermi system*.* A complete set of eigenfunctions for a Bose system is the collection of those eigenfunctions of \mathcal{H} that are symmetric under an interchange of any pair of particle coordinates. A complete set of eigenfunctions for a Fermi system is the collection of those eigenfunctions of \mathcal{H} that are antisymmetric under an interchange of any pair of particle coordinates. Particles forming a Bose system are called *bosons*, and particles forming a Fermi system are called *fermions*.

*See the Appendix, Section A.1

In addition to these two types of systems we define, for mathematical comparison, what is called a Boltzmann system. It is defined as a system of particles whose eigenfunctions are *all* the eigenfunctions of \mathcal{H} ; but the rule for counting these eigenfunctions shall be the "correct Boltzmann counting." The set of eigenfunctions for a Boltzmann system includes those for a Bose system, those for a Fermi system, and more. There is no known system of this type in nature. It is a useful model, however, because at high temperatures the thermodynamic behavior of both the Bose system and the Fermi system approaches that of the Boltzmann system.

For noninteracting identical particles we have three cases: The ideal Bose gas, the ideal Fermi gas, and the ideal Boltzmann gas. We first work out the thermodynamics of these ideal gases in the formalism of the microcanonical ensemble. For this purpose it is necessary to find out, for each of the three cases, the number of states $\Gamma(E)$ of the system having an energy eigenvalue that lies between E and $E + \Delta$. That is, we must learn how to count.

To avoid all unnecessary complications we confine our discussion to spinless particles. Any energy eigenvalue of an ideal system is a sum of single-particle energies, called *levels*. These are given by

$$\epsilon_{\mathbf{p}} = \frac{p^2}{2m} \tag{8.30}$$

where $p \equiv |\mathbf{p}|$ and \mathbf{p} is the momentum eigenvalue of the single particle:

$$\mathbf{p} = \frac{2\pi\hbar}{L} \mathbf{n} \tag{8.31}$$

in which \mathbf{n} is a vector whose components are 0 or \pm integers and L is the cube root of the volume of the system:

$$L \equiv V^{1/3}$$

In the limit as $V \rightarrow \infty$ the possible values of \mathbf{p} form a continuum. Then a sum over \mathbf{p} can sometimes be replaced by an integration

$$\sum_{\mathbf{p}} \rightarrow \frac{V}{h^3} \int d^3p \tag{8.32}$$

where $h = 2\pi\hbar$ is Planck's constant.*

A state of an ideal system can be specified by specifying a set of occupation numbers $\{n_{\mathbf{p}}\}$ so defined that there are $n_{\mathbf{p}}$ particles having the momentum \mathbf{p} in the state under consideration. Obviously the total energy E and the total number of particles N of the state are given by

$$\begin{aligned} E &= \sum_{\mathbf{p}} \epsilon_{\mathbf{p}} n_{\mathbf{p}} \\ N &= \sum_{\mathbf{p}} n_{\mathbf{p}} \end{aligned} \tag{8.33}$$

*For an explanation of (8.31) and (8.32), see the Appendix, Section A.2.

For spinless bosons and fermions $\{n_p\}$ uniquely defines a state of the system. The allowed values for any n_p are

$$n_p = \begin{cases} 2, \dots & \text{(for bosons)} \\ 0, 1, 2, \dots & \text{(for fermions)} \end{cases} \quad (8.34)$$

For a Boltzmann gas $n_p = 0, 1, 2, \dots$, but $\{n_p\}$ specifies $N! / \prod_p (n_p!)$ states of the N-particle system. This is because an interchange of the momenta of two particles in the system in general leads to a new state but leaves $\{n_p\}$ unchanged.

The total energy is a given number E to within a small uncertainty A, whose value is unimportant. Hence $\Gamma(E)$ may be found as follows. As $V \rightarrow \infty$, the levels (8.30) form a continuum. Let us divide the spectrum of (8.30) into groups of levels containing respectively g_1, g_2, \dots levels, as shown in Fig. 8.1. Each group is called a cell and has an average energy ϵ_i . The occupation number of the ith cell, denoted by n_i , is the sum of n_p over all the levels in the ith cell. Each g_i is assumed to be very large, but its exact value is unimportant. Let

$$W\{n_i\} \equiv \text{no. of states of the system corresponding to the set of occupation numbers } \{n_i\} \quad (8.35)$$

Then

$$\Gamma(E) = \sum_{\{n_i\}} W\{n_i\} \quad (8.36)$$

where the sum extends over all sets of integers $\{n_i\}$ satisfying the conditions

$$E = \sum_i \epsilon_i n_i \quad (8.37)$$

$$N = \sum_i n_i \quad (8.38)$$

To find $W\{n_i\}$ for a Bose gas and a Fermi gas it is sufficient to find w_i , the number of ways in which n_i particles can be assigned to the ith cell (which

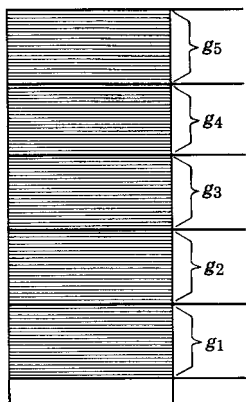
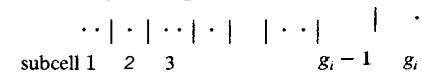


Fig. 8.1 Division of the single-particle energy spectrum into cells.

contains g_i levels). Since interchanging particles in different cells does not lead to a new state of the system, we have $W\{n_i\} = \prod_j w_j$. For a Boltzmann gas interchanging particles in different cells leads to a new state of the system, and we consider all N particles together. The three cases are worked out as follows.

Bose Gas. Each level can be occupied by any number of particles. Picture the ith cell to have g_i subcells, with $g_i - 1$ partitions, as follows:



The number w_i is the number of permutations of the n_i particles plus the $g_i - 1$ partitions that give rise to distinct arrangements:

$$w_i = \frac{(n_i + g_i - 1)!}{n_i! (g_i - 1)!}$$

Hence

$$W\{n_i\} = \prod_i w_i = \prod_i \frac{(n_i + g_i - 1)!}{n_i! (g_i - 1)!} \quad \text{(Bose)} \quad (8.39)$$

Fermi Gas. The number of particles in each of the g_i subcells of the ith cell is either 0 or 1. Therefore w_i is equal to the number of ways in which n_i things can be chosen from g_i things:

$$w_i = \binom{g_i}{n_i} = \frac{g_i!}{n_i! (g_i - n_i)!} \quad (8.40)$$

Hence

$$W\{n_i\} = \prod_i w_i = \prod_i \frac{g_i!}{n_i! (g_i - n_i)!} \quad \text{(Fermi)} \quad (8.41)$$

Boltzmann Gas. The N particles are first placed into cells, the ith cell having n_i particles. There are $N! / \prod_i (n_i!)$ ways to do this. Within the ith cell there are g_i levels. Among the n_i particles in the ith cells, the first one can occupy these levels g_i ways. The second and all subsequent ones also can occupy the levels g_i ways. Therefore there are $(g_i)^{n_i}$ ways in which n_i particles can occupy the g_i levels. The total number of ways to obtain $\{n_i\}$ is therefore

$$N! \prod_i \frac{g_i^{n_i}}{n_i!}$$

However, $W\{n_i\}$ is defined to be $1/N!$ of the last quantity:

$$W\{n_i\} = \prod_i \frac{g_i^{n_i}}{n_i!} \quad \text{(Boltzmann)} \quad (8.42)$$

This definition corresponds to the rule of "correct Boltzmann counting" and does

not correspond to any physical property of the particles in the system. It is just a rule that defines the mathematical model.

The fact that the rule for the counting of states is different for the three cases gives rise to the terminology Bose statistics, Fermi statistics, and Boltzmann statistics, which refer to the three rules of counting, respectively.

To obtain the entropy $S = k \log \Gamma(E)$ we need to sum $W\{n_i\}$ over $\{n_i\}$ in accordance with (8.35). This is a formidable task. For the Boltzmann gas it was explicitly done in Section 6.5. As we might correctly guess, however, $\Gamma(E)$ is quite well approximated by $W\{\bar{n}_i\}$, where $\{\bar{n}_i\}$ is the set of occupation numbers that maximizes $W\{\bar{n}_i\}$ subject to (8.37) and (8.38). We adopt this approximation and verify its correctness by showing that the fluctuations are small. Accordingly the entropy is taken to be

$$S = k \log W\{\bar{n}_i\} \quad (8.43)$$

To find $\{\bar{n}_i\}$ we maximize $W\{n_i\}$ by varying the n_i subject to (8.37) and (8.38). The details of this calculation are similar to that in Section 4.3 and will not be reproduced. We merely give the answers:

$$\bar{n}_i = \begin{cases} \frac{g_i}{z^{-1} e^{\beta \epsilon_i} \mp 1} & \text{(Bose and Fermi)} \\ g_i z e^{-\beta \epsilon_i} & \text{(Boltzmann)} \end{cases} \quad (8.44)$$

We deduce from this that

$$\bar{n}_{\mathbf{p}} = \begin{cases} \frac{1}{z^{-1} e^{\beta \epsilon_{\mathbf{p}}} \mp 1} & \text{(Bose and Fermi)} \\ z e^{-\beta \epsilon_{\mathbf{p}}} & \text{(Boltzmann)} \end{cases} \quad (8.45)$$

The parameters z and β are two Lagrange multipliers to be determined from the conditions

$$\begin{aligned} \sum_{\mathbf{p}} \epsilon_{\mathbf{p}} \bar{n}_{\mathbf{p}} &= E \\ \sum_{\mathbf{p}} \bar{n}_{\mathbf{p}} &= N \end{aligned} \quad (8.46)$$

The first of these leads to the identification $\beta = 1/kT$, and the second identifies z as the fugacity.

Using Stirling's approximation and neglecting 1 compared to g_i we have from (8.43) and (8.44),

$$\frac{S}{k} = \log W\{\bar{n}_i\} = \begin{cases} \left[\sum_i \left[\bar{n}_i \log \left(1 + \frac{g_i}{\bar{n}_i} \right) + g_i \log \left(1 + \frac{\bar{n}_i}{g_i} \right) \right] \right. & \text{(Bose)} \\ \left. \sum_i \left[\bar{n}_i \log \left(\frac{g_i}{\bar{n}_i} - 1 \right) - g_i \log \left(1 - \frac{\bar{n}_i}{g_i} \right) \right] \right. & \text{(Fermi)} \\ \left. \sum_i \bar{n}_i \log (g_i / \bar{n}_i) \right. & \text{(Boltzmann)} \end{cases} \quad (8.47)$$

More explicitly,

$$\begin{aligned} S &= \left(\sum_i g_i \left[\frac{\beta \epsilon_i - \log z}{z^{-1} e^{\beta \epsilon_i} - 1} - \log (1 - z e^{-\beta \epsilon_i}) \right] \right) & \text{(Bose)} \\ &= \left(\int g_i \left[\frac{\beta \epsilon_i - \log z}{z^{-1} e^{\beta \epsilon_i} + 1} + \log (1 + z e^{-\beta \epsilon_i}) \right] \right) & \text{(Fermi)} \\ &= \left(z \sum_i g_i e^{-\beta \epsilon_i} (\beta \epsilon_i - \log z) \right) & \text{(Boltzmann)} \end{aligned} \quad (8.48)$$

The validity of these equations depends on the assumption that

$$\overline{n_i^2} - \bar{n}_i^2 \ll \bar{n}_i^2 \quad (8.49)$$

This is best discussed in the grand canonical ensemble (see Problem 8.4). From (8.48) all other thermodynamic functions can be determined after z is determined in terms of N from (8.46).

The Boltzmann gas will be worked out explicitly. From (8.38) and (8.44) we have

$$N = z \sum_i g_i e^{-\beta \epsilon_i} = z \sum_{\mathbf{p}} e^{-\beta \epsilon_{\mathbf{p}}} = \frac{zV}{h^3} \int_0^\infty dp \, 4\pi p^2 e^{-\beta p^2/2m} = \frac{zV}{\lambda^3} \quad (8.50)$$

where

$$\lambda = \sqrt{\frac{2\pi \hbar^2}{mkT}} \quad (8.51)$$

This quantity is called the thermal wavelength because it is of the order of the de Broglie wavelength of a particle of mass m with the energy kT . Writing $v = V/N$ we obtain

$$z = \frac{\lambda^3}{v} \quad (8.52)$$

The condition $E = \sum n_i \epsilon_i$ requires that

$$E = z \sum_i g_i \epsilon_i e^{-\beta \epsilon_i} = z \sum_{\mathbf{p}} \epsilon_{\mathbf{p}} e^{-\beta \epsilon_{\mathbf{p}}} = \frac{zV}{h^3} \int_0^\infty dp \, 4\pi p^2 \left(\frac{p^2}{2m} \right) e^{-\beta p^2/2m} = \frac{3}{2} NkT \quad (8.53)$$

Therefore T is the absolute temperature. The entropy is, by (8.48) and (8.46),

$$\begin{aligned} \frac{S}{k} &= z \sum_{\mathbf{p}} e^{-\beta \epsilon_{\mathbf{p}}} (\beta \epsilon_{\mathbf{p}} - \log z) = \beta E - N \log z \\ &= \frac{3}{2} N - N \log \left[\frac{N}{V} \left(\frac{2\pi \hbar^2}{mkT} \right)^{3/2} \right] \end{aligned} \quad (8.54)$$

This is the Sackur-Tetrode equation. The fact that the constant $h = 2\pi \hbar$ is

Planck's constant follows from (8.31), where A first makes its appearance. The equation of state is deduced from the function $U(S, V)$, which is E expressed in terms of S and V . We straightforwardly find $PV = NkT$. It is to be noted that (8.54) does not satisfy the third law of thermodynamics. This should cause no concern, because a Boltzmann gas is not a physical system. It is only a model toward which the Bose and Fermi gas converge at high temperatures. This shows, however, that the third law of thermodynamics is not an automatic consequence of the general principles of quantum mechanics, but depends on the nature of the density of states near the ground state.

The Bose and Fermi gases can be worked out along similar lines. They are more conveniently discussed, however, in the grand canonical ensemble, which we consider in the next section.

8.6 THE IDEAL GASES: GRAND CANONICAL ENSEMBLE

The partition functions for the ideal gases are

$$Q_N(V, T) = \sum_{\{n_p\}} g\{n_p\} e^{-\beta E\{n_p\}} \quad (8.55)$$

where

$$E\{n_p\} = \sum_p \epsilon_p n_p \quad (8.56)$$

and the occupation numbers are subject to the condition

$$\sum_p n_p = N \quad (8.57)$$

For a Bose gas and a Boltzmann gas $n_p = 0, 1, 2, \dots$. For a Fermi gas $n_p = 0, 1$. The number of states corresponding to $\{n_p\}$ is

$$g\{n_p\} = \begin{cases} 1 & \text{(Bose and Fermi)} \\ \frac{1}{N!} \left(\frac{N!}{\prod_p n_p!} \right) & \text{(Boltzmann)} \end{cases} \quad (8.58)$$

We first work out the Boltzmann gas:

$$Q_N = \sum_{\substack{n_0, n_1, \dots \\ \sum n_i = N}} \left(\frac{e^{-\beta n_0 \epsilon_0}}{n_0!} \frac{e^{-\beta n_1 \epsilon_1}}{n_1!} \dots \right) = \frac{1}{N!} (e^{-\beta \epsilon_0} + e^{-\beta \epsilon_1} + \dots)^N$$

This equality is the multinomial theorem. In the limit as $V \rightarrow \infty$ we can write

$$\sum_p e^{-\beta \epsilon_p} = \frac{V}{h^3} \int_0^\infty dp 4\pi p^2 e^{-\beta p^2/2m} = V \left(\frac{mkT}{2\pi \hbar^2} \right)^{3/2} \quad (8.59)$$

Therefore

$$\frac{1}{N} \log Q_N = \log \left[\frac{V}{N} \left(\frac{mkT}{2\pi \hbar^2} \right)^{3/2} \right] \quad (8.60)$$

from which easily follows the Sackur-Tetrode equation for the entropy and the equation of state $PV = NkT$. The grand partition function is trivial and will not be considered.

For the Bose gas and the Fermi gas the partition function cannot be evaluated easily because of the condition (8.57). Instead of the partition function we consider the grand partition function

$$\begin{aligned} \mathcal{Q}(z, V, T) &= \sum_{N=0}^m z^N Q_N(V, T) = \sum_{N=0}^{\infty} \sum_{\substack{\{n_p\} \\ \sum n_p = N}} z^N e^{-\beta \sum \epsilon_p n_p} \\ &= \sum_{N=0}^{\infty} \sum_{\substack{\{n_p\} \\ \sum n_p = N}} \prod_p (z e^{-\beta \epsilon_p})^{n_p} \end{aligned} \quad (8.61)$$

Now it is to be noted that the double summation just given is equivalent to summing each n_p independently. To prove this we must show that every term in one case appears once and only once in the other, and vice versa. This is easily done mentally. Therefore

$$\begin{aligned} \mathcal{Q}(z, V, T) &= \sum_{n_0} \sum_{n_1} \dots [(z e^{-\beta \epsilon_0})^{n_0} (z e^{-\beta \epsilon_1})^{n_1} \dots] \\ &= \left[\sum_{n_0} (z e^{-\beta \epsilon_0})^{n_0} \right] \left[\sum_{n_1} (z e^{-\beta \epsilon_1})^{n_1} \right] \dots \\ &= \prod_p \left[\sum_n (z e^{-\beta \epsilon_p})^n \right] \end{aligned}$$

where the sum \sum_n extends over the values $n = 0, 1, 2, \dots$ for the Bose gas and the values $n = 0, 1$ for the Fermi gas. The results are

$$\mathcal{Q}(z, V, T) = \begin{cases} \prod_p \frac{1}{1 - z e^{-\beta \epsilon_p}} & \text{(Bose)} \\ \prod_p (1 + z e^{-\beta \epsilon_p}) & \text{(Fermi)} \end{cases} \quad (8.62)$$

The equations of state are

$$\frac{PV}{kT} = \log \mathcal{Q}(z, V, T) = \begin{cases} - \sum_p \log(1 - z e^{-\beta \epsilon_p}) & \text{(Bose)} \\ \sum_p \log(1 + z e^{-\beta \epsilon_p}) & \text{(Fermi)} \end{cases} \quad (8.63)$$

from which z is to be eliminated with the help of the equations

$$N = z \frac{\partial}{\partial z} \log \mathcal{Q}(z, V, T) = \begin{cases} \sum_{\mathbf{p}} \frac{z e^{-\beta \epsilon_{\mathbf{p}}}}{1 - z e^{-\beta \epsilon_{\mathbf{p}}}} & \text{(Bose)} \\ \sum_{\mathbf{p}} \frac{e^{-\beta \epsilon_{\mathbf{p}}}}{1 + z e^{-\beta \epsilon_{\mathbf{p}}}} & \text{(Fermi)} \end{cases} \quad (8.64)$$

The average occupation numbers $\langle n_{\mathbf{p}} \rangle$ are given by

$$\begin{aligned} \langle n_{\mathbf{p}} \rangle &\equiv \frac{1}{N} \sum_{N=0}^{\infty} z^N \sum_{\substack{\{n_{\mathbf{p}}\} \\ \sum n_{\mathbf{p}} = N}} n_{\mathbf{p}} e^{-\beta \sum \epsilon_{\mathbf{p}} n_{\mathbf{p}}} = -\frac{1}{\beta} \frac{\partial}{\partial \epsilon_{\mathbf{p}}} \log \mathcal{Q} \\ &= \frac{z e^{-\beta \epsilon_{\mathbf{p}}}}{1 \mp z e^{-\beta \epsilon_{\mathbf{p}}}} \quad \text{(Bose and Fermi)} \end{aligned} \quad (8.65)$$

which are the same as (8.45). The equations (8.64) are non other than the statement

$$N = \sum_{\mathbf{p}} \langle n_{\mathbf{p}} \rangle \quad (8.66)$$

The results here are completely equivalent to those in the microcanonical ensemble, as they should be.

Now we let $V \rightarrow \infty$, and replace sums over \mathbf{p} by integrals over \mathbf{p} in the manner indicated in (8.32), whenever possible. Such a replacement is clearly valid if the summand in question is finite for all \mathbf{p} . In (8.63) and (8.64), the fugacity z is nonnegative for both the ideal Fermi gas and the ideal Bose gas because, if z were negative, then (8.64) cannot be satisfied for positive N . We see immediately that for the ideal Fermi gas it is permissible to replace the sums in (8.63) and (8.64) by integrals over \mathbf{p} . We then obtain the following equation of state.

Ideal Fermi Gas

$$\begin{cases} \frac{P}{kT} = \frac{4\pi}{h^3} \int_0^{\infty} dp p^2 \log(1 + z e^{-\beta p^2/2m}) \\ \frac{1}{v} = \frac{4\pi}{h^3} \int_0^{\infty} dp p^2 \frac{1}{z^{-1} e^{\beta p^2/2m} + 1} \end{cases} \quad (8.67)$$

where $v = V/N$. It can be verified in a straightforward fashion that (8.67) can also be written in the form

$$\begin{cases} \frac{P}{kT} = \frac{1}{\lambda^3} f_{5/2}(z) \\ \frac{1}{v} = \frac{1}{\lambda^3} f_{3/2}(z) \end{cases} \quad (8.68)$$

where $\lambda = \sqrt{2\pi\hbar^2/mkT}$ and

$$f_{5/2}(z) \equiv \frac{4}{\sqrt{\pi}} \int_0^{\infty} dx x^2 \log(1 + z e^{-x^2}) = \sum_{l=1}^{\infty} \frac{(-1)^{l+1} z^l}{l^{5/2}} \quad (8.69)$$

$$f_{3/2}(z) \equiv z \frac{\partial}{\partial z} f_{5/2}(z) = \sum_{l=1}^{\infty} \frac{(-1)^{l+1} z^l}{l^{3/2}} \quad (8.70)$$

For the ideal Bose gas the summands appearing in (8.63) and (8.64) diverge as $z \rightarrow 1$, because the single term corresponding to $\mathbf{p} = 0$ diverges. Thus the single term $\mathbf{p} = 0$ may be as important as the entire sum.* We split off the terms in (8.63) and (8.64) corresponding to $\mathbf{p} = 0$ and replace the rest of the sums by integrals. We then obtain the following equation of state.

Ideal Bose Gas

$$\begin{cases} \frac{P}{kT} = -\frac{4\pi}{h^3} \int_0^{\infty} dp p^2 \log(1 - z e^{-\beta p^2/2m}) - \frac{1}{V} \log(1 - z) \\ \frac{1}{v} = \frac{4\pi}{h^3} \int_0^{\infty} dp p^2 \frac{1}{z^{-1} e^{\beta p^2/2m} - 1} + \frac{1}{V} \frac{z}{1 - z} \end{cases} \quad (8.71)$$

where $v = V/N$. It can be verified in a straightforward fashion that (8.71) can also be written in the form

$$\begin{cases} \frac{P}{kT} = \frac{1}{\lambda^3} g_{5/2}(z) - \frac{1}{V} \log(1 - z) \\ \frac{1}{v} = \frac{1}{\lambda^3} g_{3/2}(z) + \frac{1}{V} \frac{z}{1 - z} \end{cases} \quad (8.72)$$

where $\lambda = \sqrt{2\pi\hbar^2/mkT}$, and

$$g_{5/2}(z) \equiv -\frac{4}{\sqrt{\pi}} \int_0^{\infty} dx x^2 \log(1 - z e^{-x^2}) = \sum_{l=1}^{\infty} \frac{z^l}{l^{5/2}} \quad (8.73)$$

$$g_{3/2}(z) \equiv z \frac{\partial}{\partial z} g_{5/2}(z) = \sum_{l=1}^{\infty} \frac{z^l}{l^{3/2}} \quad (8.74)$$

As (8.65) implies, the quantity $z/(1 - z)$ is the average occupation number $\langle n_0 \rangle$ for the single-particle level with $\mathbf{p} = 0$:

$$\frac{z}{1 - z} = \langle n_0 \rangle \quad (8.75)$$

This term contributes significantly to (8.72) if $\langle n_0 \rangle/V$ is a finite number, i.e., if a

*That this is in fact the case is shown in Section 12.3 in connection with the Bose-Einstein condensation.

finite fraction of all the particles in the system occupy the single level with $\mathbf{p} = 0$. We shall see in Section 11.3 that such a circumstance gives rise to the phenomenon of Bose-Einstein condensation.

The internal energy for both the Fermi and the Bose gases may be found from the formula

$$U(z, V, T) = \frac{1}{\mathcal{Q}} \sum_{N=0}^{\infty} z^N \sum_{\substack{\{n_{\mathbf{p}}\} \\ \sum n_{\mathbf{p}} = N}} \left[e^{-\beta \sum \epsilon_{\mathbf{p}} n_{\mathbf{p}}} \sum_{\mathbf{p}} \epsilon_{\mathbf{p}} n_{\mathbf{p}} \right] = - \frac{\partial}{\partial \beta} [\log \mathcal{Q}(z, V, T)] \tag{8.76}$$

Since $\log \mathcal{Z} = PV/kT$, we obtain from (8.68) and (8.72) the results

$$\frac{1}{V} U(z, V, T) = \begin{cases} \frac{3}{2} \frac{kT}{\lambda^3} f_{5/2}(z) & \text{(Fermi)} \\ \frac{3}{2} \frac{kT}{\lambda^3} g_{5/2}(z) & \text{(Bose)} \end{cases} \tag{8.77}$$

To express U in terms of N , V , and T , we must eliminate z . The result would be a very complicated function. A comparison between (8.77), (8.68), and (8.72), however, shows that U is directly related to the pressure by*

$$U = \frac{3}{2} PV \quad \text{(Bose and Fermi)} \tag{8.78}$$

This relation also holds for the ideal Boltzmann gas.

The detailed study of the ideal gases together with their applications is taken up in Chapters 11 and 12.

8.7 FOUNDATIONS OF STATISTICAL MECHANICS

The present section contains no derivations. It merely furnishes an orientation on the subject of the derivation of statistical mechanics from molecular dynamics.†

It is recalled that a special case of statistical mechanics, the classical kinetic theory of gases, can be derived almost rigorously from molecular dynamics. The only ad hoc assumption in that derivation is the assumption of molecular chaos, which, however, plays a well-understood role, namely, the reduction of reversible microscopic phenomena to irreversible macroscopic phenomena. Since irreversibility is a necessary result of any successful derivation, an assumption of this kind is not only unavoidable but also desirable, because it serves to mark clearly the point at which irreversibility enters. An improvement on the existing derivation consists of replacing this assumption by one less ad hoc, but not of doing away with it altogether.

*It is assumed that the term $V^{-1} \log(1 - z)$ in (8.72) can be neglected. This is justified in Section 12.3.

†For a source of literature see *Fundamental Problems in Statistical Mechanics*, edited by E. G. D. Cohen (North-Holland, Amsterdam, 1962).

The derivation of the classical kinetic theory of gases may be considered largely satisfactory. When we consider the more general problem of the derivation of statistical mechanics, we may well keep this theory in mind as a model example. From this example, we learn that a satisfactory derivation of statistical mechanics must simultaneously fulfill two requirements:

- (a) It must clearly display the connection between microscopic reversibility and macroscopic irreversibility.
- (b) It must provide a detailed description of the approach to equilibrium.

Thus a satisfactory derivation of statistical mechanics must satisfy not only the philosophical desire of the physicist to base all natural phenomena on molecular dynamics, but also the practical desire of the physicist to calculate numbers with which to compare with experiments.

Logically speaking, it suffices to derive quantum statistical mechanics, of which classical statistical mechanics is a special case. If we want to understand nonequilibrium phenomenon in the classical domain, however, it is expedient to use classical mechanics as a starting point. For this reason attempts to derive classical statistical mechanics from classical mechanics can be of great practical value.

Attempts to derive statistical mechanics have so far been one of two types: Some appeal to the ergodic theorem, while others aim at establishing the "master equation." Only the latter seems capable of fulfilling both the requirements set forth previously.

The master equation is an equation governing the time development of the quantity $P_n(t)$, which is the probability that at the instant t the system is in the state n . If the word "state" is appropriately interpreted, $P_n(t)$ can be defined either in classical or quantum mechanics. To justify statistical mechanics, we have to show that $P_n(t)$ approaches the quantity (c_n, c_n) of (8.5) when t is much longer than a characteristic time of the system called the relaxation time, e.g., molecular collision time.

The master equation is

$$\frac{dP_n(t)}{dt} = \sum_m [W_{nm} P_m(t) - W_{mn} P_n(t)] \tag{8.79}$$

where W_{mn} is the transition probability per second from the state n to the state m . It was first derived by Pauli under the assumption that n refers to a single quantum state of the system and that the coefficients in the expansion (8.1) have random phases at all times. All subsequent work after Pauli's has been concerned with the improvement of these assumptions and with the solution of the master equation itself.

It can be shown that solutions to the master equation approach the desired limit as $t \rightarrow \infty$. Hence the task of deriving statistical mechanics reduces to the justification of the master equation and the calculation of the relaxation time.

The similarity between the master equation and the Boltzmann transport equation may be noted, although we should remember that the latter refers to μ space whereas the former refers to Γ space. The random-phase assumption here is similar to the assumption of molecular chaos in the Boltzmann transport equation. In both cases the solution for $t \rightarrow \infty$ is relatively easy to obtain, but the relaxation time is difficult to calculate.

The approach involving the master equation seems to hold greater promise for a satisfactory derivation of statistical mechanics and the concomitant understanding of general nonequilibrium phenomena. Further discussion of the master equation, however, is beyond the scope of this book.*

PROBLEMS

8.1 Find the density matrix for a partially polarized incident beam of electrons in a scattering experiment, in which a fraction f of the electrons are polarized along the direction of the beam and a fraction $1 - f$ is polarized opposite to the direction of the beam.

8.2 Derive the equations of state (8.67) and (8.71), using the microcanonical ensemble.

8.3 Prove (7.14) in quantum statistical mechanics.

8.4 Verify (8.49) for Fermi and Bose statistics, i.e., the fluctuations of cell occupations are small.

Solution. By (8.65),

$$\langle n_k \rangle = -\frac{1}{\beta} \frac{\partial}{\partial \epsilon_k} \log \mathcal{Q}$$

Differentiating this again with respect to ϵ_k leads to

$$\langle n_k^2 \rangle - \langle n_k \rangle^2 = -\frac{1}{\beta} \frac{\partial}{\partial \epsilon_k} \langle n_k \rangle$$

from which we can deduce

$$\langle n_k^2 \rangle - \langle n_k \rangle^2 = \langle n_k \rangle \pm \langle n_k \rangle^2 \tag{A}$$

with the plus sign for Bose statistics, and the minus sign for Fermi statistics. (For Fermi statistics the result is obvious because $n_k^2 = n_k$.) The fluctuations are not necessarily small. Note, however, that (A) refers to the fluctuations of the occupation of individual states, and not the cell occupations.

As a calculation useful for later purposes, we note

$$\langle n_k n_p \rangle - \langle n_k \rangle \langle n_p \rangle = -\frac{1}{\beta} \frac{\partial}{\partial \epsilon_k} \langle n_p \rangle, \quad (p \neq k)$$

The right side is zero because $\langle n_p \rangle$ depends only on ϵ_p . Thus we have

$$\langle n_k n_p \rangle = \langle n_k \rangle \langle n_p \rangle, \quad (p \neq k) \tag{B}$$

*For a general discussion of the master equation, see N. G. Van Kampen, in Cohen, *op. cit.* An improvement on the random phase approximation is described by L. Van Hove, in Cohen, *op. cit.*

In the infinite-volume limit the spectrum of states becomes a continuum. The physically interesting question concerns the fluctuations in the occupation of a group of states, or a cell. Let

$$n_i = \sum_k n_k$$

where the sum extends over a group of states in cell i . We are interested in

$$\langle n_i^2 \rangle - \langle n_i \rangle^2 = \left\langle \left(\sum_k n_k \right)^2 \right\rangle - \left\langle \sum_k n_k \right\rangle^2$$

By using (B), it is easily shown that the right side is equal to

$$\sum_k (\langle n_k^2 \rangle - \langle n_k \rangle^2)$$

Hence using (A) we obtain

$$\langle n_i^2 \rangle - \langle n_i \rangle^2 = \langle n_i \rangle \pm \sum_k \langle n_k \rangle^2$$

where the plus sign holds for Bose statistics, and the minus sign for Fermi statistics. In the infinite-volume limit, the k sum is replaced by an integral over a region in k space. No matter how small this region is, the integral is proportional to the volume V of the system. (This is equivalent to the statement that a finite fraction of the particles occupies a cell.) Thus the left side is of order V^2 , but the right side is only of order V .

8.5 Calculate the grand partition function for a system of N **noninteracting** quantum mechanical harmonic oscillators, all of which have the same natural frequency ω_0 . Do this for the following two cases:

- (a) Boltzmann statistics
- (b) Bose statistics.

Suggestions. Write down the energy levels of the N -oscillator system and determine the degeneracies of the energy levels for the two cases mentioned.

8.6 What is the equilibrium ratio of ortho- to parahydrogen at a temperature of 300 K? What is the ratio in the limit of high temperatures? Assume that the distance between the protons in the molecule is 0.74 Å.

The following hints may be helpful.

- (a) Boltzmann statistics is valid for H_2 molecules at the temperatures considered.
- (b) The energy of a single H_2 molecule is a sum of terms corresponding to contributions from rotational motion, vibrational motion, translational motion, and excitation of the electronic cloud. Only the rotational energy need be taken into account.

(c) The rotational energies are

$$E_{\text{para}} = \frac{\hbar^2}{2I} l(l+1) \quad (l = 0, 2, 4, \dots)$$

$$E_{\text{ortho}} = \frac{\hbar^2}{2I} l(l+1) \quad (l = 1, 3, 5, \dots)$$

where I is the moment of inertia of the H_2 molecule.

Answer. Let T = absolute temperature and $\beta = 1/kT$. Then

$$\frac{N_{\text{ortho}}}{N_{\text{para}}} = \frac{3 \sum_{l \text{ odd}} (2l+1) e^{-(\beta \hbar^2 / 2I) l(l+1)}}{\sum_{l \text{ even}} (2l+1) e^{-(\beta \hbar^2 / 2I) l(l+1)}}$$

CHAPTER

9

GENERAL PROPERTIES OF
THE PARTITION FUNCTION

9.1 THE DARWIN-FOWLER METHOD

Although the canonical ensemble may be derived from the microcanonical ensemble, as we have shown in Section 7.1, it may also be derived directly. Indeed, if we are not too concerned with rigor, the derivation is very simple. Consider an ensemble of M systems such that the energy averaged over all the systems is a given number U . We wish to find the most probable distribution of energies among these M systems in the limit as $M \rightarrow \infty$. By definition of an ensemble, the systems do not interact with one another; they may be considered one at a time, and they are consequently distinguishable from one another. Therefore our problem is mathematically identical with the problem of the most probable distribution for a classical ideal gas of particles. The answer as we know is the Maxwell-Boltzmann distribution, i.e., the energy value E_n occurs among the systems with relative probability $e^{-\beta E_n}$, where β is determined by the average energy U . This ensemble is the canonical ensemble. It is obvious that this derivation holds equally well in quantum and in classical statistical mechanics.

We want to present here a more rigorous derivation that avoids the use of Stirling's approximation, which is necessary in the usual derivation of the Maxwell-Boltzmann distribution. The purpose of this presentation is not only to derive the canonical ensemble directly but also to introduce the method of saddle point integration, which is a useful mathematical tool in statistical mechanics. The considerations that follow hold equally well for quantum and for classical statistical mechanics.

The method we shall describe is due to Darwin and Fowler. Assume that a system in the ensemble may have any one of the energy values E_k ($k = 0, 1, 2, \dots$). By choosing the unit of energy to be sufficiently small, we can regard E_k as an

integer. Among the systems in the ensemble let

$$\begin{aligned} m_0 & \text{ systems have energy } E_0 \\ m_1 & \text{ systems have energy } E_1 \\ & \dots \\ m_k & \text{ systems have energy } E_k \end{aligned} \quad (9.1)$$

The set of integers $\{m_k\}$ describes an arbitrary distribution of energy among the systems. It must satisfy the conditions

$$\begin{aligned} \sum_{k=0}^{\infty} m_k &= M \\ \sum_{k=0}^{\infty} E_k m_k &= MU \end{aligned} \quad (9.2)$$

where both M and U are integers. Our purpose is to find the most probable set $\{\bar{m}_k\}$.

Given an arbitrary set $\{m_k\}$ satisfying (9.2) there are generally more ways than one to construct an ensemble corresponding to (9.1), because the interchange of any two systems (which are distinguishable) leaves $\{m_k\}$ unchanged. Let $W\{m_k\}$ be the number of distinct ways in which we can assign energy values to systems so as to satisfy (9.1). Obviously

$$W\{m_k\} = \frac{M!}{m_0! m_1! m_2! \dots} \quad (9.3)$$

For the present case the postulate of equal a priori probability states that all distributions in energy among the systems are equally probable, subject to the conditions (9.2). Thus $\{\bar{m}_k\}$ is the set that maximizes (9.3). In anticipation of the fact that in the limit as $M \rightarrow \infty$ almost all possible sets $\{m_k\}$ are identical with $\{\bar{m}_k\}$, we can also find $\{\bar{m}_k\}$ by calculating the value of m_k averaged over all possible distributions in energy:

$$\langle m_k \rangle \equiv \frac{\sum' m_k W\{m_i\}}{\sum' W\{m_i\}} \quad (9.4)$$

where a prime over the sums indicate that they are sums over all sets $\{m_k\}$ subject to (9.2). We must also calculate the mean square fluctuation $\langle m_k^2 \rangle - \langle m_k \rangle^2$. If this vanishes as $M \rightarrow \infty$, then in that limit $\langle m_k \rangle \rightarrow \bar{m}_k$.

For convenience we modify the definition of $W\{m_k\}$ to

$$W\{m_k\} = \frac{M! g_0^{m_0} g_1^{m_1} \dots}{m_0! m_1! \dots} \quad (9.5)$$

where g_k is a number which at the end of the calculation will be set equal to unity. Let

$$\Gamma(M, U) \equiv \sum'_{\{m_i\}} W\{m_i\} \quad (9.6)$$

Then

$$\langle m_k \rangle = g_k \frac{\partial}{\partial g_k} \log \Gamma \quad (9.7)$$

The mean square fluctuation can be obtained as follows:

$$\begin{aligned} \langle m_i^2 \rangle &= \frac{1}{\Gamma} \sum'_{\{m_i\}} m_i^2 W\{m_i\} = \frac{1}{\Gamma} g_k \frac{\partial}{\partial g_k} \left(g_k \frac{\partial \Gamma}{\partial g_k} \right) \\ &= g_k \frac{\partial}{\partial g_k} \left(\frac{1}{\Gamma} g_k \frac{\partial \Gamma}{\partial g_k} \right) - \left(\frac{\partial}{\partial g_k} \frac{1}{\Gamma} \right) g_k^2 \frac{\partial \Gamma}{\partial g_k} \\ &= g_k \frac{\partial}{\partial g_k} \left(g_k \frac{\partial}{\partial g_k} \log \Gamma \right) + \left(g_k \frac{\partial}{\partial g_k} \log \Gamma \right)^2 \end{aligned}$$

Therefore

$$\langle m_k^2 \rangle - \langle m_k \rangle^2 = g_k \frac{\partial}{\partial g_k} \left(g_k \frac{\partial}{\partial g_k} \log \Gamma \right) \quad (9.8)$$

Thus it is sufficient to calculate $\log \Gamma$

By (9.6) and (9.5)

$$\Gamma = M! \sum'_{m_0, m_1, \dots} \left(\frac{g_0^{m_0}}{m_0!} \cdot \frac{g_1^{m_1}}{m_1!} \dots \right) \quad (9.9)$$

This cannot be explicitly evaluated because of the restriction (9.2). We are, however, only interested in this quantity in the limit as $M \rightarrow \infty$. To proceed, we define a generating function for Γ in the following manner. For any complex number z , let

$$G(M, z) \equiv \sum_{U=0}^{\infty} z^{MU} \Gamma(M, U) \quad (9.10)$$

Using (9.9) and (9.2) we obtain

$$G(M, z) = M! \sum_{U=0}^{\infty} \sum'_{m_0, m_1, \dots} \left[\frac{(g_0 z^{E_0})^{m_0}}{m_0!} \cdot \frac{(g_1 z^{E_1})^{m_1}}{m_1!} \dots \right] \quad (9.11)$$

It is easily seen that the double sum in (9.11) is equivalent to summing over all sets $\{m_k\}$ subject only to the condition $\sum m_k = M$. To show this we need only verify that every term in one sum appears once in the other and vice versa. Hence

$$\begin{aligned} G(M, z) &= \sum_{\substack{m_0, m_1, \dots \\ \sum m_k = M}} \frac{M!}{m_0! m_1! \dots} \left[(g_0 z^{E_0})^{m_0} (g_1 z^{E_1})^{m_1} \dots \right] \\ &= (g_0 z^{E_0} + g_1 z^{E_1} + \dots)^M \end{aligned} \quad (9.12)$$

The last step follows by use of the multinomial theorem. Let

$$f(z) \equiv \sum_{k=0}^{\infty} g_k z^{E_k} \quad (9.13)$$

Then

$$G(M, z) = [f(z)]^M \quad (9.14)$$

To obtain $\Gamma(M, U)$ from $G(M, z)$ we note that by definition $\Gamma(M, U)$ is the coefficient of z^{MU} in the expansion of $G(M, z)$ in powers of z . Therefore

$$\Gamma(M, U) = \frac{1}{2\pi i} \oint dz \frac{[f(z)]^M}{z^{MU+1}} \quad (9.15)$$

where the contour of integration is a closed path in the complex z plane about $z = 0$.

We may assume without loss of generality that the sequence E_0, E_1, \dots is a sequence of nondecreasing integers with no common divisor, because any common division τ can be removed by choosing the unit of energy τ times larger. Furthermore, we can set $E_0 = 0$, since this would only change the zero point of the energy. In so doing U would be changed to $U - E_0$, which we can again call U . The numbers g_0, g_1, \dots are as close to unity as we wish. For the immediate calculations we omit them temporarily. Hence

$$f(z) = 1 + z^{E_1} + z^{E_2} + \dots \quad (E_1 \leq E_2 \leq E_3 \leq \dots) \quad (9.16)$$

where E_1, E_2, \dots are integers with no common divisor. When z is a real positive number x , $f(x)$ is a monotonically increasing function of x with a radius of convergence at, say, $x = R$. The same is true for $[f(x)]^M$, as illustrated in Fig. 9.1. The function $1/z^{MU+1}$, on the other hand, is a monotonically decreasing function along the real positive axis. The product of these two functions has a minimum at x_0 between 0 and R , as shown in Fig. 9.1. Now $f(z)$ is an analytic function for $|z| < R$, and z^{-MU-1} is analytic everywhere except at $z = 0$. Therefore the integrand of (9.18)

$$I(z) = \frac{[f(z)]^M}{z^{MU+1}} \quad (9.17)$$

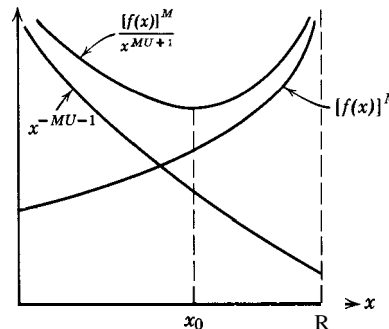


Fig. 9.1 The function $[f(x)]^M/x^{MU+1}$ for real positive x .

is analytic at $z = x_0$. An analytic function has a unique derivative at a given point. Furthermore it satisfies the Cauchy-Riemann equation

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) I(z) = 0 \quad (z \equiv x + iy) \quad (9.18)$$

Hence

$$\left(\frac{\partial I}{\partial z} \right)_{z=x_0} = 0, \quad \left(\frac{\partial^2 I}{\partial x^2} \right)_{z=x_0} > 0, \quad \left(\frac{\partial^2 I}{\partial y^2} \right)_{z=x_0} < 0 \quad (9.19)$$

That is, in the complex z plane, $I(z)$ has a minimum at $z = x_0$ along a path on the real axis but has a maximum at $z = x_0$ along a path parallel to the imaginary axis passing through x_0 . The point x_0 is a saddle point, as illustrated in Fig. 9.2. Let $g(z)$ be defined by

$$\begin{aligned} I(z) &\equiv e^{Mg(z)} \\ g(z) &= \log f(z) - U \log z \end{aligned} \quad (9.20)$$

where we have neglected $1/M$ as compared to U . Then x_0 is the root of the equation

$$g'(x_0) = 0$$

or

$$\frac{\sum_k E_k x_0^{E_k}}{\sum_k x_0^{E_k}} = U \quad (9.21)$$

Furthermore

$$\left(\frac{\partial^2 I}{\partial x^2} \right)_{z=x_0} = Mg''(x_0) \exp [Mg(x_0)] \xrightarrow{M \rightarrow \infty} \infty \quad (9.22)$$

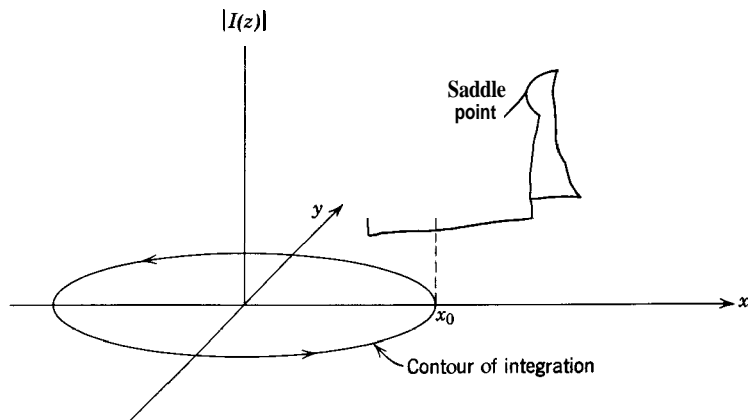


Fig. 9.2 The saddle point.

Hence the saddle point touches an infinitely sharp peak and an infinitely steep valley in the limits as $M \rightarrow \infty$. If we choose the contour of integration to be a circle about $z = 0$ with radius x_0 , the contour will pass through x_0 in the imaginary direction. Thus along the contour the integrand has an extremely sharp maximum at the point $z = x_0$. If elsewhere along the contour there is no maximum comparable in height to this one, the contribution to the integral comes solely from the neighborhood of x_0 . This is in fact true because for $z = x_0 e^{i\theta}$

$$|I(z)| = \frac{1}{x_0^{MU}} \left| 1 + (x_0 e^{i\theta})^{E_1} + (x_0 e^{i\theta})^{E_2} + \dots \right|^M \quad (9.23)$$

The series (9.23) is maximum when all terms are real. This happens when and only when $\theta E_k = 2\pi n_k$, where n_k is 0 or an integer. If $\theta \neq 0$, then $2\pi/\theta$ must be a rational number, and this would mean that $E_k = (2\pi/\theta)n_k$, which is impossible unless $\theta = 2\pi$, because the E_k have no common divisor. Hence we conclude that the largest value of $I(z)$ occurs at $z = x_0$.

To do the integral (9.15) we expand the integrand about $z = x_0$:

$$g(z) = g(x_0) + \frac{1}{2}(z - x_0)^2 g''(x_0) + \dots$$

Hence

$$\Gamma(M, U) = \frac{1}{2\pi i} \oint dz e^{Mg(z)} \approx e^{Mg(x_0)} \frac{1}{2\pi i} \oint dz e^{1/2 M g''(x_0)(z-x_0)^2}$$

Putting $(z - x_0) = iy$, we obtained

$$\Gamma(M, U) \approx e^{Mg(x_0)} \frac{1}{2\pi} \int_{-\infty}^{+\infty} dy e^{-1/2 M g''(x_0)y^2} = \frac{e^{Mg(x_0)}}{\sqrt{2\pi M g''(x_0)}} \quad (9.24)$$

Hence

$$\frac{1}{M} \log \Gamma(M, U) \approx g(x_0) - \frac{1}{2M} \log [2\pi M g''(x_0)] \quad (9.25)$$

As $M \rightarrow \infty$ the first term gives the exact result. To evaluate $g(x_0)$ we first obtain from (9.20) the formulas

$$\begin{aligned} g(x_0) &= \log f(x_0) - U \log x_0 \\ g''(x_0) &= \frac{f''(x_0)}{f(x_0)} - \frac{U(U-1)}{x_0^2} \end{aligned}$$

Using $f(x_0)$ from (9.13) (restoring now the numbers g_k) and defining a parameter β by

$$x_0 \equiv e^{-\beta} \quad (9.26)$$

we obtain

$$g(x_0) = \log \left(\sum_{k=0}^{\infty} g_k e^{-\beta E_k} \right) + \beta U$$

$$g''(x_0) = \frac{e^{2\beta} \sum_{k=0}^{\infty} g_k (E_k^2 - U^2) e^{-\beta E_k}}{\sum_{k=0}^{\infty} g_k e^{-\beta E_k}} \equiv e^{2\beta} \langle (E - U)^2 \rangle \quad (9.27)$$

Hence

$$\frac{1}{M} \log \Gamma(M, U) = \log \left(\sum_{k=0}^{\infty} g_k e^{-\beta E_k} \right) + \beta U - \frac{1}{2M} \log [2\pi M g''(x_0)] \quad (9.28)$$

from which we find, using (9.7) and (9.8),

$$\frac{\langle m_k \rangle}{M} = \frac{e^{-\beta E_k}}{\sum_{k=0}^{\infty} e^{-\beta E_k}} \quad (9.29)$$

$$\frac{\langle m_k^2 \rangle - \langle m_k \rangle^2}{M^2} = \frac{1}{M} \frac{\langle m_k \rangle}{M} \left[1 - \frac{\langle m_k \rangle}{M} - \frac{\langle m_k \rangle}{M} \frac{(E_k - U)^2}{\langle (E - U)^2 \rangle} \right] \quad (9.30)$$

This is an exact formula in the limit as $M \rightarrow \infty$. We see that the fluctuations vanish in this limit. Therefore $\langle m_k \rangle = \bar{m}_k$. The parameter β is determined by (9.21) and (9.26):

$$U = \frac{\sum_{k=0}^{\infty} E_k e^{-\beta E_k}}{\sum_{k=0}^{\infty} e^{-\beta E_k}} = \langle E \rangle \quad (9.31)$$

Hence β can be identified as $1/kT$, where T is the absolute temperature.

In the most probable distribution the probability of finding a system in the ensemble having the energy E_k is (9.29). The ensemble with such an energy distribution is the canonical ensemble.

9.2 CLASSICAL LIMIT OF THE PARTITION FUNCTION

Let \mathcal{H} be the Hamiltonian operator of a system of N identical spinless particles.* Let \mathcal{K} be the sum of two operators, the kinetic energy operator K , and the potential energy operator Ω :

$$\mathcal{H} = K + \Omega \quad (9.32)$$

*It is straightforward to generalize the following considerations to the case of particles with spin and to the case of a mixed system of two or more more different kinds of particles.

If $\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N)$ is a wave function of the system, then

$$K\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = -\frac{\hbar^2}{2m} \sum_{i=1}^N \nabla_i^2 \Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) \quad (9.33)$$

$$\Omega\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \Omega(\mathbf{r}_1, \dots, \mathbf{r}_N)\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) \quad (9.34)$$

where m is the mass of a particle, ∇_i^2 is the Laplacian operator with respect to \mathbf{r}_i , and $\Omega(\mathbf{r}_1, \dots, \mathbf{r}_N)$ is a sum of two-body potentials:

$$\Omega(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{i<j} v_{ij} \quad (9.35)$$

where

$$v_{ij} \equiv v(|\mathbf{r}_i - \mathbf{r}_j|) \quad (9.36)$$

Whenever convenient, we use the abbreviation $(1, \dots, N)$ for $(\mathbf{r}_1, \dots, \mathbf{r}_N)$.

The partition function of the system is

$$Q_N(V, T) = \text{Tr} e^{-\beta \mathcal{H}} = \sum_n (\Phi_n, e^{-\beta \mathcal{H}} \Phi_n) \quad (9.37)$$

where $\Phi_n(1, \dots, N)$ is a member of any complete orthonormal set of wave functions of the system and $\Phi_n^*(1, \dots, N)$ is its complex conjugate. For any operator \mathcal{O} ,

$$(\Phi_n, \mathcal{O} \Phi_n) \equiv \int d^{3N} \mathbf{r} \Phi_n^*(1, \dots, N) \mathcal{O} \Phi_n(1, \dots, N) \quad (9.38)$$

Each Φ_n satisfies the boundary conditions imposed on the system and is normalized in the box of volume V containing the system. It is a symmetric (antisymmetric) function of $\mathbf{r}_1, \dots, \mathbf{r}_N$, if the system is a system of bosons (fermions).

It will be shown that when the temperature is sufficiently high we can make the approximation

$$Q_N(V, T) \approx \frac{1}{N! h^{3N}} \int d^{3N} p d^{3N} \mathbf{r} e^{-\beta \mathcal{H}(p, r)} \quad (9.39)$$

where h is Planck's constant and $\mathcal{H}(p, r)$ is the classical Hamiltonian

$$\mathcal{H}(p, r) \equiv \sum_{i=1}^N \frac{p_i^2}{2m} + \Omega(\mathbf{r}_1, \dots, \mathbf{r}_N) \quad (9.40)$$

This will prove that at sufficiently high temperatures the partition function approaches the classical partition function with "correct Boltzmann counting." In the course of proving (9.39) we obtain the criterion for a "sufficiently high temperature."

Free Particles

Let us first consider an ideal gas, for which $\Omega(1, \dots, N) \equiv 0$. The eigenfunctions of the Hamiltonian are the free-particle wave functions $\Phi_p(1, \dots, N)$ described in

the Appendix, Section A.2. They are labeled by a set of N momenta

$$P \equiv \{\mathbf{p}_1, \dots, \mathbf{p}_N\} \quad (9.41)$$

and satisfy the eigenvalue equation

$$K\Phi_p(1, \dots, N) = K_p\Phi_p(1, \dots, N) \quad (9.42)$$

where

$$K_p \equiv \frac{1}{2m}(p_1^2 + \dots + p_N^2) \quad (9.43)$$

For convenience we impose periodic boundary conditions with respect to the volume V . It follows that each \mathbf{p}_i has the allowed values

$$\mathbf{p}_i = \frac{2\pi\hbar\mathbf{n}}{V^{1/3}} \quad (9.44)$$

where \mathbf{n} is a vector whose components may be $0, \pm 1, \pm 2, \dots$. More explicitly, $\Phi_p(1, \dots, N)$ is given by

$$\begin{aligned} \Phi_p(1, \dots, N) &= \frac{1}{\sqrt{N!}} \sum_P \delta_P [u_{\mathbf{p}_1}(P1) \dots u_{\mathbf{p}_N}(PN)] \\ &= \frac{1}{\sqrt{N!}} \sum_P \delta_P [u_{P\mathbf{p}_1}(1) \dots u_{P\mathbf{p}_N}(N)] \end{aligned} \quad (9.45)$$

where

$$u_{\mathbf{p}}(\mathbf{r}) = \frac{1}{\sqrt{V}} e^{i\mathbf{p}\cdot\mathbf{r}/\hbar} \quad (9.46)$$

The notation is explained in the Appendix, Section A.2. A permutation of the momenta $\mathbf{p}_1, \dots, \mathbf{p}_N$ does not produce a new state, for Φ_p is either invariant or changes sign. Therefore a sum over states is $1/N!$ times a sum over all the momenta independently. In the limit $V \rightarrow \infty$ a momentum sum may be replaced by an integral:

$$\sum_p \rightarrow \frac{V}{h^3} \int d^3p \quad (9.47)$$

Therefore

$$\text{Tr} e^{-\beta K} = \frac{V^N}{N!h^{3N}} \int d^{3N}p d^{3N}r |\Phi_p(1, \dots, N)|^2 e^{-\beta K_p} \quad (9.48)$$

Using (9.45) we can write

$$|\Phi_p(1, \dots, N)|^2 = \frac{1}{N!} \sum_P \sum_{P'} \delta_P \delta_{P'} [u_{\mathbf{p}_1}^*(P1) u_{P'\mathbf{p}_1}(1)] \dots [u_{\mathbf{p}_N}^*(PN) u_{P'\mathbf{p}_N}(N)] \quad (9.49)$$

Now every term in the P' sum will give the same contribution to the integral in

(9.48). Thus we may replace the above by $N!$ times any one term in the P' sum:

$$\begin{aligned} |\Phi(1, \dots, N)|^2 &\rightarrow \sum_P \delta_P [u_{\mathbf{p}_1}^*(P1) u_{\mathbf{p}_1}(1)] \dots [u_{\mathbf{p}_N}^*(PN) u_{\mathbf{p}_N}(N)] \\ &= \frac{1}{V^N} \sum_P \delta_P \exp \frac{i}{\hbar} [\mathbf{p}_1 \cdot (\mathbf{r}_1 - P\mathbf{r}_1) + \dots + \mathbf{p}_N \cdot (\mathbf{r}_1 - P\mathbf{r}_N)] \end{aligned} \quad (9.50)$$

When this is substituted into (9.48), each momentum integral can be expressed in term of the function

$$f(\mathbf{r}) \equiv \frac{\int d^3p e^{-\beta(p^2/2m) + i\mathbf{p}\cdot\mathbf{r}/\hbar}}{\int d^3p e^{-\beta p^2/2m}} = e^{-\pi r^2/\lambda^2} \quad (9.51)$$

where $r \equiv |\mathbf{r}|$ and $\lambda = \sqrt{2\pi\hbar^2/mkT}$, the thermal wave length. The result is

$$\begin{aligned} \text{Tr} e^{-\beta K} &= \frac{1}{N!h^{3N}} \int d^{3N}p d^{3N}r e^{-\beta(p_1^2 + \dots + p_N^2)/2m} \\ &\quad \times \sum_P \delta_P [f(\mathbf{r}_1 - P\mathbf{r}_1) \dots f(\mathbf{r}_N - P\mathbf{r}_N)] \end{aligned} \quad (9.52)$$

This is an exact identity. For high temperatures the integrand may be approximated as follows. The sum \sum_P contains $N!$ terms. The term corresponding to the unit permutation $P = 1$ is $[f(0)]^N = 1$. The term corresponding to a permutation which only interchanges \mathbf{r}_i and \mathbf{r}_j is $[f(\mathbf{r}_i - \mathbf{r}_j)]^2$. Thus by enumerating the permutations in increasing order of the number of coordinates interchanged we arrive at the expansion

$$\sum_P \delta_P [f(\mathbf{r}_1 - P\mathbf{r}_1) \dots f(\mathbf{r}_N - P\mathbf{r}_N)] = 1 \pm \sum_{i < j} f_{ij}^2 + \sum_{i, j, k} f_{ij} f_{ik} f_{kj} \pm \dots \quad (9.53)$$

where $f_{ij} \equiv f(\mathbf{r}_i - \mathbf{r}_j)$ and where the plus sign applies to bosons and the minus sign to fermions. According to (9.51), f_{ij} vanishes extremely rapidly if $|\mathbf{r}_i - \mathbf{r}_j| \gg \Lambda$. Therefore, when the temperature is so high that

$$(\text{thermal wavelength}) \ll (\text{average interparticle distance}) \quad (9.54)$$

we have

$$\text{Tr} e^{-\beta K} \approx \frac{1}{N!h^{3N}} \int d^{3N}p d^{3N}r e^{-\beta(p_1^2 + \dots + p_N^2)/2m} \quad (9.55)$$

which proves (9.39) for an ideal gas.

It is of some interest to examine the first quantum correction to the classical partition function of an ideal gas. If $|\mathbf{r}_i - \mathbf{r}_j| \gg \Lambda$, we may approximate the

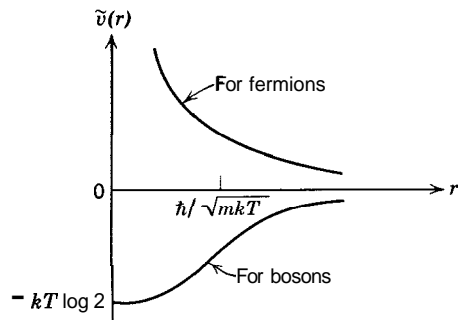


Fig. 9.3 The "statistical potential" between particles in an ideal gas arising from the symmetry properties of the N -particle wave function.

right side of (9.53) by $1 \pm \sum f_{ij}^2$. To the same order of approximation, however, we can also write

$$1 \pm \sum_{i<j} f_{ij}^2 \approx \prod_{i<j} (1 \pm f_{ij}^2) = \exp\left(-\beta \sum_{j<i} \tilde{v}_{ij}\right) \quad (9.56)$$

where

$$\tilde{v}_{ij} \equiv -kT \log(1 \pm f_{ij}^2) = -kT \log\left[1 \pm \exp\left(-\frac{2\pi|\mathbf{r}_i - \mathbf{r}_j|^2}{\lambda^2}\right)\right] \quad (9.57)$$

with the plus sign for bosons and the minus sign for fermions. Therefore an improvement over (9.55) is the formula

$$\text{Tr} e^{-\beta K} \approx \frac{1}{N! h^{3N}} \int d^{3N} p d^{3N} r \exp\left[-\beta \left(\sum_i \frac{p_i^2}{2m} + \sum_{i<j} \tilde{v}_{ij}\right)\right] \quad (9.58)$$

This shows that the first quantum correction to the partition function of an ideal gas has the same effect as that of endowing the particles with an interparticle potential* $\tilde{v}(r)$ and treating the problems classically. The potential $\tilde{v}(r)$ is attractive for bosons and repulsive for fermions, as illustrated in Fig. 9.3. In this sense we sometimes speak of the "statistical attraction" between bosons and the "statistical repulsion" between fermions. It must be emphasized, however, that $\tilde{v}(r)$ originates solely from the symmetry properties of the wave function. Furthermore, it depends on the temperature and thus cannot be regarded as a true interparticle potential.

Interacting Particles

We now turn our attention to the more general case in which the particles of the system interact with one another. For the calculation of traces we may continue

to use the free-particle wave functions Φ_p , because any complete orthonormal set of wave functions will do.

First it is to be noted that in general K does not commute with Ω . Hence

$$e^{-\beta \mathcal{H}} = e^{-\beta(K+\Omega)} \neq e^{-\beta K} e^{-\beta \Omega} \quad (9.59)$$

because the left side is invariant under the exchange of K and Ω whereas the right side is not. To find a suitable approximation for $e^{-\beta \mathcal{H}}$ when $\beta \rightarrow 0$, let us assume that the following expansion is possible:

$$e^{-\beta(K+\Omega)} = e^{-\beta K} e^{-\beta \Omega} e^{C_0} e^{\beta C_1} e^{\beta^2 C_2} \dots \quad (9.60)$$

where C_0, C_1, C_2, \dots are operators to be determined by taking the n th derivatives of both sides of (9.60) with respect to β and then setting $\beta = 0$. Letting $n = 0, 1, 2, \dots$, we successively find that*

$$\begin{aligned} C_0 &= 0 \\ C_1 &= 0 \\ C_2 &= -\frac{1}{2}[K, \Omega] \end{aligned} \quad (9.61)$$

If $[K, \Omega]$ commutes with both K and Ω , we would find that $C_n = 0$ ($n > 2$). In our case this is untrue but we shall assume that for β sufficiently small the following is a good approximation:

$$e^{-\beta(K+\Omega)} \approx e^{-\beta K} e^{-\beta \Omega} e^{-1/2\beta^2[K, \Omega]} \quad (9.62)$$

Consequently

$$Q_N(V, T) \approx \text{Tr}\left(e^{-\beta K} e^{-\beta \Omega} e^{-1/2\beta^2[K, \Omega]}\right) \quad (9.63)$$

From (9.33) and (9.34) it can be easily verified that

$$\begin{aligned} [K, \Omega] &= -\frac{\hbar^2}{2m} \sum_{i=1}^N \nabla_j^2 \Omega + 2(\nabla_i \Omega) \cdot \nabla_i \\ &= -\frac{\hbar^2}{2m} \sum_{i \neq j} (w_{ij} - 2\mathbf{F}_{ij} \cdot \nabla_i) \end{aligned} \quad (9.64)$$

where

$$\begin{aligned} w_{ij} &= w_{ji} = w(|\mathbf{r}_i - \mathbf{r}_j|), \quad w(r) \equiv \nabla^2 v(r) \\ \mathbf{F}_{ij} &= -\mathbf{F}_{ji} \equiv \nabla_i v(|\mathbf{r}_i - \mathbf{r}_j|) \end{aligned} \quad (9.65)$$

When we exponentiate (9.64), we may act as if the two terms in it commute with

*The exact expansion is known as the Baker-Campbell-Hausdorff theorem. For an elementary derivation see R. M. Wilcox, *J. Math. Phys.* **8**, 962 (1967).

*First discussed by G. E. Uhlenbeck and L. Gropper, *Phys. Rev.* **41**, 79 (1932)

each other because the correction belongs to a higher order in β than we are considering. With this in mind, we substitute (9.64) into (9.63), and again use free-particle states to calculate the trace. The operator ∇_i may then be replaced by $i\mathbf{p}_i/\hbar$. Thus we have

$$Q_N(V, T) \approx \frac{V^N}{N!h^{3N}} \int d^{3N}p d^{3N}r |\Phi_p(1, \dots, N)|^2 \times \exp \left[-\beta \sum_{j=i}^N \left(\frac{p_j^2}{2m} + \frac{i\beta\hbar}{2m} \mathbf{G}_j \cdot \mathbf{p}_j \right) - \beta \sum_{i<j} \left(v_{ij} - \frac{\beta\hbar^2}{2m} w_{ij} \right) \right] \quad (9.66)$$

where

$$\mathbf{G}_i = \sum_{\substack{j=1 \\ (j \neq i)}}^N \mathbf{F}_{ji} \quad (9.67)$$

The momentum integrations can be done, again in term of the function $f(\mathbf{r})$ of (9.51). We then obtain

$$Q_N(V, T) \approx \frac{1}{N!h^{3N}} \int d^{3N}p d^{3N}r e^{-\beta\mathcal{H}(p,r) + \beta^2\Omega'} X \quad (9.68)$$

$$X \equiv \sum_P \delta_P [f(\mathbf{r}_1 - P\mathbf{r}_1 + \mathbf{G}'_1) \dots f(\mathbf{r}_N - P\mathbf{r}_N + \mathbf{G}'_N)]$$

where

$$\Omega' = \frac{\beta\hbar^2}{2m} \sum_{i<j} w_{ij} \quad (9.69)$$

$$\mathbf{G}'_i = \frac{\beta\hbar^2}{2m} \mathbf{G}_i$$

Making an expansion similar to (9.53), and an approximation similar to (9.56), we write

$$X \approx \prod_{i=1}^N f(\mathbf{G}'_i) \prod_{i<j} \left[1 \pm \frac{f(\mathbf{r}_i - \mathbf{r}_j + \mathbf{G}'_i) f(\mathbf{r}_j - \mathbf{r}_i + \mathbf{G}'_j)}{f(\mathbf{G}'_i) f(\mathbf{G}'_j)} \right] \quad (9.70)$$

The first product can be rewritten as

$$\exp \frac{\lambda^2}{16\pi} \sum_{i,j,k} \mathbf{F}_{ji} \cdot \mathbf{F}_{ik}$$

which gives rise to an effective three-body force among the particles. The second product is generally complicated, involving many-body forces. We shall assume that the range and depth of the intermolecular potential is such that we can

neglect the terms G'_i in this product. Then we can state

$$Q_N(V, T) \approx \frac{1}{N!h^{3N}} \int d^{3N}p d^{3N}r e^{-\beta\mathcal{H}_{\text{eff}}} \quad (9.71)$$

$$\mathcal{H}_{\text{eff}} = \mathcal{H} + \Omega_1 + \Omega_2 + \Omega_3 \quad (9.72)$$

$$\Omega_1 = \sum_{i<j} \tilde{v}_{ij} \quad (9.73)$$

$$\Omega_2 = \frac{\lambda^2}{4\pi} \sum_{i<j} w_{ij} \quad (9.74)$$

$$\Omega_3 = \frac{\lambda^2}{16\pi} \sum_{i,j,k} \mathbf{F}_j \cdot \mathbf{F}_{ik} \quad (9.75)$$

9.3 SINGULARITIES AND PHASE TRANSITIONS

Phase transitions are manifested in experiments by the occurrence of singularities in thermodynamic functions, such as the pressure in a liquid-gas system, or the magnetization in a ferromagnet. How is it possible that such singularities arise from the partition function, which seems to be an analytic function of its arguments? The answer lies in the fact that a macroscopic body is close to the idealized thermodynamic limit—the limit of infinite volume with particle density held fixed. As we approach this limit, the partition function can develop singularities, because the limit function of a sequence of analytic functions need not be analytic.

Yang and Lee propose a definite scenario for the occurrence of singularities in the thermodynamic limit, which we shall now describe. It is formal in character, and belongs to a field sometimes known as "rigorous statistical mechanics."*

As a concrete model consider a classical system consisting of N molecules in volume V , interacting with one another through a pairwise potential as depicted in Fig. 9.4. Each molecule is taken to be a hard sphere surrounded by an attractive potential of finite range. Thus, a finite volume V can accommodate at most a finite number of molecules $M(V)$. For $N > M(V)$ the partition function vanishes because at least two molecules must "touch," rendering the energy infinite:

$$Q_N(V) = 0, \quad \text{for } N > M(V) \quad (9.76)$$

where we have suppressed the temperature to simplify the notation. The grand

*C. N. Yang and T. D. Lee, *Phys. Rev.* **87**, 404 (1952); T. D. Lee and C. N. Yang, *Phys. Rev.* **87**, 410 (1952). For rigorous stuff see also D. Ruelle, *Statistical Mechanics* (Benjamin, New York, 1969), Chapters 3 and 5; and J. Glimm and A. Jaffe, *Quantum Physics* (Springer-Verlag, New York, 1981), Chapter 2.

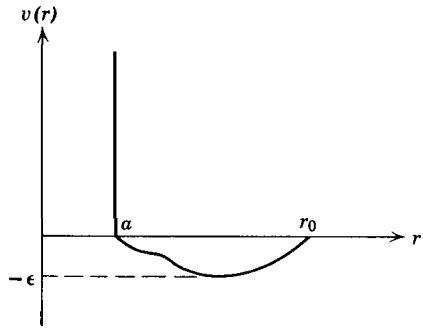


Fig. 9.4 Idealized interparticle potential.

partition function is a polynomial of degree $M(V)$ in the fugacity z :

$$\mathcal{Q}(z, V) = 1 + zQ_1(V) + z^2Q_2(V) + \dots + z^M Q_M \quad (9.77)$$

Since all the coefficients $Q_N(V)$ are positive, the polynomial has no real positive roots. The parametric form of the equation of state is

$$\begin{aligned} \frac{P}{kT} &= V^{-1} \log \mathcal{Q}(z, V) \\ \frac{1}{v} &= V^{-1} z \frac{\partial}{\partial z} \log \mathcal{Q}(z, V) \end{aligned} \quad (9.78)$$

For any finite value of V , however large, both P and v are analytic functions of z in a region of the complex z plane that includes the entire real axis. Therefore P is an analytic function of v in a region of the complex v plane that include all physical values of v , i.e., the real axis. Hence all thermodynamic functions must be free of singularities. From (9.78) and (9.77) we see that $P > 0$, and

$$\frac{\partial P}{\partial v} - \frac{\partial P}{\partial z} \frac{\partial z}{\partial v} - \frac{1}{vz(\partial v/\partial z)} = \frac{kT}{v^4[\langle n^2 \rangle - \langle n \rangle^2]} \leq 0 \quad (9.79)$$

where n is the density. To have the possibility of singularities, we must go to the limit $V \rightarrow \infty$ at fixed v —the thermodynamic limit:

$$\begin{aligned} \frac{P}{kT} &= \lim_{V \rightarrow \infty} V^{-1} \log \mathcal{Q}(z, V) \\ \frac{1}{v} &= \lim_{V \rightarrow \infty} V^{-1} z \frac{\partial}{\partial z} \log \mathcal{Q}(z, V) \end{aligned} \quad (9.80)$$

Note that, in the second equation, the order of the operations \lim and $z(\partial/\partial z)$ can be interchanged only if the limits above are approached uniformly. (For an example see Problem 9.5.)

The above conclusions are valid also for a quantum mechanical system since they depend only on the assumption that a finite volume can accommodate at most a finite number of molecules, which is true for molecules with a hard core, even in quantum mechanics.

Yang and Lee show that phase transitions are controlled by the distribution of roots of the grand partition function in the complex z plane. A phase transition occurs whenever a root approaches the real axis in the limit $V \rightarrow \infty$. The precise results are stated in the form of two theorems.

THEOREM 1

The limit

$$F_\infty(z) \equiv \lim_{V \rightarrow \infty} \frac{1}{V} \log \mathcal{Q}(z, V) \quad (9.81)$$

exists for all $z > 0$, and is a continuous nondecreasing function of z . This limit is also independent of the shape of V , if the surface area of V increases no faster than $V^{2/3}$.

THEOREM 2

Suppose R is a region in the complex z plane that includes a segment of the positive real axis, and contains no roots of the grand partition function. Then in this region the quantity $V^{-1} \log \mathcal{Q}$ converges uniformly to its limit as $V \rightarrow \infty$. The limit is analytic for all z in R .

We refer the reader to the original literature for the proofs, and merely discuss their consequences here.

A thermodynamic phase is defined by those values of z contained in any single region R of theorem 2. Since in any region R the convergence to the limit $F_\infty(z)$ is uniform, we can interchange the order of \lim and $z(\partial/\partial z)$ in (9.80). In any single phase, therefore, the equation of state is given in parametric form by

$$\begin{aligned} \frac{P(z)}{kT} &= F_\infty(z) \\ \frac{1}{v(z)} &= z \frac{\partial}{\partial z} F_\infty(z) \end{aligned} \quad (9.82)$$

The properties $P > 0$ and $\partial P/\partial v < 0$ are maintained in the thermodynamic limit. We illustrate some possible behaviors of the equation of state.

Suppose the region R includes the entire positive z axis. Then the system always exists in a single phase. The equation of state may be obtained graphically by eliminating z . The situation is illustrated in Fig. 9.5.

If on the other hand a zero of the grand partition function approaches a point z_0 on the real positive z axis, then there will be two distinct regions R_1 and R_2 in which theorem 2 holds separately. At $z = z_0$, $P(z)$ must be continuous, as required by theorem 1. However, its derivative may be discontinuous. The system then possesses two phases, corresponding respectively to the regions $z < z_0$ and $z > z_0$. Now $1/v(z)$ is a nondecreasing function of z :

$$z \frac{\partial}{\partial z} \frac{1}{v(z)} = z \frac{\partial}{\partial z} z \frac{\partial}{\partial z} V^{-1} \log \mathcal{Q}(z, V) = (n^2) - \langle n \rangle^2 \geq 0 \quad (9.83)$$

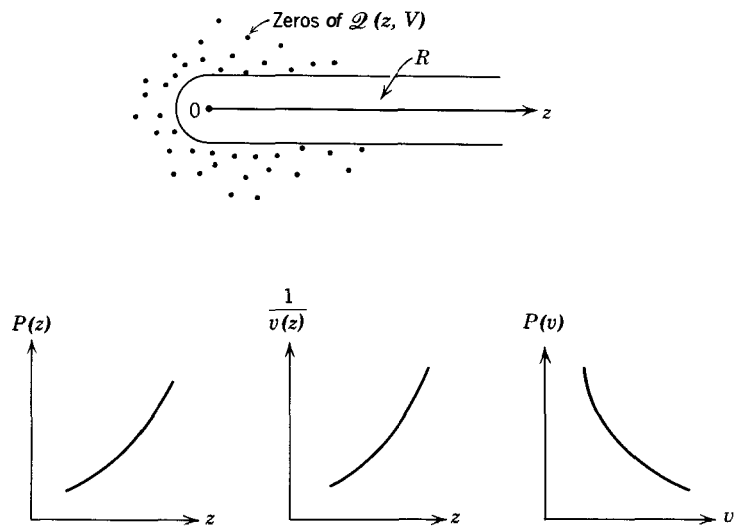


Fig. 9.5 Region R that is free of zeros of $\mathcal{Q}(z, V)$, leading to an equation of state that exhibits only a single phase.

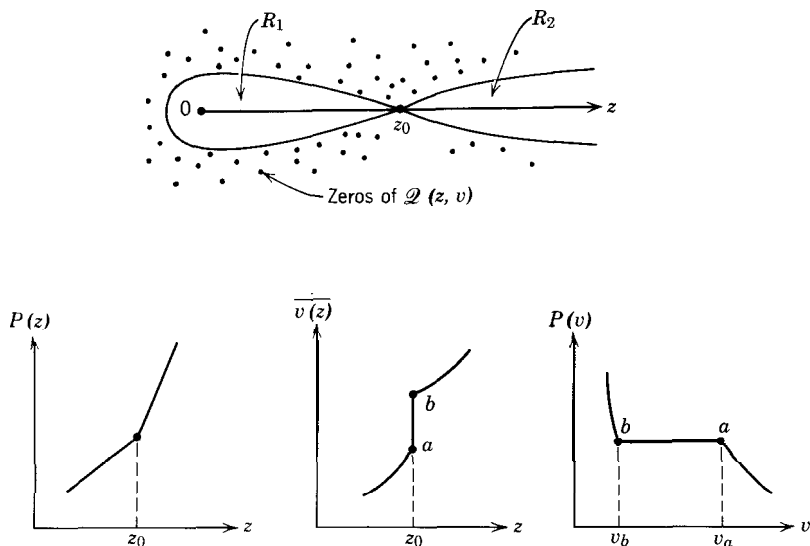


Fig. 9.6 Two regions R_1, R_2 each free of zeros of $\mathcal{Q}(z, V)$, corresponding to an equation of state with two phases connected by a first-order phase transition.

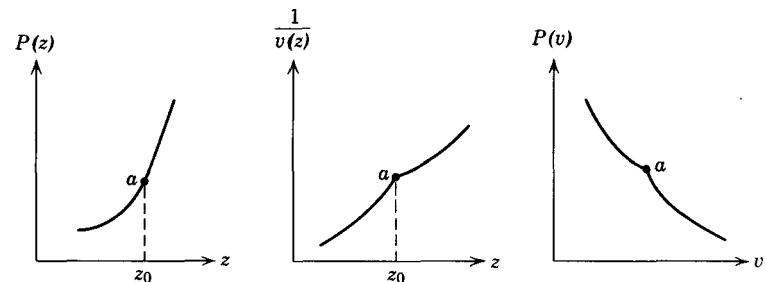


Fig. 9.7 Equation of state of system with two phases connected by a second-order transition.

Hence, if $\partial P/\partial z$ is discontinuous, $1/v(z)$ must make a discontinuous upward jump when z increases through z_0 . The result is a first-order phase transition, as depicted in Fig. 9.6. The fact that $1/v(z)$ actually assumes all the values between the points a and b follows from the fact that the curve of $1/v(z)$ is the limiting curve of a sequence of continuous curves.

If in the same example $\partial P/\partial z$ is continuous at $z = z_0$, but $\partial^2 P/\partial z^2$ is discontinuous, then we would have a second-order phase transition, as illustrated in Fig. 9.7.

9.4 THE LEE-YANG CIRCLE THEOREM

The scenario for the occurrence of a phase transition proposed above can be explicitly demonstrated in the case of a lattice gas with attractive interactions. The system consists of point atoms located on the sites of a discrete lattice, with the condition that no two atoms occupy the same site, and that atoms on different sites have attractive pairwise interactions. That is, the interatomic potential u is such that

$$u = \infty, \quad (\text{if the two atoms are on the same site})$$

$$u < 0, \quad (\text{otherwise})$$

The detailed nature of the interaction (e.g., whether it is of the nearest-neighbor type) is unspecified, as is the dimensionality or structure of the lattice. For example, the lattice does not even have to be periodic.

The Lee-Yang circle theorem states that, for the lattice gas defined above, all roots of the equation

$$\mathcal{Q}(z, v) = 0$$

lie on the unit circle in the complex z plane.

For finite V the roots occur in complex-conjugate pairs, and none can touch the real axis. Only in the limit $V \rightarrow \infty$ is it possible for any of them to approach the real axis at $z = 1$. Thus, there can be at most one transition point.

Denoting the phase angles of the roots by θ_k , we have

$$\mathcal{Q}(z, V) = \mathcal{C} \prod_k (z - e^{i\theta_k}), \quad \mathcal{C} = \prod_k (-e^{i\theta_k})^{-1} \quad (9.84)$$

In the thermodynamic limit the roots become continuously distributed on the unit circle. We can define a distribution function $g(\theta)$ by writing

$$V^{-1} \sum_k \rightarrow \int_0^{2\pi} d\theta g(\theta) \quad (9.85)$$

The fact that the roots occur in complex conjugate pairs means that $g(\theta) = g(-\theta)$. Using this fact, we can deduce the equation of state in the form

$$\begin{aligned} \frac{P}{kT} &= \int_0^\pi d\theta g(\theta) \log(1 - 2z \cos \theta + z^2) \\ \frac{1}{v} &= 2z \int_0^\pi d\theta g(\theta) \frac{z - \cos \theta}{1 - 2z \cos \theta + z^2} \end{aligned} \quad (9.86)$$

We can see from this general form that, as z varies along the real axis, the only point where singularities can occur is $z = 1$. A singularity at this point will occur only if the integrals diverge at $\theta = 0$. Therefore, no phase transitions occur if $g(0) = 0$. On the other hand, if $g(0) \neq 0$, then $z = 1$ will be a singular point, and (9.86) will give different functions for $z > 1$ and for $z < 1$, which cannot be analytically continued into each other.

PROBLEMS

9.1 Derive with the help of the saddle point integration method a formula for the partition function for an ideal Bose gas of N particles.

9.2 (a) Find the equations of state for an ideal Bose gas and an ideal Fermi gas in the limit of high temperatures. Include the first correction due to quantum effects. (Consultation with Problem 7.6 may be helpful.)

(b) Estimate, for each of the following ideal gases, the temperature below which quantum effects would become important: H_2 , He, N_2 .

9.3 Pair Correlation Function. The pair correlation function $D(\mathbf{r}_1, \mathbf{r}_2)$ of a system of particles is defined as follows:

$$D(\mathbf{r}_1, \mathbf{r}_2) d^3r_1 d^3r_2 \equiv \text{probability of simultaneously finding a particle in the volume element } d^3r_1 \text{ about } \mathbf{r}_1 \text{ and a particle in the volume element } d^3r_2 \text{ and } \mathbf{r}_2.$$

Calculate $D(\mathbf{r}_1, \mathbf{r}_2)$ for an ideal Bose gas and an ideal Fermi gas in the limit of high temperatures. Include quantum corrections only to the lowest approximation.

Solution. Classically we have

$$D(\mathbf{r}_1, \mathbf{r}_2) = \frac{N(N-1) \int d^3N p d^3r_3 \dots d^3r_N e^{-\beta \mathcal{H}(p, r)}}{\int d^3N p d^3N r e^{-\beta \mathcal{H}(p, r)}}$$

For our problem we use this formula with

$$\mathcal{H}(p, r) = \sum_{i=j}^N \frac{p_i^2}{2m} + \sum_{i<j} \tilde{v}_{ij}$$

To avoid complications assume that the density of the gas is almost zero. The limit $N \rightarrow \infty, V \rightarrow \infty$ should be so taken that $N/V \rightarrow 0$. Then

$$\begin{aligned} D(\mathbf{r}_1, \mathbf{r}_2) &\approx \frac{N(N-1) V^{N-2} \left[1 \pm f_{12}^2 \pm \frac{N(N-1)}{2V} \int d^3r f^2(r) \right]}{1 \pm \frac{N(N-1)}{2V} \int d^3r f^2(r)} \\ &\approx \frac{1}{v^2} \left[1 \pm \exp\left(-\frac{2\pi}{\lambda^2} |\mathbf{r}_1 - \mathbf{r}_2|^2\right) \right] \end{aligned}$$

This result continues to hold for finite v with $\lambda^3/v \ll 1$, although our derivation did not justify such a conclusion.

9.4 Show that the equation of state (9.86) of the Lee-Yang lattice gas has the following electrostatic analog:

(a) Consider a circular cylinder of unit radius perpendicular to the complex z plane, cutting it at the unit circle. Suppose the cylinder is charged with a surface charge density that depends only on the angle θ around the unit circle (with $\theta = 0$ corresponding to $z = 1$). The charge density (per unit area) is equal to $g(\theta)$, with $g(\theta) = g(-\theta)$. Let $\phi(z)$ and $E(z)$ be, respectively, the electrostatic potential and the electric field at a point z on the real axis. Then

$$\begin{aligned} P/kT &= -\frac{1}{2} \phi(z) \\ n &= \frac{1}{2} z E(z) \end{aligned}$$

where $n = 1/v$ is the density.

(b) Assume $g(0) \neq 0$. Show by electrostatic argument that P is continuous at $z = 1$, but n jumps discontinuously. This shows that there is a first-order phase transition. Using Gauss' theorem in electrostatics, show the discontinuity in density is given by $\Delta n = 2\pi g(0)$.

9.5 Consider the grand partition function

$$\mathcal{Q}(z, V) = (1 + z)^V (1 + z^a)^V$$

where a is a positive constant.

(a) Write down the equation of state in a parametric form, eliminate z graphically, and show that there is a first-order phase transition. Find the specific volumes of the two phases.

(b) Find the roots of $\mathcal{Q}(z, V) = 0$ in the complex z plane, at fixed V . Show that as $V \rightarrow \infty$ the roots converge toward the real axis at $z = 1$.

(c) Find the equation of state in the "gas" phase. Show that a continuation of this equation beyond the phase-transition density fails to show any sign of the transition. This will demonstrate that the order of the operations $z(\partial/\partial z)$ and $V \rightarrow \infty$ can be interchanged only within a single-phase region.

CHAPTER 10

APPROXIMATE METHODS

10.1 CLASSICAL CLUSTER EXPANSION

Many systems of physical interest can be treated classically. A large class of such systems is described by a classical Hamiltonian for N particles of the form

$$\mathcal{H} = \sum_{i=1}^N \frac{p_i^2}{2m} + \sum_{i<j} v_{ij} \quad (10.1)$$

where p_i is the momentum of the i th particle and $v_{ij} = v(|\mathbf{r}_i - \mathbf{r}_j|)$ is the potential energy of interaction between the i th and the j th particle. If the system occupies a volume V , the partition function is

$$Q_N(V, T) = \frac{1}{N! h^{3N}} \int d^{3N}p d^{3N}r \exp\left(-\beta \sum_i \frac{p_i^2}{2m} - \beta \sum_{i<j} v_{ij}\right) \quad (10.2)$$

where each coordinate \mathbf{r}_i is integrated over the volume V . The integrations over momenta can be immediately effected, leading to

$$Q_N(V, T) = \frac{1}{\lambda^{3N} N!} \int d^{3N}r \exp\left(-\beta \sum_{i<j} v_{ij}\right) \quad (10.3)$$

where $\lambda = \sqrt{2\pi\hbar^2/mkT}$ is the thermal wavelength. The integral in (10.3) is called the *configuration integral*. For potentials u_{ij} of the usual type between molecules, a systematic method for the calculation of the configuration integral consists of expanding the integrand in powers of $\exp(-\beta v_{ij}) - 1$. This leads to the cluster expansion of Ursell and Mayer.* As we shall see, this expansion is of practical use if the system is a dilute gas.

*For original literature, see J. E. Mayer and M. G. Mayer, *Statistical Mechanics* (Wiley, New York, 1940), Chapter 13.

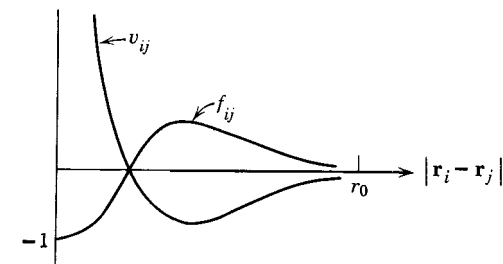


Fig. 10.1 Intermolecular potential v_{ij} and the function f_{ij} .

Let the configuration integral be denoted by $Z_N(V, T)$:

$$Z_N(V, T) \equiv \int d^3r_1 \cdots d^3r_N \exp\left(-\beta \sum_{i<j} v_{ij}\right) \quad (10.4)$$

in terms of which the partition function may be written as

$$Q_N(V, T) = \frac{1}{N! \lambda^{3N}} Z_N(V, T) \quad (10.5)$$

and the grand partition function as

$$\mathcal{Q}(z, V, T) = \sum_{N=0}^{\infty} \left(\frac{z}{\lambda^3}\right)^N \frac{Z_N(V, T)}{N!} \quad (10.6)$$

Let f_{ij} be defined by

$$e^{-\beta v_{ij}} \equiv 1 + f_{ij} \quad (10.7)$$

For the usual type of intermolecular potentials, u_{ij} and f_{ij} have the qualitative forms shown in Fig. 10.1. Thus f_{ij} is everywhere bounded and is negligibly small when $|\mathbf{r}_i - \mathbf{r}_j|$ is larger than the range of the intermolecular potential. In terms of f_{ij} the configuration integral may be represented by

$$Z_N(V, T) = \int d^3r_1 \cdots d^3r_N \prod_{i<j} (1 + f_{ij}) \quad (10.8)$$

in which the integrand is a product of $\frac{1}{2}N(N-1)$ terms, one for each distinct pair of particles. Expanding this product we obtain

$$Z_N(V, T) = \int d^3r_1 \cdots d^3r_N [1 + (f_{12} + f_{13} + \cdots) + (f_{12}f_{13} + f_{12}f_{14} + \cdots) + \cdots] \quad (10.9)$$

A convenient way to enumerate all the terms in the expansion (10.9) is to associate each term with a graph, defined as follows:

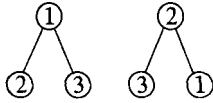
An N -particle graph is a collection of N distinct circles numbered $1, 2, \dots, N$, with any number of lines joining the same number of distinct pairs of circles. If the

distinct pairs joined by lines are the pairs a, β, \dots, A , then the graph represents the term

$$\int d^3r_1 \cdots d^3r_N f_a f_\beta \cdots f_\gamma \quad (10.10)$$

appearing in the expansion (10.9).

If the set of distinct pairs $\{a, \beta, \dots, \gamma\}$ is joined by lines in a given graph, replacing this set by a set $\{a', \beta', \dots, \gamma'\}$ that is not identical with $\{a, \beta, \dots, \gamma\}$ gives rise to a graph that is counted as distinct from the original one (although the integrals represented by the respective graphs have the same numerical value). For example, for $N = 3$, the following graphs are distinct:



but the following graphs are identical:



We may regard a graph as a picturesque way of writing the integral (10.10). For example, we may write, for $N = 10$,

$$\left[\begin{array}{ccc} \textcircled{1} & \textcircled{3}-\textcircled{9} & \textcircled{6}-\textcircled{8} \\ | & & / \backslash \\ \textcircled{2} & \textcircled{4} & \textcircled{5} & \textcircled{7} & \textcircled{10} \end{array} \right] = \int d^3r_1 \cdots d^3r_{10} f_{12} f_{39} f_{67} f_{68} f_{8,10} f_{6,10} f_{78} \quad (10.11)$$

With such a convention, we can state that

$$Z_N = (\text{sum of all distinct } N\text{-particle graphs}) \quad (10.12)$$

The proof is obvious.

Any graph can in general be decomposed into smaller units. For example, the graph (14.11) is a product of five factors, namely

$$\left[\begin{array}{ccc} \textcircled{1} & \textcircled{3}-\textcircled{9} & \textcircled{6}-\textcircled{8} \\ | & & / \backslash \\ \textcircled{2} & \textcircled{4} & \textcircled{5} & \textcircled{7} & \textcircled{10} \end{array} \right] = [\textcircled{4}] \cdot [\textcircled{5}] \cdot [\textcircled{1}-\textcircled{2}] \cdot [\textcircled{3}-\textcircled{9}] \cdot \left[\begin{array}{cc} \textcircled{6}-\textcircled{8} \\ / \backslash \\ \textcircled{7} & \textcircled{10} \end{array} \right]$$

Each factor corresponds to a connected graph, in which every circle is attached to at least one line, and every circle is joined directly or indirectly to all other circles in the graph.

It would facilitate the analysis of Z_N if we first defined the basic units out of which an arbitrary graph can be composed. Accordingly we define an 1-cluster to

be an 1 particle connected graph. For example, the following is a 6-cluster:

$$= \int d^3r_1 \cdots d^3r_6 f_{12} f_{23} f_{14} f_{46} f_{56} \quad (10.13)$$

We define a cluster integral $b_l(V, T)$ by

$$b_l(V, T) \equiv \frac{1}{l! \lambda^{3l-3V}} (\text{sum of all possible } l\text{-clusters}) \quad (10.14)$$

The normalization factor is so chosen that

- (a) $b_l(V, T)$ is dimensionless;
- (b) $\bar{b}_l(T) \equiv \lim_{V \rightarrow \infty} b_l(V, T)$ is a finite number.

The property (b) follows from the fact that f_{ij} has a finite range, so that in an 1-cluster the only integration that gives rise to a factor V is the integration over the "center of gravity" of the 1 particles. Some of the cluster integrals are

$$b_1 = \frac{1}{V} [\textcircled{1}] = \frac{1}{V} \int d^3r_1 = 1 \quad (10.15)$$

$$b_2 = \frac{1}{2! \lambda^3 V} [\textcircled{1}-\textcircled{2}] = \frac{1}{\lambda^3 2V} \int d^3r_1 d^3r_2 f_{12} = \frac{1}{2\lambda^3} \int d^3r_{12} f_{12} \quad (10.16)$$

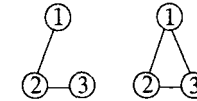
$$b_3 = \frac{1}{3! \lambda^6 V} \left[\begin{array}{c} \textcircled{1} \\ / \backslash \\ \textcircled{2}-\textcircled{3} \end{array} + \begin{array}{c} \textcircled{1} \\ | \\ \textcircled{2}-\textcircled{3} \end{array} + \begin{array}{c} \textcircled{1} \\ | \\ \textcircled{2}-\textcircled{3} \end{array} + \begin{array}{c} \textcircled{1} \\ / \backslash \\ \textcircled{2}-\textcircled{3} \end{array} \right] \quad (10.17)$$

Any N -particle graph is a product of a number of clusters, of which m_l are l -clusters, with

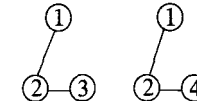
$$\sum_{l=1}^N l m_l = N \quad (10.18)$$

A given set of integers (m_l) satisfying (10.18), however, does not uniquely specify a graph, because

- (a) there are in general many ways to form an 1-cluster, e.g.,



- (b) there are in general many ways to assign which particle belongs to which cluster, e.g.,



Thus a set of integers $\{m_l\}$ specifies a collection of graphs. Let the sum of all the

graphs corresponding to $\{m_i\}$ be denoted by $S\{m_i\}$. Then

$$Z_N = \sum_{\{m_i\}} S\{m_i\} \tag{10.19}$$

where the summation extends over all sets $\{m_i\}$ satisfying (10.18).

By definition, $S\{m_i\}$ can be obtained as follows. First write down an arbitrary N-particle graph that contains m_1 1-clusters, m_2 2-clusters, etc.; e.g.,

$$\frac{\{[\bigcirc] \dots [\bigcirc]\} \{[\bigcirc-\bigcirc] \dots [\bigcirc-\bigcirc]\}}{m_1 \text{ factors} \quad m_2 \text{ factors}} \times \underbrace{\left\{ \left[\begin{array}{c} \bigcirc \\ / \quad \backslash \\ \bigcirc-\bigcirc \end{array} \right] \left[\begin{array}{c} \bigcirc \\ / \quad \backslash \\ \bigcirc-\bigcirc \end{array} \right] \left[\begin{array}{c} \bigcirc \\ / \quad \backslash \\ \bigcirc-\bigcirc \end{array} \right] \dots \left[\begin{array}{c} \bigcirc \\ / \quad \backslash \\ \bigcirc-\bigcirc \end{array} \right] \right\}}_{m_3 \text{ factors}} \dots \tag{10.20}$$

There are exactly N circles appearing in (10.20), and these N circles are to be filled in by the numbers 1, 2, ..., N in an arbitrary but definite order. We can write down many more examples like (10.20); e.g., we may change the choice of some of the 3-clusters (there being four distinct topological shapes for a 3-cluster). Again we may permute the numbering of all the N circles in (10.20), and that would lead to a distinct graph. If we add up all these possibilities, we obtain $S\{m_i\}$. Thus we may write

$$S\{m_i\} = \sum_P [\bigcirc]^{m_1} [\bigcirc-\bigcirc]^{m_2} \times \left[\begin{array}{c} \bigcirc \\ / \quad \backslash \\ \bigcirc-\bigcirc \end{array} + \begin{array}{c} \bigcirc \\ / \quad \backslash \\ \bigcirc-\bigcirc \end{array} + \begin{array}{c} \bigcirc \\ / \quad \backslash \\ \bigcirc-\bigcirc \end{array} + \begin{array}{c} \bigcirc \\ / \quad \backslash \\ \bigcirc-\bigcirc \end{array} \right]^{m_3} \left[\dots \right]^{m_4} \dots \tag{10.21}$$

The meaning of this formula is as follows. Each bracket contains the sum over all I-clusters. If all the brackets $[\dots]^{m_i}$ are expanded in multinomial expansions, the summand of \sum_P will itself be a sum of a large number of terms in which every term contains exactly N circles. The sum \sum_P extends over all distinct ways of numbering these circles from 1 to N.

Now each graph is an integral whose value is independent of the way its circles are numbered. Therefore $S\{m_i\}$ is equal to the number of terms in the sum \sum_P times the value of any term in the sum. The number of terms in the sum \sum_P can be found by observing that

- (a) there are m_i I-clusters, and a permutation of these m_i things does not lead to a new graph;
- (b) in the sum over all I-clusters, such as (10.17), a permutation of the I particles within it does not lead to a new graph. Hence the number of

terms in the sum \sum_P is*

$$\frac{N!}{[(1!)^{m_1} (2!)^{m_2} \dots] [m_1! m_2! \dots]} \tag{10.22}$$

and the value of any term is

$$(1!Vb_1)^{m_1} (2!\lambda^3 Vb_2)^{m_2} (3!\lambda^6 Vb_3)^{m_3} \dots \tag{10.23}$$

Therefore

$$S\{m_i\} = N! \prod_{i=1}^N \frac{(V\lambda^{3i-3} b_i)^{m_i}}{m_i!} = N! \lambda^{3N} \prod_{i=1}^N \frac{1}{m_i!} (\lambda^3 b_i)^{m_i} \tag{10.24}$$

From (10.5), (10.9), and (10.24) we obtain

$$Q_N(V, T) = \sum_{\{m_i\}} \prod_{i=1}^N \frac{1}{m_i!} \left(\frac{V}{\lambda^3} b_i \right)^{m_i} \tag{10.25}$$

This formula is complicated by the restriction (10.18). The grand partition function is simpler in appearance:

$$\mathcal{Q}(z, V, T) = \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} \dots \left[\frac{1}{m_1!} \left(\frac{V}{\lambda^3} z b_1 \right)^{m_1} \frac{1}{m_2!} \left(\frac{V}{\lambda^3} z^2 b_2 \right)^{m_2} \dots \right]$$

or

$$\frac{1}{V} \log \mathcal{Q}(z, V, T) = \frac{1}{\lambda^3} \sum_{i=1}^{\infty} b_i z^i \tag{10.26}$$

from which we obtain the equation of state in parametric form:

$$\begin{cases} \frac{P}{kT} = \frac{1}{\lambda^3} \sum_{i=1}^{\infty} b_i z^i \\ \frac{1}{v} = \frac{1}{\lambda^3} \sum_{i=1}^{\infty} i b_i z^i \end{cases} \tag{10.27}$$

This is known as the cluster expansion for the equation of state.†

What we have described is historically the first graphical representation of a perturbation series. Graphs have become indispensable tools in the many-body problem and in quantum field theory, in which the analog of (10.26), known generally by the name of the *linked cluster theorem*, plays an important role. Generally it states that the sum of all graphs is the exponential of the sum of all connected graphs.

*To understand the method of counting the reader is advised to work out some simple examples.

†Compare this derivation with that outlined in Problem 7.6.

If the system under consideration is a dilute gas, we may expand the pressure in powers of $1/v$ and obtain the virial expansion. For this purpose we may take the equation of state to be

$$\begin{cases} \frac{P}{kT} = \frac{1}{\lambda^3} \sum_{l=1}^{\infty} \bar{b}_l z^l \\ \frac{1}{v} = \frac{1}{\lambda^3} \sum_{l=1}^{\infty} l \bar{b}_l z^l \end{cases} \quad (10.28)$$

where

$$\bar{b}_l(T) \equiv \lim_{V \rightarrow \infty} b_l(V, T) \quad (10.29)$$

The virial expansion of the equation of state is defined to be

$$\frac{Pv}{kT} = \sum_{l=1}^{\infty} a_l(T) \left(\frac{\lambda^3}{v} \right)^{l-1} \quad (10.30)$$

where $a_l(T)$ is called the l th virial coefficient. We can find the relationship between the virial coefficients a_l and the cluster integrals \bar{b}_l by substituting (10.30) into (10.28) and requiring that the resulting equation be satisfied for every z :

$$\sum_{l=1}^{\infty} a_l \left(\sum_{n=1}^{\infty} n \bar{b}_n z^n \right)^{l-1} = \frac{\sum_{l=1}^{\infty} \bar{b}_l z^l}{\sum_{l=1}^{\infty} l \bar{b}_l z^l} \quad (10.31)$$

This is equivalent to the condition

$$\begin{aligned} & (\bar{b}_1 z + 2\bar{b}_2 z^2 + 3\bar{b}_3 z^3 + \dots) \left[a_1 + a_2 \left(\sum_{n=1}^{\infty} n \bar{b}_n z^n \right) + a_3 \left(\sum_{n=1}^{\infty} n \bar{b}_n z^n \right)^2 + \dots \right] \\ & = \bar{b}_1 z + \bar{b}_2 z^2 + \bar{b}_3 z^3 + \dots \end{aligned} \quad (10.32)$$

By equating the coefficient of each power of z we obtain

$$\begin{aligned} a_1 &= \bar{b}_1 = 1 \\ a_2 &= -\bar{b}_2 \\ a_3 &= 4\bar{b}_2^2 - 2\bar{b}_3 \\ a_4 &= -20\bar{b}_2^3 + 18\bar{b}_2\bar{b}_3 - 3\bar{b}_4 \\ &\dots \end{aligned} \quad (10.33)$$

Each virial coefficient therefore involves only a straightforward computation of a number of integrals.

Note that (10.28) differs from (10.27) in that the limit $V \rightarrow \infty$ is taken term by term in (10.28). In so doing we have lost all information about possible phase transitions, as we have remarked earlier in Section 9.3. The equation of state

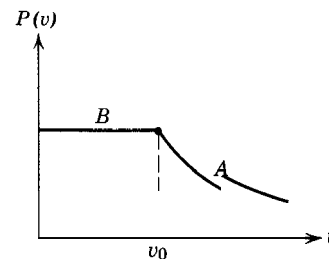


Fig. 10.2 Equation of state obtained by taking the virial expansion to be exact.

(10.30) of the gas phase cannot tell us if and when a phase transition will occur. Mayer* has demonstrated that the equation of state (10.30) has the general form shown in Fig. 10.2. The portion of the isotherm marked A is valid for $v > v_1$, but the value of v_1 is unrelated to v_0 , and cannot be obtained from (10.30). The portion marked B is purely mathematical, and unrelated to how the isotherm actually behaves in that region.

10.2 QUANTUM CLUSTER EXPANSION

Kahn and Uhlenbeck† develop a cluster expansion in quantum statistical mechanics. The method they introduce applies equally well to classical statistical mechanics.

Consider N identical particles enclosed in a volume V . Let the Hamiltonian \mathcal{H} of the system have the same form as (10.1) but be an operator instead of a number. In the coordinate representation, $\mathbf{p}_j = -i\hbar\nabla_j$, and v_{ij} is the same function of the number $|\mathbf{r}_i - \mathbf{r}_j|$ as that shown in Fig. 10.1. The partition function is

$$Q_N(V, T) = \text{Tr} e^{-\beta H} = \int d^{3N}r \sum_{\alpha} \Psi_{\alpha}^*(1, \dots, N) e^{-\beta \mathcal{H}} \Psi_{\alpha}(1, \dots, N) \quad (10.34)$$

where $\{\Psi_{\alpha}\}$ is a complete set of orthonormal wave functions appropriate to the system considered, and the set of coordinates $\{\mathbf{r}_1, \dots, \mathbf{r}_N\}$ is denoted in abbreviation by $\{1, \dots, N\}$. It is important to use symmetric or antisymmetric wave functions, as required by the statistics of the particles (see Problem 10.4). Let us define

$$W_N(1, \dots, N) \equiv N! \lambda^{3N} \sum_{\alpha} \Psi_{\alpha}^*(1, \dots, N) e^{-\beta \mathcal{H}} \Psi_{\alpha}(1, \dots, N) \quad (10.35)$$

The partition function can be written in the form

$$Q_N(V, T) = \frac{1}{N! \lambda^{3N}} \int d^{3N}r W_N(1, \dots, N) \quad (10.36)$$

*See Mayer and Mayer, *loc. cit.*

†B. Kahn and G. E. Uhlenbeck, *Physica* 5, 399 (1938).

The integral appearing in (10.36) approaches the classical configuration integral in the limit of high temperatures. Some properties of the function $W_N(1, \dots, N)$ are

$$(a) W_1(1) = 1$$

Proof

$$\begin{aligned} W_1(1) &= W_1(\mathbf{r}_1) = \frac{\lambda^3}{V} \sum_{\mathbf{p}} e^{-i\mathbf{p} \cdot \mathbf{r}_1 / \hbar} e^{(\beta \hbar^2 / 2m) \nabla^2} e^{i\mathbf{p} \cdot \mathbf{r}_1 / \hbar} \\ &= \left(\frac{\lambda}{\hbar} \right)^3 \int d^3 p e^{-\beta p^2 / 2m} = 1 \quad \blacksquare \end{aligned}$$

(b) $W_N(1, \dots, N)$ is a symmetric function of its arguments.

(c) $W_N(1, \dots, N)$ is invariant under a unitary transformation of the complete set of wave functions $\{\Psi_\alpha\}$ appearing in (10.35).

Proof

$$\begin{aligned} \text{Suppose } \Psi_\alpha &= \sum_A S_{\alpha\lambda} \Phi_\lambda, \text{ where } S_{\alpha\lambda} \text{ is a unitary matrix:} \\ &\sum_a S_{\alpha\lambda}^* S_{\alpha\gamma} = \delta_{\lambda\gamma} \end{aligned}$$

Then

$$\sum_o (\Psi_\alpha, e^{-\beta \mathcal{H}} \Psi_\alpha) = \sum_{o, A} S_{\alpha\lambda}^* S_{\alpha\gamma} (\Phi_\lambda, e^{-\beta \mathcal{H}} \Phi_\gamma) = \sum_A (\Phi_\lambda, e^{-\beta \mathcal{H}} \Phi_\lambda) \quad \blacksquare$$

The following property appears to be intuitively obvious, but it is difficult to establish quantitatively. Suppose the coordinates $\mathbf{r}_1, \dots, \mathbf{r}_N$ have such values that they can be divided into two groups containing respectively A and B coordinates, with the property that any two coordinates \mathbf{r}_i and \mathbf{r}_j , belonging to different groups must satisfy the condition

$$\begin{aligned} |\mathbf{r}_i - \mathbf{r}_j| &\gg r_0 \\ |\mathbf{r}_i - \mathbf{r}_j| &\gg \lambda \end{aligned} \quad (10.37)$$

Then

$$W_N(\mathbf{r}_1, \dots, \mathbf{r}_N) \approx W_A(r_A) W_B(r_B) \quad (10.38)$$

where r_A and r_B denote collectively the respective coordinates in the two groups.

Consider first the case $N = 2$. According to (10.38) we should expect that as $|\mathbf{r}_1 - \mathbf{r}_2| \rightarrow \infty$,

$$W_2(1, 2) \rightarrow W_1(1) W_2(2) \quad (10.39)$$

If we define a function $U_2(1, 2)$ by $W_2(1, 2) = W_1(1) W_2(2) + U_2(1, 2)$, we should expect that, as $|\mathbf{r}_1 - \mathbf{r}_2| \rightarrow \infty$,

$$U_2(1, 2) \rightarrow 0 \quad (10.40)$$

Hence the integral of $U_2(1, 2)$ over \mathbf{r}_1 and \mathbf{r}_2 should be the analog of the 2-cluster in classical statistical mechanics.

We proceed systematically in the following manner. Let a sequence of cluster functions $U_l(1, \dots, l)$ be successively defined by the following scheme, in which the l th equation is a definition of $U_l(1, \dots, l)$:

$$W_1(1) = U_1(1) = 1 \quad (10.41)$$

$$W_2(1, 2) = U_1(1) U_1(2) + U_2(1, 2) \quad (10.42)$$

$$\begin{aligned} W_3(1, 2, 3) &= U_1(1) U_1(2) U_1(3) + U_1(1) U_2(2, 3) \\ &\quad + U_1(2) U_2(3, 1) + U_1(3) U_2(1, 2) + U_3(1, 2, 3) \end{aligned} \quad (10.43)$$

The last equation in this scheme, defining $U_N(1, \dots, N)$, is

$$\begin{aligned} W_N(1, \dots, N) &= \sum_{\{m_l\}} \sum_P \underbrace{[U_1(\) \cdots U_1(\)]}_{m_1 \text{ factors}} \underbrace{[U_2(\) \cdots U_2(\)]}_{m_2 \text{ factors}} \cdots \underbrace{[U_N(\) \cdots U_N(\)]}_{m_N \text{ factors}} \end{aligned} \quad (10.44)$$

where m_l is zero or a positive integer and the set of integers $\{m_l\}$ satisfies the condition

$$\sum_{l=1}^N l m_l = N \quad (10.45)$$

The sum over $\{m_l\}$ in (10.44) extends over all sets $\{m_l\}$ satisfying (10.45). The arguments of the U_l are left blank in (10.44). There are exactly N such blanks, and they are to be filled by the N coordinates $\mathbf{r}_1, \dots, \mathbf{r}_N$ in any order. The sum \sum_P is a sum over all *distinct* ways of filling these blanks.

We can solve the equations (10.41)–(10.44) successively for U_1, U_2 , etc., and obtain

$$U_1(1) = W_1(1) = 1 \quad (10.46)$$

$$U_2(1, 2) = W_2(1, 2) - W_1(1) W_1(2) \quad (10.47)$$

$$\begin{aligned} U_3(1, 2, 3) &= W_3(1, 2, 3) - W_2(1, 2) W_1(3) - W_2(2, 3) W_1(1) \\ &\quad - W_2(3, 1) W_1(2) + 2 W_1(1) W_1(2) W_1(3) \end{aligned} \quad (10.48)$$

We see that $U_l(1, \dots, l)$ is a symmetric function if its arguments and is determined by all the $W_{N'}$ with $N' \leq l$. By the property (10.38) we expect that $U_l \rightarrow 0$ as $|\mathbf{r}_i - \mathbf{r}_j| \rightarrow \infty$, where r_i and r_j are any two of the arguments of U_l .

The l cluster integral $b_l(V, T)$ is defined by

$$b_l(V, T) \equiv \frac{1}{l! \lambda^{3l-3V}} \int d^3 r_1 \cdots d^3 r_l U_l(1, \dots, l) \quad (10.49)$$

It is clear that b_l is dimensionless. If U_l vanishes sufficiently rapidly whenever any two of its arguments are far apart from each other, the integral appearing in (10.49) is proportional to V as $V \rightarrow \infty$, and the limit $b_l(\infty, T)$ may be expected

to exist. Whether this is true depends on the nature of the interparticle potential. We assume that it is.

We now show that the partition function is expressible directly in terms of the cluster integrals. According to (10.36) we need to integrate W_N over all the coordinates. Let us make use of the formula (10.44). An integration over all the coordinates will yield the same result for every term in the sum \sum . Thus the result of the integration is the number of terms in the sum \sum times \int the integral of any term in the sum \sum . The number of terms in the sum \sum is given by (10.22). Hence

$$\begin{aligned} & \int d^{3N}r W(1, \dots, N) \\ &= \sum_{\{m_i\}} \frac{N!}{[(1!)^{m_1} (2!)^{m_2} \dots] (m_1! m_2! \dots)} \\ & \quad \times \int d^{3N}r [(U_1 \dots U_1)(U_2 \dots U_2) \dots] \\ &= N! \sum_{\{m_i\}} \frac{m_1!}{m_1!} \left[\frac{1}{1!} \int d^3r_1 U_1(1) \right]^{m_1} \frac{1}{m_2!} \left[\frac{1}{2!} \int d^3r_1 d^3r_2 U_2(1, 2) \right]^{m_2} \dots \\ &= N! \sum_{\{m_i\}} \prod_{l=1}^N \frac{(V \lambda^{3l-3} b_l)^{m_l}}{m_l!} = N! \lambda^{3N} \sum_{\{m_i\}} \prod_{l=1}^N \frac{1}{m_l!} \left(\frac{V}{\lambda^3} b_l \right)^{m_l} \end{aligned} \quad (10.50)$$

Therefore the partition function is given by

$$Q_N(V, T) = \sum_{\{m_i\}} \prod_{l=1}^N \frac{1}{m_l!} \left(\frac{V}{\lambda^3} b_l \right)^{m_l} \quad (10.51)$$

This is of precisely the same form as (10.25) for the classical partition function. The discussion following (10.25) therefore applies equally well to the present case and will not be repeated. We point out only the main differences between the quantum cluster integrals and the classical ones.

For an ideal gas we have seen in earlier chapters that

$$b_l^{(0)} = \begin{cases} l^{-5/2} & \text{(ideal Bose gas)} \\ (-1)^{l+1} l^{-5/2} & \text{(ideal Fermi gas)} \end{cases} \quad (10.52)$$

Thus for a Bose and a Fermi gas b_l does not vanish for $l > 1$, even in the absence of interparticle interactions, in contradistinction to the classical ideal gas.

The calculation of b_l in the classical case only involves the calculation of a number of integrals—a finite task. In the quantum case, however, the calculation of b_l necessitates a knowledge of U_l , which in turn necessitates a knowledge of $W_{N'}$ for $N' \leq l$. Thus to find b_l for $l > 1$ we would have to solve an l -body problem. There is no finite prescription for doing this except for the case $l = 2$, which is the subject of the next section.

10.3 THE SECOND VIRIAL COEFFICIENT

To calculate the second virial coefficient a , for any system it is sufficient to calculate b_2 , since $a = -b_2$. A general formula for b_2 (in fact, for all b_l) has already been given for the classical case. Only the quantum case is considered here.*

To find b_2 we need to know $W_2(1, 2)$, which is a property of the two-body system. Let the Hamiltonian for the two-body system in question be

$$\mathcal{H} = -\frac{\hbar^2}{2m} (\nabla_1^2 + \nabla_2^2) + v(|\mathbf{r}_1 - \mathbf{r}_2|) \quad (10.53)$$

and let its normalized eigenfunctions be $\Psi_\alpha(1, 2)$, with eigenvalues E_α :

$$\mathcal{H} \Psi_\alpha(1, 2) = E_\alpha \Psi_\alpha(1, 2) \quad (10.54)$$

Let

$$\begin{aligned} \mathbf{R} &= \frac{1}{2}(\mathbf{r}_1 + \mathbf{r}_2) \\ \mathbf{r} &= \mathbf{r}_2 - \mathbf{r}_1 \end{aligned} \quad (10.55)$$

Then

$$\begin{aligned} \Psi_\alpha(1, 2) &= \frac{1}{\sqrt{V}} e^{i\mathbf{P} \cdot \mathbf{R}} \psi_n(\mathbf{r}) \\ E_\alpha &= \frac{P^2}{4m} + \epsilon_n \end{aligned} \quad (10.56)$$

where the quantum number α refers to the set of quantum numbers (P, n) . The relative wave function $\psi_n(\mathbf{r})$ satisfies the eigenvalue equation

$$\left[-\frac{\hbar^2}{m} \nabla^2 + v(r) \right] \psi_n(\mathbf{r}) = \epsilon_n \psi_n(\mathbf{r}) \quad (10.57)$$

with the normalization condition

$$\int d^3r |\psi_n(\mathbf{r})|^2 = 1 \quad (10.58)$$

Using (10.56) to be the wave functions for the calculation of $W_2(1, 2)$, we find from (10.35) that

$$W_2(1, 2) = 2\lambda^6 \sum_{\alpha} |\Psi_\alpha(1, 2)|^2 e^{-\beta E_\alpha} = \frac{2\lambda^6}{V'} \sum_{\mathbf{P}} \sum_n |\psi_n(\mathbf{r})|^2 e^{-\beta P^2/4m} e^{-\beta \epsilon_n} \quad (10.59)$$

In the limit as $V \rightarrow \infty$ the sum over \mathbf{P} can be effected immediately:

$$\frac{1}{V} \sum_{\mathbf{P}} e^{-\beta P^2/4m} = \frac{4\pi}{h^3} \int_0^\infty dP P^2 e^{-\beta P^2/4m} = \frac{2^{3/2}}{\lambda^3} \quad (10.60)$$

*The following development is due to E. Beth and G. E. Uhlenbeck, *Physica* 4, 915 (1937).

where $\lambda = \sqrt{2\pi\hbar^2/mkT}$, the thermal wavelength. Therefore

$$W_2(1, 2) = 2^{5/2}\lambda^3 \sum_n |\psi_n(\mathbf{r})|^2 e^{-\beta\epsilon_n} \quad (10.61)$$

If we repeat all the calculations so far for a two-body system of noninteracting particles, we obtain

$$W_2^{(0)}(1, 2) = 2^{5/2}\lambda^3 \sum_n |\psi_n^{(0)}(\mathbf{r})|^2 e^{-\beta\epsilon_n^{(0)}} \quad (10.62)$$

where the superscript (0) refers to quantities of the noninteracting system. From (10.49) and (10.47) we have

$$\mathfrak{b}_2 = \frac{1}{2\lambda^3 V} \int d^3r_1 d^3r_2 U_2(1, 2) = \frac{1}{2\lambda^3 V} \int d^3R d^3r [W_2(1, 2) - 1]$$

Hence

$$\begin{aligned} \mathfrak{b}_2 - \mathfrak{b}_2^{(0)} &= \frac{1}{2\lambda^3 V} \int d^3R d^3r [W_2(1, 2) - W_2^{(0)}(1, 2)] \\ &= 2\sqrt{2} \int d^3r \sum_n [|\psi_n(\mathbf{r})|^2 e^{-\beta\epsilon_n} - |\psi_n^{(0)}(\mathbf{r})|^2 e^{-\beta\epsilon_n^{(0)}}] \\ &= 2\sqrt{2} \sum (e^{-\beta\epsilon_n} - e^{-\beta\epsilon_n^{(0)}}) \end{aligned} \quad (10.63)$$

where

$$\mathfrak{b}_2^{(0)} = \begin{cases} 2^{-5/2} & \text{(ideal Bose gas)} \\ -2^{-5/2} & \text{(ideal Fermi gas)} \end{cases} \quad (10.64)$$

To analyze (10.63) further we must study the energy spectra $\epsilon_n^{(0)}$ and ϵ_n . For the noninteracting system, $\epsilon_n^{(0)}$ forms a continuum. We write

$$\epsilon_n^{(0)} = \frac{\hbar^2 k^2}{m} \quad (10.65)$$

which defines the relative wave number k . For the interacting system the spectrum of ϵ_n in general contains a discrete set of values ϵ_B , corresponding to two-body bound states, and a continuum. In the continuum, we define the wave number k for the interacting system by putting

$$\epsilon_n = \frac{\hbar^2 k^2}{m} \quad (10.66)$$

Let $g(k) dk$ be the number of states with wave number lying between k and $k + dk$, and let $g^{(0)}(k) dk$ denote the corresponding quantity for the noninteracting system. Then (10.63) can be written in the form

$$\mathfrak{b}_2 - \mathfrak{b}_2^{(0)} = 2^{3/2} \left\{ \sum_B e^{-\beta\epsilon_B} + \int_0^\infty dk [g(k) - g^{(0)}(k)] e^{-\beta\hbar^2 k^2/m} \right\} \quad (10.67)$$

where ϵ_B denotes the energy of a bound state of the interacting two-body system.

We remark in passing that the factor $2^{3/2}$ in front of (10.67) is the ratio $(\lambda/\lambda_{cm})^{3/2}$, where λ is the thermal wavelength, and λ_{cm} is the thermal wavelength of the center-of-mass motion of the two-body system.

Let $\eta_l(k)$ be the scattering phase shift of the potential $v(r)$ for the l th partial wave of wave number k . It will be shown that

$$g(k) - g^{(0)}(k) = \frac{1}{\pi} \sum'_l (2l+1) \frac{\partial \eta_l(k)}{\partial k} \quad (10.68)$$

where the sum \sum' extends over the values

$$l = \begin{cases} 0, 2, 4, 6, \dots & \text{(bosons)} \\ 1, 3, 5, \dots & \text{(fermions)} \end{cases} \quad (10.69)$$

Therefore

$$\mathfrak{b}_2 - \mathfrak{b}_2^{(0)} = 2^{3/2} \left\{ \sum_B e^{-\beta\epsilon_B} + \frac{1}{\pi} \int_0^\infty dk \sum'_l (2l+1) \frac{\partial \eta_l(k)}{\partial k} e^{-\beta\hbar^2 k^2/m} \right\} \quad (10.70)$$

A partial integration leads finally to the formula

$$\mathfrak{b}_2 - \mathfrak{b}_2^{(0)} = 2^{3/2} \left\{ \sum_B e^{-\beta\epsilon_B} + \frac{\lambda^2}{\pi^2} \sum'_l (2l+1) \int_0^\infty dk k \eta_l(k) e^{-\beta\hbar^2 k^2/m} \right\} \quad (10.71)$$

It remains to prove (10.68). We may choose both $\psi_n(\mathbf{r})$ and $\psi_n^{(0)}(\mathbf{r})$ to be pure spherical harmonics, because $v(r)$ does not depend on the angles of \mathbf{r} with respect to any fixed axis. Thus we write

$$\begin{aligned} \psi_{klm}(\mathbf{r}) &= A_{klm} Y_l^m(\theta, \phi) \frac{u_{kl}(r)}{r} \\ \psi_{klm}^{(0)}(\mathbf{r}) &= A_{klm}^{(0)} Y_l^m(\theta, \phi) \frac{u_{kl}^{(0)}(r)}{r} \end{aligned} \quad (10.72)$$

For bosons $\psi(\mathbf{r}) = +\psi(-\mathbf{r})$, and for fermions $\psi(\mathbf{r}) = -\psi(-\mathbf{r})$. Therefore

$$l = \begin{cases} 0, 2, 4, 6, \dots & \text{(bosons)} \\ 1, 3, 5, 7, \dots & \text{(fermions)} \end{cases} \quad (10.73)$$

Let the boundary conditions be

$$u_{kl}(R) = u_{kl}^{(0)}(R) = 0 \quad (10.74)$$

where R is a very large radius which approaches infinity at the end of the calculation. The asymptotic forms of u_{kl} and $u_{kl}^{(0)}$ are

$$\begin{aligned} u_{kl}(r) &\xrightarrow{r \rightarrow \infty} \sin \left[kr + \frac{l\pi}{2} + \eta_l(k) \right] \\ u_{kl}^{(0)}(r) &\xrightarrow{r \rightarrow \infty} \sin \left(kr + \frac{l\pi}{2} \right) \end{aligned} \quad (10.75)$$

This defines $\eta_l(k)$. The eigenvalues k are determined by the boundary conditions (10.74):

$$kR + \frac{1}{2}a + \eta_l(k) = \pi n \quad (\text{interacting system}) \quad (10.76)$$

$$kR + \frac{1}{2}a = \pi n \quad (\text{noninteracting system})$$

where $n = 0, 1, 2, \dots$. It is seen that the eigenvalues k depends on n and l but not on m . Since there are $2l + 1$ spherical harmonics Y_l^m for a given l , each eigenvalue k is $(2l + 1)$ -fold degenerate.

For a given l , changing n by one unit causes k to change by the respective amounts Δk , $\Delta k^{(0)}$:

$$\Delta k = \frac{a}{R + [\partial \eta_l(k) / \partial k]} \quad (10.77)$$

$$\Delta k^{(0)} = \frac{\pi}{R}$$

These are the spacings of eigenvalues for a given l . Let the number of states of a given l with wave number lying between k and $k + dk$ be denoted by $g_l(k) dk$ and $g_l^{(0)}(k) dk$ for the two cases. We must have

$$\frac{g_l(k) \Delta k}{2l + 1} = 1$$

$$\frac{g_l^{(0)}(k) \Delta k^{(0)}}{2l + 1} = 1 \quad (10.78)$$

$$g_l(k) = \frac{2l + 1}{\pi} \left[R + \frac{\partial \eta_l(k)}{\partial k} \right] \quad (10.79)$$

$$g_l^{(0)}(k) = \frac{2l + 1}{\pi} R$$

Therefore

$$g_l(k) - g_l^{(0)}(k) = \frac{2l + 1}{\pi} \frac{\partial \eta_l(k)}{\partial k} \quad (10.80)$$

Summing (10.80) over all l consistent with (10.73) we obtain (10.68).

For $l > 2$ there is no known formula for b_l comparable in simplicity to (14.71), because there is no known treatment of the l -body problem for $l > 2$ comparable to the phase shift analysis of the two-body problem.*

Lastly, we remark that there is no essential difference between a sharp scattering resonance and a bound state, as far as the second virial coefficient is concerned. In the neighborhood of a sharp resonance, the scattering phase shift

increases by π over a small energy interval. In the idealized limit of an infinitely sharp resonance, we can represent the phase shift by

$$\frac{\partial \eta(k)}{\partial k} = \pi \delta(k - k_0) \quad (10.81)$$

where k_0 marks the position of the resonance. From (10.70), it is clear that each sharp resonance contributes to the second virial coefficient a term of the same form as that from a bound state. This supports what we expect, namely, that a sharp resonance can be treated as a particle.

10.4 VARIATIONAL PRINCIPLES

In quantum mechanics we are familiar with the variational principle, which states that the lowest energy eigenvalue of the system is the minimum of the expectation value of the Hamiltonian, taken with respect to a wave function that is completely arbitrary, except for normalization and the imposed boundary conditions of the problem. By using a trial wave function with adjustable parameters, one can use the variational principle to obtain an upper bound for the ground state energy, and improve on the bound by giving the trial wave function more freedom to vary. We shall describe here similar variational principles for the partition function.

Gibbs Variational Principle

Let p denote a normalized density function for an ensemble, classical or quantum mechanical. That is, it is a real positive quantity in the classical case, and a Hermitian operator with positive eigenvalues in the quantum case, and that

$$\text{Tr} p = 1 \quad (10.82)$$

In the classical case the operation Tr means $\int \phi p d\phi$. Now define

$$\psi(p) \equiv \text{Tr}(\mathcal{H}p) + \beta^{-1} \text{Tr}(p \log p) \quad (10.83)$$

where \mathcal{H} is the Hamiltonian of the system under consideration, and β is a constant. The Gibbs variational principle states the following:

- (a) Minimize $\psi(p)$ by varying p , subject only to the condition that it be a legitimate normalized density function. The function \bar{p} that minimizes $\psi(p)$ is the density function of the canonical ensemble with $kT = \beta^{-1}$.
- (b) The Helmholtz free energy is given by $A = \psi(\bar{p})$.

To prove this, first calculate the variations of ψ when p changes by δp :

$$\delta \psi = \text{Tr} \left\{ [\mathcal{H} + \beta^{-1}(1 + \log p)] \delta p \right\}$$

$$\delta^2 \psi = \text{Tr} \left[(\beta p)^{-1} (\delta p)^2 \right] \quad (10.84)$$

*An analysis of b_3 is given by A. Pais and G. E. Uhlenbeck, *Phys. Rev.* **116**, 250 (1959).

The second of these shows that D is a convex function, so that $\delta\psi = 0$ gives a minimum. We now vary p , taking the normalization constraint into account through a Lagrange multiplier λ :

$$0 = \delta\psi + \lambda\delta(\text{Tr}\rho) = \text{Tr} \left\{ \left[\mathcal{H} + \beta^{-1}(1 + \log \rho) + \lambda \right] \delta\rho \right\} \quad (10.85)$$

Solving for p , and determining λ by (10.82), we obtain

$$\bar{p} = e^{-\beta\mathcal{H}} / \text{Tr} e^{-\beta\mathcal{H}} \quad (10.86)$$

which proves (a). Substituting this into (10.83) gives

$$\psi(\bar{p}) = -\beta^{-1} \log \text{Tr} e^{-\beta\mathcal{H}} \quad (10.87)$$

which proves (b). ■

Peierls Variational Principle

Consider a quantum mechanical system with Hamiltonian \mathcal{H} , and partition function $Q = \text{Tr} \exp(-\beta\mathcal{H})$. The Peierls variational principle* states that

$$Q \geq \sum_n e^{-\beta(\Phi_n, \mathcal{H}\Phi_n)} \quad (10.88)$$

where $\{\Phi_n\}$ is an arbitrary set of wave functions of the system. Obviously, the equality holds when $\{\Phi_n\}$ is the set of eigenfunctions of \mathcal{H} . Note that the set $\{\Phi_n\}$ does not have to be complete, for the inequality holds a fortiori for an incomplete set since the terms on the right side of (10.88) are all positive. Thus it suffices to prove (10.88) under the assumption that $\{\Phi_n\}$ is a complete set of wave functions.

The Peierls variational principle is a special case of a more general theorem on convex functions. Suppose $f(x)$ is a real convex function of a real variable x , (i.e., $f''(x) \geq 0$.) Let us denote by \bar{f} the average of $f(x)$ over a selected set of x 's, with specified weights:

$$\bar{f}(x) \equiv \sum c_n f(x_n) \quad (10.89)$$

where $\{x_n\}$ is an arbitrary set of real numbers, and $\{c_n\}$ is a set of real numbers such that

$$c_n \geq 0, \quad \sum_n c_n = 1 \quad (10.90)$$

By the mean-value theorem,

$$f(x) = f(\bar{x}) + \frac{1}{2}(x - \bar{x})^2 f''(x_1) \quad (10.91)$$

for some x_1 . Now average both sides:

$$\bar{f}(x) = f(\bar{x}) + \frac{1}{2}(x - \bar{x})^2 \overline{f''(x_1)} \quad (10.92)$$

*R. E. Peierls, Phys. Rev. 54, 918 (1938).

Since $f''(x) \geq 0$, we have the following theorem:

$$\overline{f(x)} \geq f(\bar{x}) \quad (10.93)$$

Now let $\{\Phi_n\}$ be a complete set of wave functions, and S the unitary matrix which relates it to the eigenfunctions $\{\Psi_n\}$ of \mathcal{H} :

$$\Phi_n = \sum_m S_{nm} \Psi_m \quad (10.94)$$

$$\sum_n |S_{nm}|^2 = 1$$

from which we can see that

$$(\Phi_n, \mathcal{H}\Phi_n) = \sum_m |S_{nm}|^2 E_m \quad (10.95)$$

where $\{E_m\}$ are the eigenvalues of \mathcal{H} . We can write the partition function in the form

$$Q = \sum_n e^{-\beta E_n} = \sum_n \sum_m |S_{nm}|^2 e^{-\beta E_m} \quad (10.96)$$

and define

$$q = \sum_n e^{-\beta(\Phi_n, \mathcal{H}\Phi_n)} = \sum_n \exp \left[-\beta \sum_m |S_{nm}|^2 E_m \right] \quad (10.97)$$

Let $f(x) = \exp(-\beta x)$. For each n , the following definitions fulfill the requirements of the previous theorem:

$$\begin{aligned} \bar{E}_n &= \sum_m |S_{nm}|^2 E_m \\ \overline{f(E)}_n &= \sum_m |S_{nm}|^2 f(E_m) \end{aligned} \quad (10.98)$$

Thus we can write

$$Q - q = \sum_n \left[\overline{f(E)}_n - f(\bar{E}_n) \right] \quad (10.99)$$

According to the theorem, $Q - q \geq 0$ term by term. ■

10.5 IMPERFECT GASES AT LOW TEMPERATURES

An imperfect gas is an extremely dilute system of particles that interact among themselves through an interparticle potential of finite range and of such a nature that there exists no two-particle bound state. The diluteness of the gas enables us to treat the interparticle interaction as a small perturbation on the ideal gas. An imperfect gas, therefore, is the first improvement on the ideal gas as a model for a physical gas. We shall consider an imperfect gas at extremely low temperatures.

For such a system there are two important parameters of the dimension of length: the thermal wavelength Λ and the average interparticle separation $v^{1/3}$. These two lengths may be of comparable magnitude, but they must be much larger than the range of the interparticle potential, or any other length in the problem, except that size of the container.

In quantum mechanics a particle cannot be localized within its de Broglie wavelength, which in the present case may be replaced by the thermal wavelength. Thus in the present case a particle "spreads" over a distance much larger than the range of the interaction potential. Within the range of interaction of any given particle, the probability of finding another particle is small. Therefore

- (a) the effective interaction experienced by a particle is small, even though the interparticle potential may have large values;
- (b) the details of the interparticle potential are unimportant, because a particle that is spread out in space sees only an averaged effect of the potential.

In the quantum theory of scattering it is known that at low energies the scattering of a particle by a potential does not depend on the shape of the potential, but depends only on a single parameter obtainable from the potential — the scattering length a . The total scattering cross section at low energies is $4\pi a^2$. Hence roughly speaking a is the effective diameter of the potential. We may also say that at low energies the scattering from a potential looks like that from a hard sphere of diameter a . This makes it plausible that at extremely low temperatures it is possible to describe an imperfect gas solely in terms of the three parameters Λ , $v^{1/3}$, and a . Our problem is to formulate a method by which all the thermodynamic functions of the imperfect gas can be obtained to lowest order in the small parameters a/Λ and $a/v^{1/3}$.

We first show that, for the purpose of calculating the low-lying energy levels of an imperfect gas, the Hamiltonian of the system may be replaced by an effective Hamiltonian in which only scattering parameters, such as the scattering length, appear explicitly. The partition function of the imperfect gas can then be calculated with the help of the effective Hamiltonian. This method, first introduced by Fermi,* is known as the method of *pseudopotentials*.

Consider first a system of two particles interacting through a finite-ranged potential which has no bound state. The object of the method of pseudopotentials is to obtain all the energy levels of the system in terms of the scattering phase shifts of the potential. For the sake of concreteness we first assume that the potential is the hard-sphere potential with diameter a . The wave function for the two particles may be written in the form

$$\Psi(\mathbf{r}_1, \mathbf{r}_2) = e^{i\mathbf{P}\cdot\mathbf{R}} \psi(\mathbf{r}) \quad (10.100)$$

*E. Fermi, *Ricerca Sci.* 7, 13 (1936). Our presentation follows that of K. Huang and C. N. Yang, *Phys. Rev.* 105, 767 (1957).

where

$$\begin{aligned} \mathbf{R} &= \frac{1}{2}(\mathbf{r}_1 + \mathbf{r}_2) \\ \mathbf{r} &= \mathbf{r}_2 - \mathbf{r}_1 \end{aligned} \quad (10.101)$$

and \mathbf{P} is the total momentum vector. The Schrodinger equation in the center-of-mass system is

$$\begin{aligned} (\nabla^2 + k^2)\psi(\mathbf{r}) &= 0 & (r > a) \\ \psi(\mathbf{r}) &= 0 & (r \leq a) \end{aligned} \quad (10.102)$$

The hard-sphere potential is no more than a boundary condition for the relative wave function $\psi(\mathbf{r})$. It is understood that some boundary condition for $r \rightarrow \infty$ is specified, but what it is is irrelevant to our considerations. The number k is the relative wave number, and (10.102) presents an eigenvalue problem for k . When the allowed values of k are known, the energy eigenvalues of the system are given by

$$E(\mathbf{P}, \mathbf{k}) = \frac{P^2}{2M} + \frac{\hbar^2 k^2}{2\mu}$$

where M is the total mass and μ the reduced mass of the system.

The aim of the method of pseudopotentials is to replace the hard-sphere boundary condition by an inhomogeneous term for the wave equation. Such an idea is familiar in electrostatics, where to find the electrostatic potential in the presence of a metallic sphere (with some given boundary condition at infinity) we may replace the sphere by a distribution of charges on the surface of the sphere and find the potential set up by the fictitious charges. We can further replace the surface charges by a collection of multipoles at the center of the sphere with appropriate strengths. If we solve the Poisson equation with these multipole sources, we obtain the exact electrostatic potential outside the sphere. In an analogous way, the method of pseudopotentials replaces the boundary condition on $\psi(\mathbf{r})$ by a collection of sources at the point $r = 0$. Instead of producing electrostatic multipole potentials, however, these sources will produce scattered S waves, P waves, D waves, etc.

Let us first consider spherically symmetric (S wave) solutions of (10.102) at very low energies ($k \rightarrow 0$). The equations (10.102) become

$$\begin{aligned} \frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d\psi}{dr} \right) &= 0 & (r > a) \\ \psi(r) &= 0 & (r \leq a) \end{aligned} \quad (10.103)$$

The solution is obviously

$$\psi(r) = \begin{cases} \text{const.} \left(1 - \frac{a}{r} \right) & (r > a) \\ 0 & (r \leq a) \end{cases} \quad (10.104)$$

Now define an extended wave function $\psi_{\text{ex}}(r)$ such that

$$(\nabla^2 + k^2)\psi_{\text{ex}}(r) = 0 \quad (\text{everywhere except at } r = 0) \quad (10.105)$$

with the boundary condition

$$\psi_{\text{ex}}(a) = 0 \quad (10.106)$$

For $k \rightarrow 0$ we have

$$\psi_{\text{ex}}(r) \xrightarrow{r \rightarrow 0} \left(1 - \frac{a}{r}\right)\chi \quad (10.107)$$

where χ is a constant that depends on the boundary condition at $r = \infty$. We can avoid explicit use of this boundary condition by writing

$$\chi = \left[\frac{\partial}{\partial r}(r\psi_{\text{ex}}) \right]_{r=0} \quad (10.108)$$

which is an immediate consequence of (10.107). To eliminate the explicit requirement (10.106), we generalize the equation (10.106) to include the point $r = 0$. This can be easily done by finding the behavior of $(\nabla^2 + k^2)\psi_{\text{ex}}$ near $r = 0$, as required by (10.107). Since $k \rightarrow 0$, it is sufficient to note that according to (10.107)

$$\nabla^2\psi_{\text{ex}}(r) \xrightarrow{r \rightarrow 0} 4\pi a \delta(\mathbf{r})\chi = 4\pi a \delta(\mathbf{r}) \frac{\partial}{\partial r}(r\psi_{\text{ex}}) \quad (10.109)$$

Therefore as $k \rightarrow 0$ the function $\psi_{\text{ex}}(r)$ everywhere satisfies the equation

$$(\nabla^2 + k^2)\psi_{\text{ex}}(r) = 4\pi a \delta(\mathbf{r}) \frac{\partial}{\partial r}(r\psi_{\text{ex}}) \quad (10.110)$$

The operator $\delta(\mathbf{r})(\partial/\partial r)r$ is the pseudopotential.* For small k and for $r \geq a$, $\psi_{\text{ex}}(r)$ satisfies the same equation and the same boundary condition as $\psi(r)$. Therefore $\psi_{\text{ex}}(r) = \psi(r)$ for $r \geq a$, and the eigenvalues of k are the same in both cases.

The equation (10.110) is not the exact equation we desire, because only the S-wave solutions with small k coincide with the actual solutions of the physical problem. To obtain an equation for an extended wave function that rigorously coincides with $\psi(r)$ for $r \geq a$ it is necessary to generalize (10.110) to arbitrary values of k and to nonspherically symmetric solutions.† It suffices for the present to state that the result of the generalization consists of the following modifications of (10.110):

(a) The exact S-wave pseudopotential is

$$-\frac{4\pi}{k \cot \eta_0} \delta(\mathbf{r}) \frac{\partial}{\partial r} r \quad (10.111)$$

*The foregoing derivation is due to J. M. Blatt and V. F. Weisskopf, *Theoretical Nuclear Physics* (Wiley, New York, 1952), p. 74.

†See Huang and Yang, *op. cit.*

where η_0 is the S-wave phase shift for the hard-sphere potential:

$$-\frac{1}{k \cot \eta_0} = \frac{\tan ka}{k} = a \left[1 + \frac{1}{3}(ka)^2 + \dots \right] \quad (10.112)$$

(b) An infinite series of pseudopotentials is added to the right side of (10.110), representing the effects of P-wave scattering, D-wave scattering, etc. The l th-wave pseudopotential is proportional to a^{2l+1} .

From these results it is seen that (10.110) is correct up to the order a^2 . That is, if the wave function $\psi(r)$ and the eigenvalue k are expanded in a power series in a , then (10.110) correctly gives the coefficients of a and a^2 .

The differential operator $(\partial/\partial r)r$ in the pseudopotential (10.110) may be replaced by unity if $\psi_{\text{ex}}(\mathbf{r})$ is well behaved at $r = 0$, for then

$$\left[\frac{\partial}{\partial r}(r\psi_{\text{ex}}) \right]_{r=0} = \psi_{\text{ex}}(0) + \left[r \frac{\partial}{\partial r} \psi_{\text{ex}} \right]_{r=0} = \psi_{\text{ex}}(0) \quad (10.113)$$

If $\psi_{\text{ex}}(\mathbf{r}) \xrightarrow{r \rightarrow 0} Ar^{-1} + B$, however, then

$$\left[\frac{\partial}{\partial r}(r\psi_{\text{ex}}) \right]_{r=0} = B \quad (10.114)$$

An illustration of the effect of $(\partial/\partial r)r$ is given in Problem 10.7.

We now turn to the method of pseudopotentials for the case of two particles interacting through a general finite-ranged potential which has no bound state. Here (10.102) is replaced by the equation

$$\frac{\hbar^2}{2\mu}(\nabla^2 + k^2)\psi(\mathbf{r}) = v(r)\psi(\mathbf{r}) \quad (10.115)$$

with some given boundary condition for $r \rightarrow \infty$. At low energies only S-wave scattering is important. Therefore let us consider only spherically symmetric solutions. Then (10.115) reduces to

$$u''(r) + k^2u(r) = \frac{\mu}{\hbar^2}v(r)u(r) \quad (10.116)$$

where

$$u(r) \equiv r\psi(r) \quad (10.117)$$

By assumption $v(r)$ is finite-ranged and has no bound state. Therefore, as $r \rightarrow \infty$, $u(r)$ approaches a sinusoidal function:

$$u(r) \xrightarrow{r \rightarrow \infty} u_\infty(r) \quad (10.118)$$

where

$$u_\infty(r) \equiv r\psi_\infty(r) = \text{const.} (\sin kr + \tan \eta_0 \cos kr) \quad (10.119)$$

where η_0 is by definition the S-wave phase shift. For $k \rightarrow 0$,

$$\psi_\infty(r) \xrightarrow{r \rightarrow 0} \text{const.} \left(1 + \frac{\tan \eta_0}{kr} \right) \quad (10.120)$$

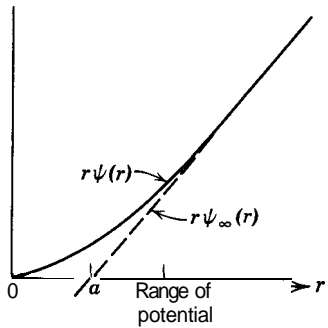


Fig. 10.3 Wave function in a repulsive potential with positive scattering length.

In general η_0 is a function of k . For small k there is a well-known expansion analogous to (10.112), known as the effective range expansion:

$$k \cot \eta_0 = -\frac{1}{a} + \frac{1}{2} k^2 r_0 + \dots \quad (10.121)$$

where a is called the scattering length and r_0 the effective range. The meaning of the scattering length can be seen by substituting (10.121) into (10.119). For $k \rightarrow 0$ we obtain (10.117). As illustrated in Figs. 10.3 and 10.4, the scattering length is the intercept of the asymptotic wave function $r\psi_\infty(r)$ with the coordinate axis. For the hard-sphere potential the scattering length is the hard-sphere diameter. In general a may be either positive or negative. It is positive for a predominantly repulsive potential (Fig. 10.3) and negative for a predominantly attractive potential (Fig. 10.4).

At low energies we may neglect all terms in (10.121) except $-1/a$ and obtain

$$-\frac{1}{k \cot \eta_0} = a$$

This approximation, known as the "shape-independent approximation," states that at low energies the potential acts as if it were a hard-sphere potential of

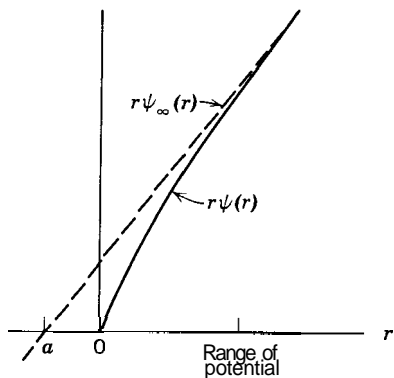


Fig. 10.4 Wave function in an attractive potential with negative scattering length.

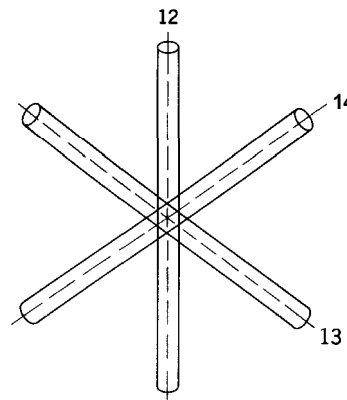


Fig. 10.5 The three-like hypersurface in the $3N$ -dimensional configuration space. The hard-sphere interactions are equivalent to the boundary condition that the wave function vanishes on the surface of the "tree."

diameter a . Therefore (10.110) can be taken over.* In general (10.110) is certainly valid for the calculation of the energy to the lowest order in the scattering length a . Whether it is still meaningful to use (10.110) for higher orders in a depends on the potential.

Having introduced the pseudopotentials in the two-body problem we are now in a position to discuss the generalization to the N -body problem. The considerations that follow are independent of statistics.

Let us first consider the n -body problem with hard-sphere interactions. The Schrodinger equation for the system is

$$-\frac{\hbar^2}{2m} (\nabla_1^2 + \dots + \nabla_N^2) \Psi = E \Psi \quad (|\mathbf{r}_i - \mathbf{r}_j| > a, \text{ all } i \neq j) \quad (10.122)$$

$$\Psi = 0 \quad (\text{otherwise})$$

We also require that Ψ satisfies some boundary condition on the surface of a large cube, e.g., that Ψ satisfies periodic boundary conditions. The hard-sphere interactions are equivalent to a boundary condition that requires Ψ to vanish whenever $|\mathbf{r}_i - \mathbf{r}_j| = a$, for all $i \neq j$. In the $3N$ -dimensional configuration space the collection of all points for which $|\mathbf{r}_i - \mathbf{r}_j| = a$ represents a tree-like hypersurface, a portion of which we schematically represent by Fig. 10.5. Thus we draw a cylinder, labeled 12, to represent the surface in which $|\mathbf{r}_1 - \mathbf{r}_2| = a$, whereas $\mathbf{r}_3, \dots, \mathbf{r}_N$ may have arbitrary values. The whole "tree" is the totality of all such cylinders, $\frac{1}{2}N(N-1)$ in number, which mutually intersect in a complicated way. If the hard-sphere diameter a is small, these cylinders have a small radius. To find the wave function outside the "tree," it is natural to replace the "tree" by a series of "multipoles" at the "axes," i.e., at the lines $|\mathbf{r}_i - \mathbf{r}_j| = 0$.

It can be easily shown that replacing the effect of each cylinder by multipoles along its axis amount to introducing the two-body pseudopotentials described in the previous section. Our extended wave function would then satisfy a

*The derivation of (10.110) remains valid if a is negative.

Schrodinger equation containing the sum of $\frac{1}{2}N(N-1)$ two-body pseudopotentials. These two-body pseudopotentials, however, do not exactly replace the effect of the "tree." Although they correctly give the behavior of Ψ near a cylinder and far away from any intersection of cylinders, they do not necessarily give the correct behavior of Ψ near an intersection of two or more cylinders. For example, the intersection corresponding to $|\mathbf{r}_1 - \mathbf{r}_2| = a$ and $|\mathbf{r}_1 - \mathbf{r}_3| = a$ represents a configuration in which particles 1, 2, and 3 collide simultaneously—an intrinsically three-body effect which has not been taken into account in the two-body pseudopotentials. The sum of two-body pseudopotentials accounts only for the effects of binary collisions.

Using our geometrical picture, we see that in addition to the two-body pseudopotentials it may be necessary to place additional multipoles (pseudopotentials) at each intersection of two or more cylinders. To find the exact magnitude of these three- and more-body pseudopotentials we would have to solve three- and more-body problems. Their dependence on the hard-sphere diameter a , however, can be found by a dimensional argument.

As an example, the three-body pseudopotential needed at the intersection of the lines $|\mathbf{r}_1 - \mathbf{r}_2| = 0$ and $|\mathbf{r}_1 - \mathbf{r}_3| = 0$ must appear in the three-body Schrodinger equation in the form

$$\begin{aligned} & (\nabla_1^2 + \nabla_2^2 + \nabla_3^2 + k^2)\Psi \\ & = (\text{sum of two-body pseudopotentials}) + 6(\mathbf{r}_1 - \mathbf{r}_2) \cdot 6(\mathbf{r}_1 - \mathbf{r}_3) K \Psi \end{aligned}$$

The quantity K must be of the dimension (length)⁴. At low energies ($k \rightarrow 0$) the only length in the problem is a . Therefore K must be of the order a^4 . In a similar way we deduce that four-body pseudopotentials are of the order a^7 , and so forth. These pseudopotentials may be ignored, if we are only interested in an accuracy up to the order a^2 . The necessity for such n -body pseudopotentials shows that the pseudopotentials are not additive. This is analogous to the well-known situation in electrostatics that image charges are not additive. For example, the images of a point charge in front of two mutually orthogonal plane conductors are not simply the two images produced by each plane conductor taken separately.

If the interparticle potential is not the hard-sphere potential but a finite-ranged potential that has no bound state, the considerations just given can be taken over. The effective Hamiltonian for an imperfect gas of N identical particles of mass m may be taken to be

$$\mathcal{H} = \frac{\hbar^2}{2m} (\nabla_1^2 + \cdots + \nabla_N^2) + \frac{4\pi a \hbar^2}{m} \sum_{i < j} \delta(\mathbf{r}_i - \mathbf{r}_j) \frac{\partial}{\partial r_{ij}} r_{ij} \quad (10.123)$$

where a is the scattering length. This is valid for both fermions and bosons. The eigenvalues of this Hamiltonian will be the correct eigenvalues for an imperfect hard-sphere gas up to order a^2 . For a general imperfect gas they will be correct to the lowest order in a .

We note that (10.123) is not a Hermitian operator because $(\partial/\partial r)r$ is not a Hermitian operator. This need not cause concern because, by its derivation, (10.123) has been shown to have real eigenvalues that are the approximate eigenvalues of the real problem. The non-Hermiticity reflects the fact that the eigenfunctions of (10.123) do not everywhere coincide with the eigenfunctions of the real problem, but do so only in the asymptotic region. The fact, however, that (10.123) is not Hermitian means that we cannot find its eigenvalues by variational methods.

If the pseudopotentials in (10.123) are regarded as small perturbations to be treated only to the first order in perturbation theory, then the operators $(\partial/\partial r)r$ will always act on unperturbed free-particle wave functions, which are well-behaved. Hence the operators $(\partial/\partial r)r$ can be set equal to unity, and we can work with the Hamiltonian

$$\mathcal{H}' = -\frac{\hbar^2}{2m} (\nabla_1^2 + \cdots + \nabla_N^2) + \frac{4\pi a \hbar^2}{m} \sum_{i < j} \delta(\mathbf{r}_i - \mathbf{r}_j) \quad (10.124)$$

It is to be emphasized that this Hamiltonian is *valid* only for the purpose of applying *first-order* perturbation theory. We must not diagonalize (10.124) exactly, because the exact eigenvalues are the same as those for a free-particle system—it being well known that a three-dimensional S-function potential produces no scattering.

The first-order energy levels of \mathcal{H}' are calculated in the Appendix a. The result for bosons is given in (A.36); that for fermions in (A.42).

PROBLEMS

10.1 (a) Calculate δ_2 and δ_3 for a classical hard-sphere gas with hard-sphere diameter a .

(b) Express the equation of state of a classical hard-sphere gas in the form of a virial expansion. Include terms up to the third virial coefficient.

10.2 Find b_2 for an ideal Bose gas and compare it with δ_2 . Is the difference significant? (See (8.72).)

10.3 Calculate the second virial coefficients for a spinless hard-sphere Bose gas and a spinless hard-sphere Fermi gas to the two lowest nonvanishing orders in a/λ , where a is the hard sphere diameter and λ is the thermal wavelength.

Answers.

$$\delta_2 = 2^{-5/2} - \frac{2a}{\lambda} - \frac{10\pi^2}{3} \left(\frac{a}{\lambda}\right)^5 + \cdots \quad (\text{Bose})$$

$$\delta_2 = -2^{-5/2} - 6\pi \left(\frac{a}{\lambda}\right)^3 + 18\pi^2 \left(\frac{a}{\lambda}\right)^5 + \cdots \quad (\text{Fermi})$$

10.4 In calculating W_N defined in (10.35), the symmetry or antisymmetry of the wave function makes the calculation complicated. The following is a method to deal with this problem.

Let the free-particle wave functions for a system of distinguishable particles be

$$\chi_p(1, \dots, N) \equiv \frac{1}{V^{N/2}} e^{i(\mathbf{p}_1 \cdot \mathbf{r}_1 + \dots + \mathbf{p}_N \cdot \mathbf{r}_N)}$$

Let

$$\begin{aligned} & (1, \dots, N | e^{-\beta \mathcal{H}} | 1', \dots, N') \\ &= \sum_{\mathbf{P}_1} \dots \sum_{\mathbf{P}_N} \chi_p^*(1, \dots, N) e^{-\beta \mathcal{H}} \chi_p(1', \dots, N') \end{aligned}$$

The symbol $|1, \dots, N\rangle$ may be regarded as an eigenvector of the position operators of N distinguishable particles. Show that with the help of this quantity (10.35) may be expressed in the form

$$\frac{1}{\lambda^{3N}} W_N(1, \dots, N) = \sum_{\mathbf{P}} \delta_{\mathbf{P}} \langle 1, \dots, N | e^{-\beta H} | \mathbf{P}1, \dots, \mathbf{P}N \rangle$$

10.5 Models for Ferromagnetism. Consider a lattice of N fixed atoms of spin $\frac{1}{2}$. The quantum mechanical spin operators of the i th atom are the Pauli spin matrices \mathbf{a}_i . Assuming that only nearest neighbors interact via a spin-spin interaction, we obtain the *Heisenberg model of ferromagnetism*. The Hamiltonian is

$$\mathcal{H}_{\text{Heisenberg}} = -\epsilon \sum_{\langle ij \rangle} \sigma_i \cdot \sigma_j - \mu \sum_{i=1}^N \sigma_i \cdot \mathbf{H}$$

where $\langle ij \rangle$ denotes a nearest-neighbor pair, \mathcal{H} is a uniform external magnetic field, and ϵ and μ are positive constants.

Another model, the *Ising model*, is constructed by associating with the i th atom a *number* s_i that is either $+1$ or -1 and taking the Hamiltonian to be

$$\mathcal{H}_{\text{Ising}} = -\epsilon \sum_{\langle ij \rangle} s_i s_j - \mu \sum_{i=1}^N s_i H$$

where H is the z component of \mathbf{H}

Using the Peierls variational principle prove that, for the same temperature, the Helmholtz free energy of the Heisenberg model is not greater than that of the Ising model.

10.6 Mean-Field Approximation. Consider the Ising model, whose Hamiltonian is given in the last problem. In the mean-field approximation one assumes that each spin sees a mean field due to all its neighbors. Determine this mean field with the help of the Gibbs variational principle, as follows:

(a) Assume a product form for the trial density function

$$\begin{aligned} \rho(s_1, \dots, s_N) &= g(s_1) \dots g(s_N) \\ g(s) &= C e^{Bs} \end{aligned}$$

Find C by normalizing $g(s)$. The mean field B is to be determined.

(b) Instead of B , use as variational parameter the magnetization per spin

$$m = \sum_s s g(s)$$

Show that $B = \tanh^{-1} m$.

(c) Show that the Gibbs function $\psi(\rho)$, as defined in (10.83) is given by

$$\psi = N \left[-\frac{1}{2} \epsilon \gamma m^2 - \mu H m + kT (B m + \log C) \right]$$

where y is the number of nearest neighbors.

(d) Show that ψ is minimized by \bar{m} , which satisfies

$$\bar{m} = \tanh [(\epsilon \gamma \bar{m} + \mu H)/kT]$$

(e) Show that the Helmholtz free energy per spin is given by

$$\frac{A}{N} = -\frac{1}{2} \epsilon \gamma \bar{m}^2 + \mu H \bar{m} + kT \left[\frac{1 + \bar{m}}{2} \log \frac{1 + \bar{m}}{2} + \frac{1 - \bar{m}}{2} \log \frac{1 - \bar{m}}{2} \right]$$

These results are the same as those of the Bragg-Williams approximation, which we shall derive and analyze in Chapter 14.

10.7 (a) Find all spherically symmetric solutions and corresponding eigenvalues of the equation

$$(\nabla^2 + k_n^2) \psi_n(r) = 0$$

in the region between two concentric spheres of radii R and a ($R > a$), with the boundary conditions

$$\psi(R) = \psi(a) = 0$$

(b) Expand the eigenvalues k_n^2 in powers of a , keeping terms up to order a^2

(c) Using the method of pseudopotentials, calculate the eigenvalue k_n^2 up to order a^2 and show that it agrees with the answer to (b).

Reference. K. Huang and C. N. Yang, *Phys. Rev.* 105, 767 (1957), §2(b).

CHAPTER

11

FERMI SYSTEMS

In this chapter we study various examples of systems of fermions. The dominant common characteristic is the existence of the Fermi surface, which is a direct consequence of the Pauli exclusion principle.

11.1 THE EQUATION OF STATE OF AN IDEAL FERMI GAS

The equation of state of a spinless ideal Fermi gas is obtained by eliminating z from Eqs. (8.67). We first study the behavior of z as determined by the second equation of (8.67), namely

$$\frac{\lambda^3}{v} = f_{3/2}(z) \quad (11.1)$$

where $v = V/N$, $\lambda = \sqrt{2\pi\hbar^2/mkT}$ is the thermal wavelength, and

$$f_{3/2}(z) = \frac{4}{\sqrt{\pi}} \int_0^\infty dx \frac{x^2}{z^{-1}e^{x^2} + 1} \quad (11.2)$$

is a monotonically increasing function of z . For small z we have the power series expansion

$$f_{3/2}(z) = z - \frac{z^2}{2^{3/2}} + \frac{z^3}{3^{3/2}} - \frac{z^4}{4^{3/2}} + \dots \quad (11.3)$$

For large z an asymptotic expansion may be obtained through a method due to Sommerfeld, as follows. For convenience put $z = e^v$, so that v is related to the chemical potential μ by

$$v = \log z = \mu/kT \quad (11.4)$$

Then

$$\begin{aligned} f_{3/2}(z) &= \frac{4}{\sqrt{\pi}} \int_0^\infty dx \frac{x^2}{e^{x^2-v} + 1} = \frac{2}{\sqrt{\pi}} \int_0^\infty dy \frac{\sqrt{y}}{e^{y-v} + 1} \\ &= \frac{4}{3\sqrt{\pi}} \int_0^\infty dy \frac{y^{3/2}e^{y-v}}{(e^{y-v} + 1)^2} \end{aligned} \quad (11.5)$$

The last step is obtained through a partial integration. Expanding $y^{3/2}$ in a Taylor series about v , we obtain

$$\begin{aligned} f_{3/2}(z) &= \frac{4}{3\sqrt{\pi}} \int_0^\infty dy \frac{e^{y-v}}{(e^{y-v} + 1)^2} \\ &\quad \times \left[v^{3/2} + \frac{3}{2}v^{1/2}(y-v) + \frac{3}{8}v^{-1/2}(y-v)^2 + \dots \right] \\ &= \frac{4}{3\sqrt{\pi}} \int_{-v}^\infty dt \frac{e^t}{(e^t + 1)^2} (v^{3/2} + \frac{3}{2}v^{1/2}t + \frac{3}{8}v^{-1/2}t^2 + \dots) \end{aligned} \quad (11.6)$$

Now we write

$$\int_{-v}^\infty = \int_{-\infty}^{+\infty} - \int_{-\infty}^{-v}$$

The second integral is of order e^{-v} . Therefore

$$\begin{aligned} f_{3/2}(z) &= \frac{4}{3\sqrt{\pi}} \int_{-\infty}^{+\infty} dt \frac{e^t}{(e^t + 1)^2} (v^{3/2} + \frac{3}{2}v^{1/2}t + \frac{3}{8}v^{-1/2}t^2 + \dots) + O(e^{-v}) \\ &= \frac{4}{3\sqrt{\pi}} (I_0 v^{3/2} + \frac{3}{2}I_1 v^{1/2} + \frac{3}{8}I_2 v^{-1/2} + \dots) + O(e^{-v}) \end{aligned} \quad (11.7)$$

where

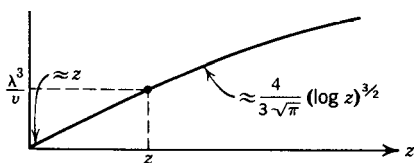
$$I_n \equiv \int_{-\infty}^{+\infty} dt \frac{t^n e^t}{(e^t + 1)^2} \quad (11.8)$$

Apart from the factor t^n , the integrand is an even function of t . Hence $I_n = 0$ for odd n . For $n = 0$ we have

$$I_0 = -2 \int_0^\infty dt \frac{d}{dt} (e^{-t/2} + 1) = 1 \quad (11.9)$$

and for even $n > 0$,

$$\begin{aligned} I_n &= -2 \left[\frac{\partial}{\partial \lambda} \int_0^\infty dt \frac{t^{n-1}}{e^{\lambda t} + 1} \right]_{\lambda=1} = 2n \int_0^\infty du \frac{u^{n-1}}{e^u + 1} \\ &= (n-1)!(2n)(1 - 2^{1-n})\zeta(n) \end{aligned} \quad (11.10)$$

$f_{3/2}(z)$

Fig. 11.1 The function $f_{3/2}(z)$.

where $\zeta(n)$ is the Riemann zeta function,* some special value of which are

$$\zeta(2) = \frac{\pi^2}{6}, \quad \zeta(4) = \frac{\pi^4}{90}, \quad \zeta(6) = \frac{\pi^6}{945}$$

Hence

$$f_{3/2}(z) = \frac{4}{3\sqrt{\pi}} \left[(\log z)^{3/2} + \frac{\pi^2}{8} (\log z)^{-1/2} + \dots \right] + O(z^{-1}) \quad (11.11)$$

A graph of $f_{3/2}(z)$ is shown *in* Fig. 11.1. For any given positive value of λ^3/v , the value of z determined by (11.1) can be read off such a graph. It is seen that z increases monotonically as A^3/u increases. For fixed v , z increases monotonically as the temperature decreases.

High Temperatures and Low Densities ($\lambda^3/v \ll 1$)

For $A^3/u \ll 1$ the average interparticle separation $v^{1/3}$ is much larger than the thermal wavelength λ . We expect quantum effects to be negligible. From (11.1) and (11.3),

$$\frac{A^3}{v} = z - \frac{z^2}{2^{3/2}} + \dots$$

which may be solved to give

$$z = \frac{\lambda^3}{v} + \frac{1}{2^{3/2}} \left(\frac{\lambda^3}{v} \right)^2 + \dots \quad (11.12)$$

Thus z reduces to that of the Boltzmann gas (Eq. (8.52)) when $A^3 \rightarrow 0$ ($T \rightarrow m$). The average occupation number (8.65) reduces to Maxwell-Boltzmann form:

$$\langle n_{\mathbf{p}} \rangle \approx \frac{\lambda^3}{v} e^{-\beta \epsilon_{\mathbf{p}}} \quad (11.13)$$

*cf. *Handbook of Mathematical Functions*, M. Abramowitz and I. A. Stegun, eds., (National Bureau of Standards, Washington, D.C., 1964), Chapter 23.

The equation of state (8.67) then becomes

$$\frac{Pv}{kT} = \frac{v}{\lambda^3} \left(z - \frac{z^2}{2^{5/2}} + \dots \right) = 1 + \frac{1}{2^{5/2}} \frac{\lambda^3}{v} + \dots \quad (11.14)$$

This is in the form of a virial expansion. The corrections to the classical ideal gas law, however, are not due to molecular interactions, but to quantum effects. The second virial coefficient in this case is

$$\frac{\lambda^3}{2^{5/2}} = \frac{1}{2} \left(\frac{\pi \hbar^2}{mkT} \right)^{3/2} \quad (11.15)$$

All other thermodynamic functions reduce to those for a classical ideal gas plus small corrections.

Low Temperatures and High Densities ($\lambda^3/v \gg 1$)

For $A^3/u \gg 1$ the average de Broglie wavelength of a particle is much greater than the average interparticle separation. Thus quantum effects, in particular the effects of the Pauli exclusion principle, become all important.

In the neighborhood of absolute zero we have, from (11.1) and (11.11),

$$\frac{1}{v} \left(\frac{2\pi \hbar^2}{mkT} \right)^{3/2} \approx \frac{4}{3\sqrt{\pi}} (\log z)^{3/2} \quad (11.16)$$

Hence

$$z \approx e^{\beta \epsilon_F} \quad (11.17)$$

where ϵ_F , the chemical potential at absolute zero, is called the Fermi energy:

$$\epsilon_F \equiv \frac{\hbar^2}{2m} \left(\frac{6\pi^2}{v} \right)^{2/3} \quad (11.18)$$

To study its physical significance, let us examine $\langle n_{\mathbf{p}} \rangle$ near absolute zero:

$$\langle n_{\mathbf{p}} \rangle \approx \frac{1}{e^{\beta(\epsilon_{\mathbf{p}} - \epsilon_F)} + 1} \quad (11.19)$$

If $\epsilon_{\mathbf{p}} < \epsilon_F$, then the exponential in the denominator vanishes as $T \rightarrow 0$ ($\beta \rightarrow m$). Hence $\langle n_{\mathbf{p}} \rangle = 1$. Otherwise, $\langle n_{\mathbf{p}} \rangle = 0$. Thus

$$\langle n_{\mathbf{p}} \rangle_{T=0} = \begin{cases} 1 & (\epsilon_{\mathbf{p}} < \epsilon_F) \\ 0 & (\epsilon_{\mathbf{p}} > \epsilon_F) \end{cases} \quad (11.20)$$

The physical meaning of this formula is clear. Because of the Pauli exclusion principle no two particles can be in the same state. Therefore, in the ground state of the system, the particles occupy the lowest possible levels and fill the levels up to the finite energy level ϵ_F . Thus ϵ_F is simply the single-particle energy level below which there are exactly N states. In momentum space the particles fill a sphere of radius p_F , the surface of which is called the Fermi surface.

With this interpretation, let us now calculate the Fermi energy independently, under more general conditions. Suppose all single-particle energy levels are g -fold degenerate. For example, $g = 2s + 1$ for a particle of spin s . The condition determining ϵ_F is then

$$g \sum_{\mathbf{p}} \langle n_{\mathbf{p}} \rangle_{T=0} = N \quad (11.21)$$

In view of (11.20), this states that there are N states with energy below the Fermi energy. Putting $\epsilon_F = p_F^2/2m$, we find

$$\frac{g}{(2\pi\hbar)^3} \frac{4\pi}{3} p_F^3 = \frac{N}{V} \quad (11.22)$$

Hence

$$\epsilon_F = \frac{\hbar^2}{2m} \left(\frac{6\pi^2}{gV} \right)^{2/3} \quad (11.23)$$

which reduces to (11.18) when $g = 1$. We can also interpret (11.21) as follows. Particles with different quantum numbers are not constrained by any symmetry requirement with respect to the interchange of their *positions*. Thus we may consider a system of N fermions, each with degeneracy g , to be made up of g independent Fermi gases each with N/g particles whose energies are *nondegenerate*.

To obtain the thermodynamic functions for low temperatures and high densities we first obtain the expansion for the chemical potential from (11.1) and (11.11):

$$kT\nu = kT \log z = \epsilon_F \left[1 - \frac{\pi^2}{12} \left(\frac{kT}{\epsilon_F} \right)^2 + \dots \right] \quad (11.24)$$

The expansion parameter is kT/ϵ_F . If we define the *Fermi temperature* T_F , which is a function of density, by

$$kT_F \equiv \epsilon_F \quad (11.25)$$

then low temperature and high density means $T \ll T_F$. In this domain the gas is said to be *degenerate* because the particles tend to go to the lowest energy levels possible. For this reason T_F is also called the *degeneracy temperature*.

The average occupation number is

$$\langle n_{\mathbf{p}} \rangle = \frac{1}{e^{\beta\epsilon_{\mathbf{p}} - \nu} + 1} \quad (11.26)$$

where ν is given by (11.24). Since $\epsilon_{\mathbf{p}} = p^2/2m$, $n_{\mathbf{p}}$ depends on \mathbf{p} only through p^2 . A sketch of $n_{\mathbf{p}}$ is shown in Fig. 11.2.

The internal energy is

$$U = \sum_{\mathbf{p}} \epsilon_{\mathbf{p}} \langle n_{\mathbf{p}} \rangle = \frac{V}{h^3} \frac{4\pi}{2m} \int_0^\infty dp p^4 \langle n_{\mathbf{p}} \rangle$$

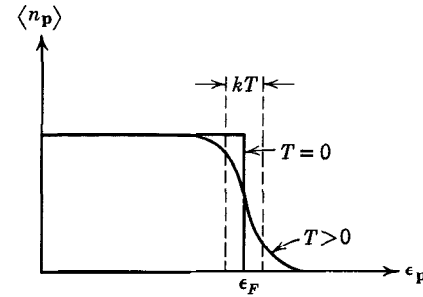


Fig. 11.2 Average occupation number in an ideal Fermi gas.

After a partial integration we obtain

$$U = \frac{V}{4\pi^2 m \hbar^3} \int_0^\infty dp \frac{p^5}{5} \left(-\frac{\partial}{\partial p} \langle n_{\mathbf{p}} \rangle \right) = \frac{\beta V}{20\pi^2 m^2 \hbar^3} \int_0^\infty dp \frac{p^6 e^{\beta\epsilon_{\mathbf{p}} - \nu}}{(e^{\beta\epsilon_{\mathbf{p}} - \nu} + 1)^2} \quad (11.27)$$

It is apparent from Fig. 11.2 that $\partial \langle n_{\mathbf{p}} \rangle / \partial p$ is sharply peaked at $p = p_F$. In fact, at absolute zero it is a δ function at $p = p_F$. Therefore the integral in (11.27) can be evaluated by expanding the factor p^6 about $p = p_F$. The procedure is similar to that used in obtaining (11.11). After inserting ν from (11.24) we obtain the asymptotic expansion

$$U = \frac{3}{5} N \epsilon_F \left[1 + \frac{5}{12} \pi^2 \left(\frac{kT}{\epsilon_F} \right)^2 + \dots \right] \quad (11.28)$$

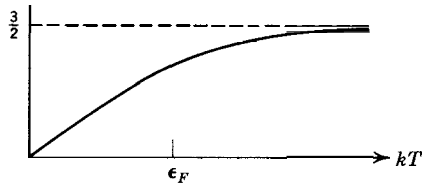
The first term is the ground state energy of the Fermi gas at the given density, as we can verify by showing the following:

$$\sum_{|\mathbf{p}| < p_F} \frac{p^2}{2m} = \frac{3}{5} N \epsilon_F \quad (11.29)$$

The specific heat at constant volume can be immediately obtained from (11.28)

$$\frac{C_V}{Nk} \approx \frac{\pi^2}{2} \frac{kT}{\epsilon_F} \quad (11.30)$$

It vanishes linearly as $T \rightarrow 0$, thus verifying the third law of thermodynamics. We know that C_V/Nk approaches $\frac{3}{2}$ as $T \rightarrow \infty$. Thus a rough sketch of C_V/Nk can be made, as shown in Fig. 11.3. The fact that it is proportional to T at these low temperatures can be understood as follows. At a temperature $T > 0$, $\langle n_{\mathbf{p}} \rangle$ differs from that at $T = 0$ because a certain number of particles are excited to energy levels $\epsilon_{\mathbf{p}} > \epsilon_F$. Roughly speaking, particles with energies of order kT below ϵ_F are excited to energies of order kT above ϵ_F (see Fig. 11.2). The number of particles excited is therefore of the order of $(kT/\epsilon_F)N$. Therefore the total excitation energy above the ground state is $\Delta U \approx (kT/\epsilon_F)NkT$, from which follows $C_V \approx (kT/\epsilon_F)Nk$.

C_V/Nk

Fig. 11.3 Specific heat of an ideal Fermi gas.

From (8.78) and (11.28) follows the equation of state

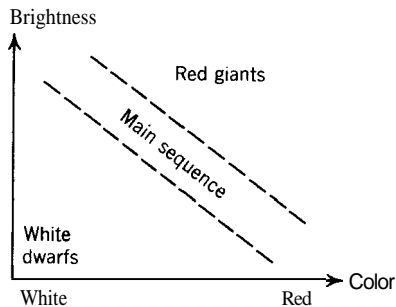
$$P = \frac{2}{3} \frac{U}{V} = \frac{2}{5} \frac{\epsilon_F}{v} \left[1 + \frac{5\pi^2}{12} \left(\frac{kT}{\epsilon_F} \right)^2 + \dots \right] \quad (11.31)$$

This shows that even at absolute zero it is necessary to contain the ideal Fermi gas with externally fixed walls because the pressure does not vanish. This is a manifestation of the Pauli exclusion principle, which allows only one particle to have zero momentum. All other particles must have finite momentum and give rise to the zero-point pressure.

To obtain the thermodynamic function for arbitrary values of λ^3/v numerical methods must be employed to calculate the functions $f_{3/2}(z)$ and $f_{5/2}(z)$.

11.2 THE THEORY OF WHITE DWARF STARS

It is an empirical rule that the brightness of a star is proportional to its color (i.e., predominant wavelength emitted). The proportionality constant is roughly the same for all stars. Thus if we make a plot of brightness against color, we obtain what is known as the Hertzsprung-Russell diagram, in which most stars fall within a linear strip called the main sequence, as shown in Fig. 11.4. There are, however, stars that are exceptions to this rule. There are the red giant stars, huge stars which are abnormally bright for their red color; and there are the white dwarf stars, small stars which are abnormally faint for their white color. The white


Fig. 11.4 Russell-Hertzsprung diagram.

dwarf star makes an interesting subject for our study, because to a good approximation it is a degenerate Fermi gas.

A detailed study of the constitution of white dwarf stars leads to the conclusion that they lack brightness because the hydrogen supply, which is the main energy source of stars, has been used up, and they are composed mainly of helium. What little brightness they have is derived from the gravitational energy released through a slow contraction of the star. Probably these stars have reached the end point of stellar evolution. One of the nearest stars to the solar system, the companion of Sirius, 8 light years from us, is a white dwarf. So faint that it escapes the naked eye, it was first predicted by the calculations of Bessel, who tried to explain why Sirius apparently moves about a point in empty space.

An idealized model of a white dwarf may be constructed from some typical data for such a star:

Content: mostly helium

Density $\approx 10^7 \text{ g/cm}^3 \approx 10^7 \rho_{\odot}$

Mass $\approx 10^{33} \text{ g} \approx M_{\odot}$

Central temperature $\approx 10^7 \text{ K} \approx T_{\odot}$

where the subscript \odot denotes quantities referring to the sun. Thus a white dwarf star is a mass of helium at an extremely high temperature and under extreme compression. The temperature 10^7 K corresponds to a thermal energy of 1000 eV . Hence the helium atoms are expected to be completely ionized, and the star may be regarded as a gas composed of helium nuclei and electrons. We regard the gas of electrons as an ideal Fermi gas, with a density of approximately 10^{30} electrons/cm³. This corresponds to a Fermi energy of

$$\epsilon_F \approx \frac{\hbar^2}{2m} \frac{1}{v^{2/3}} \approx 20 \text{ MeV}$$

and a Fermi temperature of

$$T_F \approx 10^{11} \text{ K}$$

Since the Fermi temperature is much greater than the temperature of the star, the electron gas is a highly degenerate Fermi gas, which behaves no differently from an electron gas at absolute zero. In fact we regard the electron gas to be an ideal Fermi gas in its ground state. The enormous zero-point pressure exerted by the electron gas is counteracted by the gravitational attraction that binds the star. This gravitational binding is due almost entirely to the helium nuclei in the star. The pressure due to kinetic motion of the helium nuclei, and to any radiation that may be present, will be neglected.

Thus we arrive at the following idealized model: A white dwarf is taken to be a system of N electrons in its ground state, at such a density that the electrons must be treated by relativistic dynamics. The electrons move in a background of $N/2$ motionless helium nuclei which provide the gravitational attraction to hold

the entire system together.* This model must then exhibit properties that are the combined effects of the Pauli principle, relativistic dynamics, and the gravitational law.

First let us work out the pressure exerted by a Fermi gas of relativistic electrons in the ground state. The states for a single electron are specified by the momentum p and the spin quantum number $s = \pm \frac{1}{2}$. The single-particle energy levels are independent of s :

$$\epsilon_{ps} = \sqrt{(pc)^2 + (m_e c^2)^2}$$

where m_e is the mass of an electron. The ground state energy of the Fermi gas is

$$E_0 = 2 \sum_{|\mathbf{p}| < p_F} \sqrt{(pc)^2 + (m_e c^2)^2} = \frac{2V}{h^3} \int_0^{p_F} dp 4\pi p^2 \sqrt{(pc)^2 + (m_e c^2)^2} \quad (11.32)$$

where p_F , the Fermi momentum, is defined by

$$\frac{V}{h^3} \left(\frac{4}{3} \pi p_F^3 \right) = \frac{N}{2}$$

$$p_F = \hbar \left(\frac{3\pi^2}{v} \right)^{1/3} \quad (11.33)$$

Changing the variable of integration in (11.32) to $x = p/m_e c$ we obtain

$$\frac{E_0}{N} = \frac{m_e^4 c^5}{\pi^2 \hbar^3} v f(x_F) \quad (11.34)$$

where

$$f(x_F) = \int_0^{x_F} dx x^2 \sqrt{1+x^2} = \begin{cases} \frac{1}{3} x_F^3 \left(1 + \frac{3}{10} x_F^2 + \dots \right) & (x_F \ll 1) \\ \frac{1}{4} x_F^4 \left(1 + \frac{1}{x_F^2} + \dots \right) & (x_F \gg 1) \end{cases} \quad (11.35)$$

and

$$x_F \equiv \frac{p_F}{m_e c} = \frac{\hbar}{m_e c} \left(\frac{3\pi^2}{v} \right)^{1/3} \quad (11.36)$$

*The temperature in an actual white dwarf star is so high that electron-positron pairs can be created in electron-electron collisions. These pairs in turn annihilate into radiation. Therefore in equilibrium there should be a certain number of electron-positron pairs and a certain amount of radiation present. We neglect the effects of these. It has been speculated that neutrinos can also be created in electron-electron, electron-positron, and photon-photon collisions with appreciable probability. This leads to some interesting phenomena, for neutrinos interact so weakly with matter that they do not come to thermal equilibrium with the rest of the system. They simply leave the star and cause a constant drain of energy. (H. Y. Chiu and P. Morrison, *Phys. Rev. Lett.* **5**, 573 (1960).) Our model is based on the neglect of these effects.

If the total mass of the star is M and the radius of the star is R , then

$$M = (m_e + 2m_p)N \approx 2m_p N$$

$$R = \left(\frac{3V}{4\pi} \right)^{1/3} \quad (11.37)$$

where m_p is the mass of a proton. In terms of M and R we have

$$v = \frac{8\pi m_p R^3}{3M} \quad (11.38)$$

and

$$x_F = \frac{\hbar}{m_e c} \frac{1}{R} \left(\frac{9\pi M}{8 m_p} \right)^{1/3} \equiv \frac{\bar{M}^{1/3}}{\bar{R}} \quad (11.39)$$

where

$$\bar{M} = \frac{9\pi M}{8 m_p}$$

$$\bar{R} = \frac{R}{(\hbar/m_e c)} \quad (11.40)$$

The pressure exerted by the Fermi gas is

$$P_0 = -\frac{\partial E_0}{\partial V} = \frac{m_e^4 c^5}{\pi^2 \hbar^3} \left[-f(x_F) - \frac{\partial f(x_F)}{\partial x_F} v \frac{\partial x_F}{\partial v} \right]$$

$$= \frac{m_e^4 c^5}{\pi^2 \hbar^3} \left[\frac{1}{3} x_F^3 \sqrt{1+x_F^2} - f(x_F) \right] \quad (11.41)$$

The nonrelativistic and extreme relativistic limits of P_0 are given by

$$P_0 \approx \left(\frac{m_e^4 c^5}{15\pi^2 \hbar^3} \right) x_F^5 = \frac{4}{5} K \frac{\bar{M}^{5/3}}{\bar{R}^5} \quad (\text{nonrel.: } x_F \ll 1) \quad (11.42)$$

$$P_0 \approx \left(\frac{m_e^4 c^5}{12\pi^2 \hbar^3} \right) (x_F^4 - x_F^2) = K \left(\frac{\bar{M}^{4/3}}{\bar{R}^4} - \frac{\bar{M}^{2/3}}{\bar{R}^2} \right) \quad (\text{extreme rel.: } x_F \gg 1) \quad (11.43)$$

where

$$K = \frac{m_e c^2}{12\pi^2} \left(\frac{m_e c}{\hbar} \right)^3 \quad (11.44)$$

A qualitative plot of P_0 against R for fixed M is shown in Fig. 11.5. It is seen that, for small R , P_0 becomes smaller than what is expected on the basis of nonrelativistic dynamics.

The condition for equilibrium of the star may be obtained through the following argument. Let us first imagine that there is no gravitational interaction. Then the density of the system will be uniform, and external walls will be needed to keep the Fermi gas at a given density. The amount of work that an external

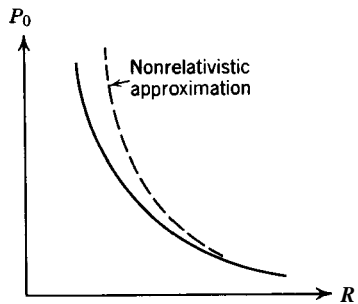


Fig. 11.5 Pressure of an ideal Fermi gas at absolute zero.

agent has to do to compress the star of given mass from a state of infinite diluteness to a state of finite density would be given by

$$-\int_{\infty}^R P_0 4\pi r^2 dr \quad (11.45)$$

where P_0 is the pressure of a uniform Fermi gas and R is the radius of the star. Now imagine that the gravitational interaction is "switched on." Different parts of the star will now attract one another, resulting in a decrease of the energy of the star by an amount that is called the gravitational self-energy. On dimensional grounds the gravitational self-energy must have the form

$$-\frac{\alpha\gamma M^2}{R} \quad (11.46)$$

where γ is the gravitational constant and α is a pure number of the order of unity. The exact value of α depends on the functional form of the density as a function of spatial distance and cannot be determined by our argument. If R is the equilibrium radius of the star, the gravitational self-energy must exactly compensate the work done in bringing the star together. Hence

$$\int_{\infty}^R P_0 4\pi r^2 dr = -\frac{\alpha\gamma M^2}{R} \quad (11.47)$$

Differentiating (11.47) with respect to R we obtain the condition for equilibrium:

$$P_0 = \frac{\alpha}{4\pi} \frac{\gamma M^2}{R^4} = \frac{\alpha}{4\pi} \gamma \left(\frac{8m_p}{9\pi} \right)^2 \left(\frac{m_e c}{\hbar} \right)^4 \frac{\bar{M}^2}{R^4} \quad (11.48)$$

Strictly speaking, (11.47) merely defines α . Its physical content is furnished by the assumption that α is of the order of unity. We now determine the relation between M and R by inserting an appropriate expression for P_0 into (11.48). This will be done for the following three different cases:

- (a) Suppose the temperature of the electron gas is much higher than the Fermi temperature. Then the electron gas may be considered as an ideal Boltzmann gas, with

$$P_0 = \frac{kT}{v} = \frac{3kT}{8\pi m_p} \frac{M}{R^3}$$

Substitution of this into (11.48) yields the linear relation

$$R = \frac{2}{3} \alpha M \frac{m_p \gamma}{kT} \quad (11.49)$$

This case, however, is never applicable for a white dwarf star.

- (b) Suppose the electron gas is at such a low density that nonrelativistic dynamics may be used ($x_F \ll 1$). Then P_0 is given by (11.42), and (11.48) leads to the equilibrium condition

$$\frac{4}{3} K \frac{\bar{M}^{5/3}}{R^5} = K' \frac{\bar{M}^2}{R^4}$$

where

$$K' = \frac{\alpha}{4\pi} \gamma \left(\frac{8m_p}{9\pi} \right)^2 \left(\frac{m_e c}{\hbar} \right)^4 \quad (11.50)$$

Thus the radius of the star decreases as the mass of the star increases:

$$\bar{M}^{1/3} \bar{R} = \frac{4}{5} \frac{K}{K'} \quad (11.51)$$

This condition is valid when the density is low. Hence it is valid for small M and large R .

- (c) Suppose the electron gas is at such a high density that relativistic effects are important ($x_F \gg 1$). Then P_0 is given by (11.43). The equilibrium condition becomes

$$K \left(\frac{\bar{M}^{4/3}}{\bar{R}^4} - \frac{\bar{M}^{2/3}}{\bar{R}^2} \right) = K' \frac{\bar{M}^2}{\bar{R}^4} \quad (11.52)$$

or

$$\bar{R} = \bar{M}^{2/3} \sqrt{1 - (\bar{M}/\bar{M}_0)^{2/3}} \quad (11.53)$$

where

$$\bar{M}_0 = \left(\frac{K}{K'} \right)^{3/2} = \left(\frac{27\pi}{64\alpha} \right)^{3/2} \left(\frac{\hbar c}{\gamma m_p^2} \right)^{3/2} \quad (11.54)$$

Numerically,

$$\frac{\hbar c}{\gamma m_p^2} \approx 10^{39} \quad (11.55)$$

This interesting pure number is the rest energy of X divided by the gravitational attraction of two protons separated by the Compton wavelength of X , where X is anything. The mass M_0 corresponding to the reduced quantity \bar{M}_0 is (taking $\alpha \approx 1$):

$$M_0 = \frac{8}{9\pi} m_p \bar{M}_0 \approx 10^{33} \text{ g} \approx M_{\odot} \quad (11.56)$$

the mass of the sun. The formula (11.53) is valid for high densities or for $R \rightarrow 0$.

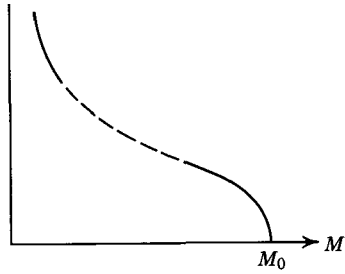


Fig. 11.6 Radius-mass relationship of a white dwarf star.

Hence it is valid for M near M_0 . Our model yields the remarkable prediction that no white dwarf star can have a mass larger than M_0 , because otherwise (11.53) would give an imaginary radius. The physical reason underlying this result is that if the mass is greater than a certain amount, the pressure coming from the Pauli exclusion principle is not sufficient to support the gas against gravitational collapse.

The radius-mass relationship of a white dwarf star, according to our model, has the form shown in Fig. 11.6, where the solid lines indicate the regions covered by formulas (11.51) and (11.53). We have not been able to calculate a , so that an exact value of M_0 cannot be obtained. More refined considerations* give the result

$$M_0 = 1.4M_{\odot} \quad (11.57)$$

This mass is known as the Chandrasekhar limit. Thus according to our model no star can become a white dwarf unless its mass is less than $1.4M_{\odot}$. This conclusion has so far been verified by astronomical observations. If the mass of a star is greater than the Chandrasekhar limit then it will eventually collapse under its own gravitational attraction. When the density becomes so high that new interactions, dormant thus far, are awakened, a new regime takes over. For example, the star could explode as a supernova.

11.3 LANDAU DIAMAGNETISM

Van Leeuwen's theorem[†] states that the phenomenon of diamagnetism is absent in classical statistical mechanics. Landau[‡] first showed how diamagnetism arises from the quantization of the orbits of charged particles in a magnetic field.

The magnetic susceptibility per unit volume of a system is defined to be

$$\chi \equiv \frac{\partial \mathcal{M}}{\partial H} \quad (11.58)$$

*S. Chandrasekhar, *Stellar Structure* (Dover, New York, 1957), Chapter XI

[†]See Problem 8.7.

[‡]L. Landau, *Z. Phys.* 64, 629 (1930).

where \mathcal{M} is the average induced magnetic moment per unit volume of the system along the direction of an external magnetic field H :

$$\mathcal{M} \equiv \frac{1}{V} \left\langle - \frac{\partial \mathcal{H}}{\partial H} \right\rangle \quad (11.59)$$

where \mathcal{H} is the Hamiltonian of the system in the presence of an external magnetic field H . For weak fields the Hamiltonian \mathcal{H} depends on H linearly. In the canonical ensemble we have

$$\mathcal{M} = kT \frac{\partial}{\partial H} \frac{\log Q_N}{V} \quad (11.60)$$

and in the grand canonical ensemble we have

$$\mathcal{M} = kT \frac{\partial}{\partial H} \left(\frac{\log \mathcal{Q}}{V} \right)_{T, V, z} \quad (11.61)$$

where z is to be eliminated in terms of N by the usual procedure.

A system is said to be diamagnetic if $\chi < 0$; paramagnetic if $\chi > 0$. To understand diamagnetism in the simplest possible terms, we construct an idealized model of a physical substance that exhibits diamagnetism. The magnetic properties of a physical substance are mainly due to the electrons in the substance. These electrons are either bound to atoms or nearly free. In the presence of an external magnetic field two effects are important for the magnetic properties of the substance: (a) The electrons, free or bound, move in quantized orbits in the magnetic field. (b) The spins of the electrons tend to be aligned parallel to the magnetic field. The atomic nuclei contribute little to the magnetic properties except through their influence on the wave functions of the electrons. They are too massive to have significant orbital magnetic moments, and their intrinsic magnetic moments are about 10^{-3} times smaller than the electron's. The alignment of the electron spin with the external magnetic field gives rise to paramagnetism, whereas the orbital motions of the electrons give rise to diamagnetism. In a physical substance these two effects compete. We completely ignore paramagnetism for the present, however. The effect of atomic binding on the electrons is also ignored. Thus we consider the idealized problem of a free spinless electron gas in an external magnetic field.

Landau Levels

The Hamiltonian of a nonrelativistic electron in an external magnetic field is

$$\mathcal{H} = \frac{1}{2m} \left(\mathbf{p} + \frac{e}{c} \mathbf{A} \right)^2 \quad (11.62)$$

where e is positive, (i.e., the charge of the electron is $-e$). The Schrodinger

equation $\mathcal{H}\psi = \epsilon\psi$ is invariant under the gauge transformation

$$\begin{aligned} \mathbf{A}(\mathbf{r}) &\rightarrow \mathbf{A}(\mathbf{r}) - \nabla\omega(\mathbf{r}) \\ \psi(\mathbf{r}) &\rightarrow \exp\left[-\frac{ie}{\hbar c}\omega(\mathbf{r})\right]\psi(\mathbf{r}) \end{aligned} \quad (11.63)$$

where $\omega(\mathbf{r})$ is an arbitrary continuous function. We consider a uniform external magnetic field \mathbf{H} pointing along the z axis, and choose the vector potential, via a gauge transformation if necessary, such that

$$A_x = -Hy, \quad A_y = A_z = 0 \quad (11.64)$$

This is called "choosing the gauge." The Hamiltonian then reads

$$\mathcal{H} = \frac{1}{2m} \left\{ [p_x - (eH/c)y]^2 + p_y^2 + p_z^2 \right\} \quad (11.65)$$

We solve the Schrodinger equation by assuming a wave function of the form

$$\psi(x, y, z) = e^{i(k_x x + k_z z)} f(y) \quad (11.66)$$

Then $f(y)$ satisfies the equation for a harmonic oscillator:

$$\begin{aligned} \left[\frac{1}{2m} p_y^2 + \frac{1}{2} m \omega_0^2 (y - y_0)^2 \right] f(y) &= \epsilon' f(y) \\ \omega_0 &= eH/mc, \quad y_0 = (\hbar c/eH) k_x \end{aligned} \quad (11.67)$$

where $\epsilon' = \epsilon - \hbar^2 k_z^2 / 2m$. The natural frequency of the harmonic oscillator ω_0 is the "cyclotron frequency," that of a classical charge moving in a circular orbit normal to a uniform magnetic field. The energy eigenvalues are thus

$$\epsilon(p_z, j) = \frac{p_z^2}{2m} + \hbar\omega_0 \left(j + \frac{1}{2} \right), \quad (j = 0, 1, 2, \dots) \quad (11.68)$$

where $p_z = \hbar k_z$. These are the Landau energy levels. Since they are independent of k_x , they have a degeneracy equal to the number of allowed values of k_x , such that y_0 lies within the container of the system.

Let us put the system in a large cube of size L , and impose periodic boundary conditions. The allowed values of k_x are of the form $2\pi n_x / L$, where $n_x = 0, \pm 1, \pm 2, \dots$. For y_0 to lie between 0 and L , the values of n_x must be positive and bounded by

$$g = (eH/\hbar c) L^2 \quad (11.69)$$

which is the degeneracy of a Landau level. The proportionality to L^2 reflects the fact that the projection of the electron orbit onto the xy plane can be centered anywhere in the plane without changing the energy. Thus, when the external field is turned on, the energy spectrum associated with the motion in the xy plane changes from a continuous spectrum to a discrete one, and the level spacing and degeneracy increases with the external field. This is illustrated in Fig. 11.7.

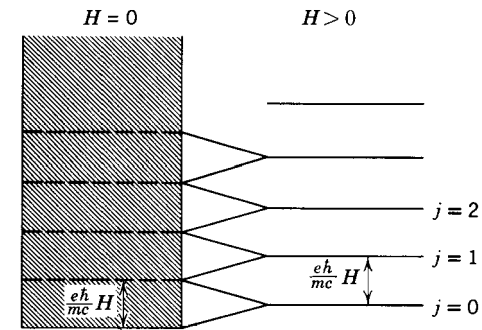


Fig. 11.7 Comparison of the energy spectra of a charged particle with and without magnetic field.

Flux Quantization

The Landau levels and the degeneracies derived above are all we need to calculate the partition function. However, we take the opportunity to discuss flux quantization briefly, to help us better understand the wave functions.

Consider a plane with a hole in it, which contains a certain amount of magnetic flux Φ , as shown in Fig. 11.8. Suppose there is no magnetic field anywhere else. Then the vector potential in the plane must be "pure gauge," i.e., of the form

$$\mathbf{A} = \nabla\omega$$

We cannot transform this to zero through any continuous gauge transformation, because necessarily

$$\int_C d\mathbf{S} \cdot \mathbf{A} = \Phi \quad (11.70)$$

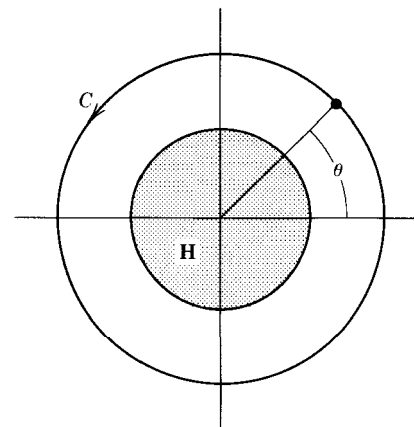


Fig. 11.8 Charged particle moving in plane with a hole containing magnetic flux. The particle will not notice the flux if either (a) it is in a localized state, or (b) the flux is quantized in units of hc/e .

where the closed path C encloses the hole, as indicated in Fig. 11.7. A solution to (11.70) is

$$\omega = \Phi\theta/2\pi \quad (11.71)$$

where θ is the angle around the hole, measured from some arbitrary axis.

Now consider an electron moving in the plane, with the boundary condition that its wave function vanish in the hole. In general it is affected by the flux, because the Schrodinger equation involves \mathbf{A} , which is nonzero where the electron moves. But since \mathbf{A} is pure gauge, we are tempted to try to remove it from the Schrodinger equation through the gauge transformation

$$\mathbf{A} \rightarrow \mathbf{A} - \nabla\omega$$

In so doing, the wave function of the electron acquires the phase factor

$$\exp\left(-\frac{ie\omega}{\hbar c}\right) = \exp\left[-i\theta\left(\frac{e\Phi}{\hbar c}\right)\right] \quad (11.72)$$

which is generally unacceptable because it will render the wave function discontinuous in space (for θ increases by 2π each time we go around the hole). The objection is circumvented under either of the following circumstances:

- The electron is "localized," i.e., its wave function is nonvanishing only in the neighborhood of some point. In this case, where (11.72) might lead to a discontinuity, the wave function vanishes anyway. This is not relevant to free electrons, but may be relevant, for example, for an electron trapped by an impurity in a metal.
- The electron is "extended," with a wave function that is phase-coherent around a closed path about the hole, but the flux is quantized in integer multiples of the flux quantum

$$\Phi_0 = hc/e \quad (11.73)$$

In this case, (11.72) becomes a periodic function of θ and represents a legitimate gauge transformation. Thus, the vector potential can be transformed away, and the electron does not "know" there is flux through the hole. This is the basis of the Aharonov-Bohm effect,* and the flux quantization in superconductivity,† which have been experimentally verified.

The relevance of flux quantization to the Landau levels lies in the fact that the degeneracy (11.69) is just the total magnetic flux measured in units of the flux quantum:

$$g = \Phi/\Phi_0 \quad (11.74)$$

*Y. Aharonov and D. Bohm, *Phys. Rev.* 115, 465 (1959).

†See P. G. De Gennes, *Superconductivity of Metal and Alloys* (Benjamin, New York, 1966), p. 149.

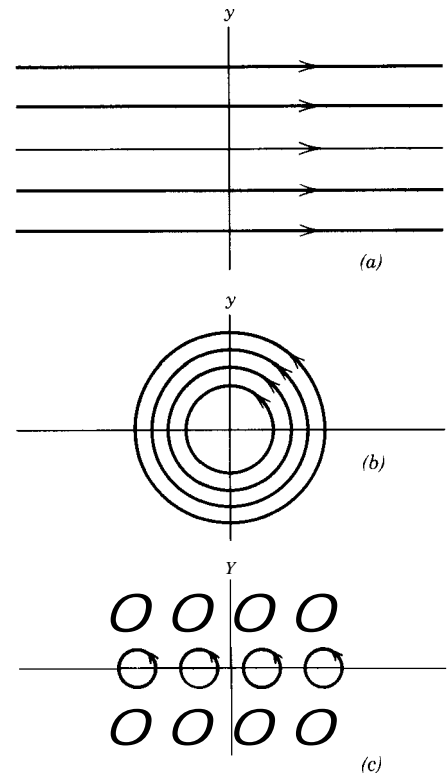


Fig. 11.9 Different bases for electron states in a Landau level, which is highly degenerate: (a) Member wave functions are peaked at different elevations (y direction), and are eigenstates of momentum in the x direction. (b) Member wave functions are eigenstates of orbital angular momentum. They are peaked at concentric circles, with equal areas between successive circles. Each ring between circles supports one magnetic flux quantum. (c) Member wave functions are "vortices" of flux quanta, forming a lattice.

We have been working in a gauge in which the wave functions have definite linear momentum in the x direction. Thus the probability densities are independent of x , and peaked about parallel ridges at $y = y_0$. The spacing between successive values of y_0 is hc/eHL , and hence the area of the strip in the xy plane between two neighboring ridges is hc/eH . Thus, exactly one flux quantum hc/e goes through the strip.

Since the energy levels are highly degenerate, we can make linear transformations on the wave functions belonging to the same Landau level to obtain equivalent sets. Such transformations are equivalent to gauge transformations. We can make them eigenfunctions of angular momentum about the z axis, in which case the probability distributions will be peaked about circles in the xy plane about the z axis, with exactly one quantum of flux going through the annular ring between two successive circles. We can also make them into individual orbitals, whose centers form a regular lattice in the xy plane. In this case each orbital will link exactly one flux quantum (a "vortex.") These different bases are illustrated in Fig. 11.9.

The qualitative fact relevant to our immediate purpose is that the flux quantum sets a finite minimum size of an orbit, and thereby provides the escape from Van Leeuwen's theorem.

Magnetic Susceptibility

The grand partition function is

$$\mathcal{Q} = \prod_{\lambda} (1 + ze^{-\beta\epsilon_{\lambda}}) \quad (11.75)$$

where λ denotes the set of quantum numbers $\{p_z, j, a\}$, with $a = 1, \dots, g$. Thus

$$\begin{aligned} \log \mathcal{Q} &= \sum_{a=1}^g \sum_{j=0}^{\infty} \sum_{p_z} \log [1 + ze^{-\beta\epsilon(p_z, j)}] \\ &= \frac{2zgL}{h} \sum_{j=0}^{\infty} \int_0^m dp \log [1 + ze^{-\beta\epsilon(p, j)}] \end{aligned} \quad (11.76)$$

The average number of electrons is

$$N = \frac{2gL}{h} \sum_{j=0}^{\infty} \int_0^m dp \frac{1}{z^{-1}e^{\beta\epsilon(p, j)} + 1} \quad (11.77)$$

To calculate the magnetization in the classical domain we take the high-temperature limit. The condition (11.76) requires that $z \rightarrow 0$ to keep N finite. Thus we expand the above equations in powers of z , and retain only the first-order term:

$$\begin{aligned} \log \mathcal{Q} &\approx \frac{2zgL}{h} \sum_{j=0}^{\infty} \int_0^m dp e^{-\beta(p^2/2m + \hbar\omega_0(j+1/2))} \\ &= \frac{zgL}{\lambda} \frac{e^{-x}}{1 - e^{-2x}} \end{aligned} \quad (11.78)$$

where $\lambda = \sqrt{2\pi\hbar^2/mkT}$ and $x = \hbar\omega_0/2kT$. We keep only the lowest-order contribution in x :

$$\log \mathcal{Q} \approx \frac{zgL}{\lambda} \frac{1}{2x} \left(1 - \frac{x^2}{6}\right) = \frac{zV}{\lambda^3} \left[1 - \frac{1}{24} \left(\frac{\hbar\omega_0}{kT}\right)^2\right] \quad (11.79)$$

from which follows

$$\chi \approx -\frac{z}{3kT\lambda^3} \left(\frac{e\hbar}{2mc}\right)^2 \quad (11.80)$$

To eliminate z , we note from (11.77) that to first order in z , N is the same as $\log \mathcal{Q}$. Hence

$$\frac{N}{V} \approx \frac{z}{\lambda^3} \quad (11.81)$$

Solving for z and substituting the result into (11.80), we obtain the final answer

$$\chi \approx -\frac{1}{3kTv} \left(\frac{e\hbar}{2mc}\right)^2 \quad (11.82)$$

which conforms to Curie's $1/T$ law. Since the lowest energy of an electron is

$$\hbar\omega_0/2 = (e\hbar/2mc)H$$

we see that the magnetic moment of the minimal orbit is just the Bohr magneton $e\hbar/2mc$.

11.4 THE DE HAAS-VAN ALPHEN EFFECT

We now turn to the low-temperature limit of electrons in Landau levels. The electrons will tend to occupy the lowest available levels. As the magnetic field is decreased, each Landau level can accommodate fewer electrons because the degeneracy is decreased. Consequently, some electrons will be forced to jump up to a higher level. This causes the de Haas-Van Alphen effect, the oscillation of the low-temperature magnetic susceptibility as the magnetic field is decreased. To study this effect in a simple context we shall assume $kT \ll \hbar\omega_0$, so that we can set $T = 0$. We shall also ignore the motion in the z direction.*

Our problem is to calculate the ground state energy of a two-dimensional electron system of total area L^2 in a uniform magnetic field H . We write the Landau levels ϵ_j' and their degeneracy g in the following notation:

$$\begin{aligned} \epsilon_j &= 2\mu_0 H(j + \frac{1}{2}), & \mu_0 &= e\hbar/2mc \\ g &= NH/H_0, & H_0 &= nhc/e \end{aligned} \quad (11.83)$$

where $n = N/L^2$ is the number of electrons per unit area. The field H_0 is the value of H above which the Landau level can hold all the N particles.

If $H/H_0 > 1$, then all particles can be accommodated in the lowest Landau level, and the ground state energy per particle is

$$E_0/N = \mu_0 H, \quad (H/H_0 > 1) \quad (11.84)$$

If $H < H_0$, then some particles will have to occupy higher levels. Suppose H is such that the j lowest Landau levels are completely filled, the $(j+1)$ th level is partially filled, and all higher levels are empty. The condition for H is

$$(j+1)g < N < (j+2)g$$

or

$$\frac{1}{j+2} < \frac{H}{H_0} < \frac{1}{j+1} \quad (11.85)$$

For H in this interval,

$$\begin{aligned} E_0/N &= g \sum_{i=0}^j \epsilon_i + [N - (j+1)g] \epsilon_{j+1} \\ &= \mu_0 (H/H_0) [2j + 3 - (j+1)(j+2)(H/H_0)] \end{aligned} \quad (11.86)$$

*The experimental effect was discovered by W. J. De Haas and P. M. Van Alphen, *Leiden Commun.*, 212 (1931). Our simplified model is that of R. E. Peierls, *Z. Phys.* **81**, 186 (1933). For a more realistic treatment see J. M. Luttinger, *Phys. Rev.* **121**, 1251 (1961).

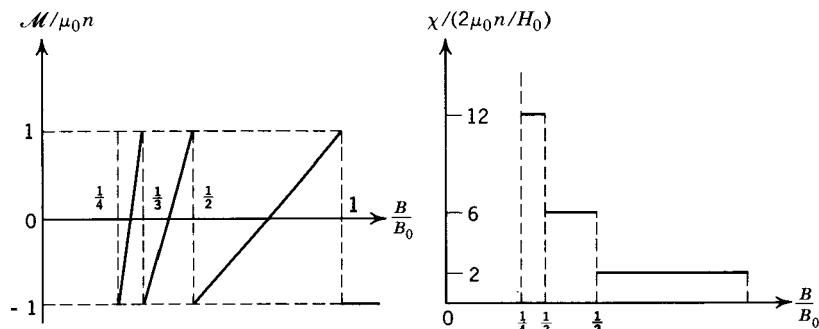


Fig. 11.10 De Haas-Van Alphen effect.

Introducing the parameter

$$x = H/H_0 \quad (11.87)$$

we can summarize the results as follows:

$$\frac{1}{N} E_0(B) = \begin{cases} \mu_0 H_0 x & (x > 1) \\ \mu_0 H_0 x [(2j+3) - (j+1)(j+2)x] & \left(\frac{1}{j+2} < x < \frac{1}{j+1}, j = 0, 1, 2, \dots \right) \end{cases} \quad (11.88)$$

The magnetization per unit volume and the magnetic susceptibility per unit volume are respectively given by

$$\mathcal{M} = \begin{cases} -\mu_0 n & (x > 1) \\ \mu_0 n [2(j+1)(j+2)x - (2j+3)] & \left(\frac{1}{j+2} < x < \frac{1}{j+1}, j = 0, 1, 2, \dots \right) \end{cases} \quad (11.89)$$

$$\chi = \begin{cases} 0 & (x > 1) \\ \frac{2\mu_0 n}{H_0} (j+1)(j+2) & \left(\frac{1}{j+2} < x < \frac{1}{j+1}, j = 0, 1, 2, \dots \right) \end{cases} \quad (11.90)$$

These are shown in Fig. 11.10

11.5 THE QUANTIZED HALL EFFECT

The Hall effect was discovered in the nineteenth century: When crossed magnetic and electric fields are applied to a metal, a voltage is induced in a direction orthogonal to the crossed fields, as evidenced by an induced current flowing in

that direction—the Hall current. This effect is easy to understand on the basis of the free electron theory of a metal, as follows. Crossed magnetic and electric fields, denoted, respectively, by \mathbf{H} and \mathbf{E} , act as velocity filters to free charges, letting through only those whose velocity \mathbf{v} is such that $\mathbf{E} + (\mathbf{v}/c)\mathbf{B} = 0$, or

$$\frac{v}{c} = \frac{E}{B} \quad (11.91)$$

For free charge carriers in a metal, the current density is

$$\mathbf{j} = qn\mathbf{v} \quad (11.92)$$

where q is the charge, and n the density. The Hall resistivity ρ_{xy} is defined as the ratio of the electric field (in the y direction) to the Hall current density (in the x direction):

$$j = \frac{E}{\rho_{xy}} \quad (11.93)$$

Substituting this into (11.92) and then into (11.91), we obtain

$$\rho_{xy} = \frac{H}{qnc} \quad (11.94)$$

Measurements of the Hall resistivity in various metals has yielded charge carrier densities and provided the first demonstrations that there are not only negative charge carriers (electrons), but also positive ones (holes).

The two-dimensional electron system used as a model in the last section can now be created in the laboratory, thanks to developments in the transistor technology. It can be made by injecting electrons into the interface of an alloy sandwich, which confines the electrons in a thin film about 500 Å thick. The Hall experiment has been performed on such two-dimensional electron systems at very low temperatures, and the direct resistivities ρ_{xx} and the Hall resistivities ρ_{xy} have been measured, as indicated in Fig. 11.11.

The experimental results are quite dramatic, as shown in Fig. 11.12. As the magnetic field H increases the degeneracy of the Landau levels increases. Since the electron density does not depend on the field the filling fraction ν of the lowest Landau level decreases:

$$\nu \equiv \frac{hcn}{eH} \quad (11.95)$$

The Hall resistivity exhibits plateaus at $\nu = 1, \frac{2}{3}, \frac{1}{3}$, with values equal to $1/\nu$, in units of h/e^2 . At the same time, the conventional resistivity ρ_{xx} drops to very low values. This indicates that in the neighborhood of these special filling fractions the two-dimensional electron fluid flows with almost no resistance. The value at $\nu = 1$, called the integer quantized Hall effect, was first observed in a MOSFET (metal-oxide semiconductor field-effect transistor) at $T = 1.5$ K. The Hall resis-

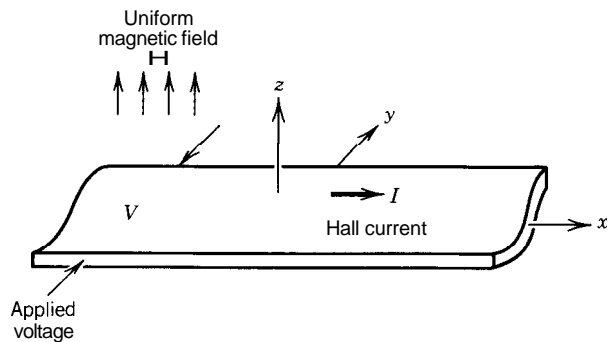


Fig. 11.11 The Hall effect. A current I flows in a direction orthogonal to crossed electric and magnetic fields. The Hall resistivity is defined as $\rho_{xy} = V/I$. The conventional resistivity ρ_{xx} can be obtained by measuring the voltage drop along the direction of the current.

tivity was found to be quantized with a precision of one part in 10^5 .^{*} The fractional values were found soon after.?

The integer effect is easy to understand on a naive basis. Since at $\nu = 1$ the lowest Landau level is completely filled, there is an energy gap above the Fermi level. Low-energy excitations are therefore impossible, and so the centers of the electron orbits flow like a free gas. Using (11.94) with $n = eH/hc$, the Landau degeneracy per unit area, we immediately obtain the desired result.

$$\rho_{xy} = \frac{h}{e^2} \quad (11.96)$$

But this does not explain why the Hall resistivity continues to be quantized even when the field is changed somewhat, so that there is a plateau of the quantized value, as seen in the data.

Laughlin^{*} offers the important insight that the integer effect is due to the phase coherence of the electronic wave function over the entire sample, and that the effect of impurities are important in producing the observed plateau. Consider a sample in the form of a ribbon forming a closed loop, as shown in Fig. 11.13. A magnetic field H pierces the ribbon everywhere normal to its surface, and a voltage V is applied across the edges of the ribbon. Our object is to deduce the relation between the Hall current I and V .

The Hall current produces a magnetic moment $\mu = IA/c$, where A is the area enclosed by the ribbon loop. Imagine that a small amount of magnetic flux $\delta\Phi$ is introduced through the loop, corresponding to an increase in the magnetic

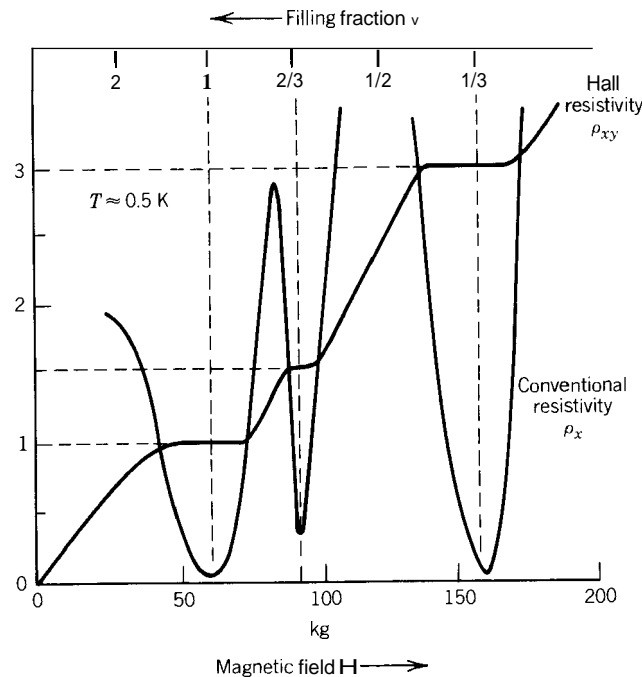


Fig. 11.12 Quantized Hall effect: Schematic representation of experimental data. The filling fraction ν is the fraction of degenerate states in the lowest Landau levels occupied by electrons. The Hall resistivity exhibits plateau of value $1/\nu$, at $\nu = 1, \frac{2}{3}, \frac{1}{3}$ (in units of h/e^2 .) The conventional resistivity becomes very small at these values. The quantization is accurate to at least one part in 10^4 .

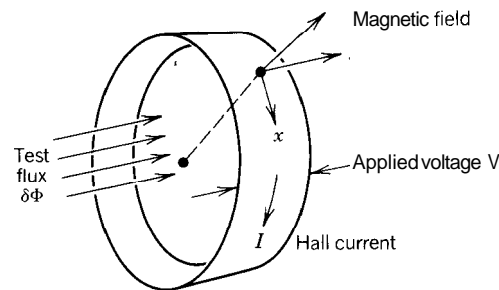


Fig. 11.13 Hall effect in idealized geometry.

^{*}K. V. Klitzing, G. Dorda, and M. Pepper, *Phys. Rev. Lett.* 45, 494 (1980).

[†]D. C. Tsui, H. L. Stormer, and A. C. Gossard, *Phys. Rev. Lett.* 48, 1559 (1982).

^{*}R. B. Laughlin, *Phys. Rev. B* 23, 5632 (1981). See also B. I. Halperin, *Phys. Rev. B* 25, 2185 (1982).

field $\delta H = \delta\Phi/A$ normal to the plane of the loop. The energy of the system increases by $\delta E = \mu \delta H = (IA/c)(\delta\Phi/A)$. Hence we can find the current from the formula

$$I/c = \frac{\delta E}{\delta\Phi} \tag{11.97}$$

We recall from our discussion of flux quantization in Section 11.3 that the "localized" electrons will not respond to the flux, but the "extended" ones may. Electrons in Landau levels do have extended wave functions, and thus will respond to the flux and contribute to the Hall current. Across the ribbon, in the y direction, the wave function of an electron is peaked about some value of y , say y_0 . The allowed values of y_0 extends from one edge of the ribbon to the other. (We are using here the "strip" representation of the wave functions, as shown in Fig. 11.9a.)

Consider now a completely filled Landau level. The electron density across the ribbon may be represented schematically as in Fig. 11.14. The electrons lying closer to the right edge have a higher electrostatic energy because of the applied voltage. Now imagine that the flux through the loop is increased slowly from zero. The electrons will respond to the change until the flux reaches the quantum value hc/e , at which point they cannot feel the flux. During the slow increase, the energy of the electrons must rise by the transfer of electrons from one edge of the ribbon to the other. When the flux reaches one quantum, the electron distribution must look exactly the same as before. Overall, therefore, the electrons play musical chairs, moving up one position per quantum of flux penetration, as indicated in Fig. 11.14. Since the gain in energy is $\delta E = eV$, and the change in flux is $\delta\Phi = hc/e$, we have from (11.97) $\mathbf{I} = (e^2/h)V$, whence

$$\rho_{xy} = \frac{h}{e^2} \tag{11.98}$$

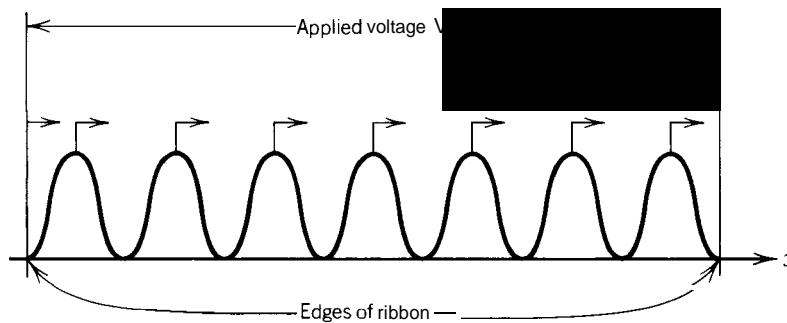


Fig. 11.14 Schematic representation of electron density across the ribbon in Fig. 11.13, when the lowest Landau level is completely filled. The electrons move to the right by one "musical chair," when one unit of test flux pushes through the loop in Fig. 11.13.

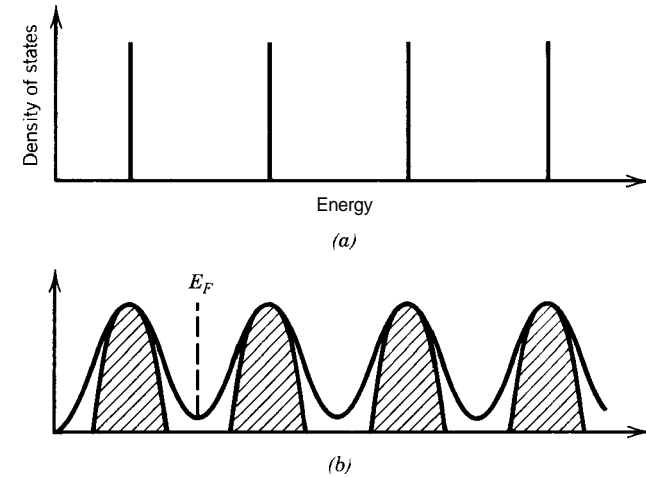


Fig. 11.15 Density of state of an electron in a magnetic field. (a) In a pure metal, a series of delta functions mark the positions of the Landau levels. (b) In the presence of impurities, the Landau levels broaden to bands (shaded region). Localized electrons states fill the gap between Landau bands. E_F denotes the Fermi level.

If the total number of electrons is fixed, then changing the magnetic field will change the filling fraction. A filled level will either become underfilled, or it will spill electrons over to a higher level. In either case, the previous analysis breaks down. However, in a physical sample there are always impurities that trap electrons into localized states. It is the presence of impurities that give rise to the stability of the effect, as shown by the plateau in the data.

In Fig. 11.15 we show the density of states in a pure sample as compared with one with impurities. In a pure sample we have a series of 6 functions at the Landau levels, while in the presence of impurities each Landau level is broadened to a band, shown shaded in Fig. 11.15. At low impurity densities these bands do not overlap. The important point is that between the Landau bands the density of states is not zero, as would be the case for a pure sample, but is filled by contributions from localized states. The Fermi level can lie in a continuum between Landau bands, and it can shift in response to a change in the occupancy of the filled Landau band, so that the band beneath it remains filled. Thus, for a certain range of the external magnetic field the lowest Landau band remains completely filled, and our argument applies.

The fractional effects are more intriguing. What accounts for the stability of the electron film when the Landau level is only one-third full? The answer must lie in the Coulomb interaction among the electrons, but so far we only have preliminary guesses.*

*R. B. Laughlin, *Phys. Rev. Lett.* 50, 1395 (1983).

11.6 PAULI PARAMAGNETISM

The Hamiltonian of a nonrelativistic free electron in an external magnetic field \mathbf{B} is given by

$$\frac{1}{2m} \left(\mathbf{p} + \frac{e}{c} \mathbf{A} \right)^2 - \mu_0 \boldsymbol{\sigma} \cdot \mathbf{H} \quad (11.99)$$

where $\mu_0 = e\hbar/2mc$, and \mathbf{a} are the Pauli spin matrices. The first term gives rise to diamagnetism, as we have studied. The second term gives rise to paramagnetism. We now consider its effect alone, and take the single-particle Hamiltonian to be*

$$\mathcal{H} = \frac{p^2}{2m} - \mu_0 \boldsymbol{\sigma} \cdot \mathbf{H} \quad (11.100)$$

The eigenvalues of $\mathbf{a} \cdot \mathbf{H}$ are sH , where $s = \pm 1$. Hence the single-particle energy levels are

$$\epsilon_{\mathbf{p},s} = \frac{p^2}{2m} - s\mu_0 H \quad (11.101)$$

An energy eigenvalue of the N -particle system may be labeled by the occupation numbers $n_{\mathbf{p},s}$ of the single-particle levels $\epsilon_{\mathbf{p},s}$:

$$E_n = \sum_{\mathbf{p}} \sum_s \epsilon_{\mathbf{p},s} n_{\mathbf{p},s} = \sum_{\mathbf{p}} \left[\left(\frac{p^2}{2m} - \mu_0 H \right) n_{\mathbf{p},+1} + \left(\frac{p^2}{2m} + \mu_0 H \right) n_{\mathbf{p},-1} \right] \quad (11.102)$$

where

$$\begin{aligned} n_{\mathbf{p},s} &= 0, 1 \\ \sum_s \sum_{\mathbf{p}} n_{\mathbf{p},s} &= N \end{aligned} \quad (11.103)$$

Let

$$\begin{aligned} n_{\mathbf{p},+1} &\equiv n_{\mathbf{p}}^+ \\ n_{\mathbf{p},-1} &\equiv n_{\mathbf{p}}^- \\ \sum_{\mathbf{p}} n_{\mathbf{p},+1} &\equiv N_+ \\ \sum_{\mathbf{p}} n_{\mathbf{p},-1} &\equiv N_- = N - N_+ \end{aligned} \quad (11.104)$$

Then an energy eigenvalue of the system can also be written in the form

$$E_n = \sum_{\mathbf{p}} (n_{\mathbf{p}}^+ + n_{\mathbf{p}}^-) \frac{p^2}{2m} - \mu_0 H (N_+ - N_-) \quad (11.105)$$

The partition function is

$$Q_N = \sum'_{\{n_{\mathbf{p}}^+\}, \{n_{\mathbf{p}}^-\}} \exp \left[-\beta \sum_{\mathbf{p}} (n_{\mathbf{p}}^+ + n_{\mathbf{p}}^-) \frac{p^2}{2m} + \beta \mu_0 H (N_+ - N_-) \right] \quad (11.106)$$

where the prime over the sum denotes the restrictions (11.103). The sum can be evaluated as follows. First we choose an arbitrary integer N_+ and sum over all sets $\{n_{\mathbf{p}}^+\}, \{n_{\mathbf{p}}^-\}$ such that $\sum_{\mathbf{p}} n_{\mathbf{p}}^+ = N_+$, and $\sum_{\mathbf{p}} n_{\mathbf{p}}^- = N - N_+$. Then we sum over all integers N_+ from 0 to N . In this manner we arrive at the formula

$$Q_N = \sum_{N_+=0}^N e^{\beta \mu_0 H (2N_+ - N)} \sum''_{\{n_{\mathbf{p}}^+\}} \exp \left(-\beta \sum_{\mathbf{p}} \frac{p^2}{2m} n_{\mathbf{p}}^+ \right) \sum'''_{\{n_{\mathbf{p}}^-\}} \exp \left(-\beta \sum_{\mathbf{p}} \frac{p^2}{2m} n_{\mathbf{p}}^- \right) \quad (11.107)$$

where \sum'' is subject to the restriction $\sum_{\mathbf{p}} n_{\mathbf{p}}^+ = N_+$, and \sum''' is subject to the restriction $\sum_{\mathbf{p}} n_{\mathbf{p}}^- = N_- = N - N_+$. Let $Q_N^{(0)}$ denote the partition function of the ideal Fermi gas of N spinless particles of mass m :

$$Q_N^{(0)} \equiv \sum_{\sum n_{\mathbf{p}} = N} \exp \left(-\beta \sum_{\mathbf{p}} \frac{p^2}{2m} n_{\mathbf{p}} \right) \equiv e^{-\beta A(N)} \quad (11.108)$$

Then

$$\begin{aligned} Q_N &= e^{-\beta \mu_0 H N} \sum_{N_+=0}^N e^{2\beta \mu_0 H N_+} Q_{N_+}^{(0)} Q_{N-N_+}^{(0)} \\ \frac{1}{N} \log Q_N &= -\beta \mu_0 H + \frac{1}{N} \log \sum_{N_+=0}^N e^{2\beta \mu_0 H N_+ - \beta A(N_+) - \beta A(N-N_+)} \end{aligned} \quad (11.109)$$

There are $N + 1$ positive terms in the sum just given. The logarithm of this sum is equal to the logarithm of the largest term in the sum plus a contribution of the order of $\log N$. Therefore, neglecting a term of order $N^{-1} \log N$, we have

$$\frac{1}{N} \log Q_N = \beta f(\bar{N}_+) \quad (11.110)$$

where

$$\begin{aligned} f(\bar{N}_+) &= \max [f(N_+)] \\ f(N_+) &\equiv \mu_0 H \left(\frac{2N_+}{N} - 1 \right) - \frac{1}{N} [A(N_+) + A(N - N_+)] \end{aligned} \quad (11.111)$$

Obviously we can interpret \bar{N}_+ as the average number of particles with spin up. If \bar{N}_+ is known, the magnetization per unit volume can be obtained through the

*Following W. Pauli, *Z. Phys.* 41, 81 (1927)

formula

$$\mathcal{M} = \frac{\mu_0(2\bar{N}_+ - N)}{V} \quad (11.112)$$

We now explicitly find \bar{N}_+ . The condition (11.111) is equivalent to the condition*

$$\left[\frac{\partial f(N_+)}{\partial N_+} \right]_{N_+ = \bar{N}_+} = 0$$

$$2\mu_0 H - \left[\frac{\partial A(N')}{\partial N'} \right]_{N' = \bar{N}_+} - \left[\frac{\partial A(N - N')}{\partial N'} \right]_{N' + \bar{N}_+} = 0 \quad (11.113)$$

Let $kT\nu(N)$ be the chemical potential of an ideal Fermi gas of N spinless particles:

$$kT\nu(N) = \frac{\partial A(N)}{\partial N} \quad (11.114)$$

Then

$$\left[\frac{\partial A(N')}{\partial N'} \right]_{N' = \bar{N}_+} = kT\nu(\bar{N}_+)$$

$$\left[\frac{\partial A(N - N')}{\partial N'} \right]_{N' = \bar{N}_+} = - \left[\frac{\partial A(N - N')}{\partial (N - N')} \right]_{N - N' = N - \bar{N}_+} = -kT\nu(N - \bar{N}_+)$$

Thus (11.113) becomes

$$kT[\nu(\bar{N}_+) - \nu(N - \bar{N}_+)] = 2\mu_0 H \quad (11.115)$$

This condition states that at a given temperature the average number of particles with spin up is such that the chemical potential of the particles with spin up is greater than that of the particles with spin down by $2\mu_0 H$. We solve (11.115) in the low-temperature and high-temperature limits.

Let the Fermi energy for the present system be

$$\epsilon_F(N) \equiv \left(\frac{3\pi^2 N}{V} \right)^{2/3} \frac{\hbar^2}{2m} \quad (11.116)$$

In the low-temperature region ($kT \ll \epsilon_F$), we can use the expansion (11.24)[†] for $kT\nu(N)$:

$$kT\nu(N) = \epsilon_F(2N) \left\{ 1 - \frac{\pi^2}{12} \left[\frac{kT}{\epsilon_F(2N)} \right]^2 + \dots \right\}$$

*We should make sure that (11.113) determines a maximum and not a minimum and that \bar{N}_+ lies between 0 and N . It can be verified that (11.113) has only one real root that automatically satisfies these requirements.

[†]Note that in (11.24) the symbol ϵ_F stands for the Fermi energy of N spinless particles and does not have the same meaning as ϵ_F here.

Thus (11.115) becomes

$$\epsilon_F(2\bar{N}_+) - \epsilon_F(2N - 2\bar{N}_+) - \frac{\pi^2(kT)^2}{12}$$

$$\times \left[\frac{1}{\epsilon_F(2\bar{N}_+)} - \frac{1}{\epsilon_F(2N - 2\bar{N}_+)} \right] + \dots = 2\mu_0 H \quad (11.117)$$

Let

$$r \equiv \frac{2\bar{N}_+}{N} - 1 \quad (-1 \leq r \leq +1) \quad (11.118)$$

Then (11.117) becomes

$$(1+r)^{2/3} - (1-r)^{2/3} - \frac{\pi^2}{12} \left(\frac{kT}{\epsilon_F} \right)^2$$

$$\times \left[(1+r)^{-2/3} - (1-r)^{-2/3} \right] + \dots = \frac{2\mu_0 H}{\epsilon_F} \quad (11.119)$$

At absolute zero, r satisfies the equation

$$(1+r)^{2/3} - (1-r)^{2/3} = \frac{2\mu_0 H}{\epsilon_F} \quad (11.120)$$

This may be solved graphically, as shown in Fig. 11.16. For $\ll \epsilon_F/2\mu$ an approximate solution is

$$r \approx \frac{3\mu_0 H}{2\epsilon_F}$$

$$\bar{N}_+ \approx \frac{N}{2} \left(1 + \frac{3\mu_0 H}{2\epsilon_F} \right) \quad (11.121)$$

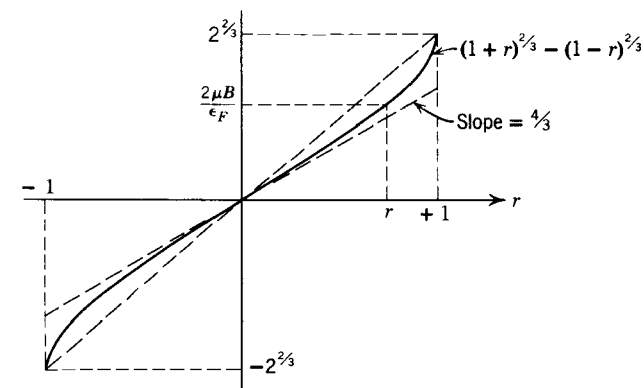


Fig. 11.16 Graphical solution of (11.122).

Thus when $H = 0$ half the particles have spin up, and the other half spin down. When $H > 0$ the balance shifts in favor of spin up. From (11.112) and (11.121) we obtain, for absolute zero,

$$\begin{aligned} \mathcal{M} &= \frac{\mu_0 r}{v} \approx \frac{3\mu_0^2}{2\epsilon_F v} \\ \chi &\approx \frac{3\mu_0^2 H}{2\epsilon_F v} \end{aligned} \quad (11.122)$$

For $0 < kT \ll \epsilon_F$ and $\mu \ll \epsilon_F$ we can solve (11.119) by expanding the left side in powers of r , and we obtain

$$\begin{aligned} r &\approx \frac{3\mu_0 H}{2\epsilon_F} \left[1 - \frac{\pi^2}{12} \left(\frac{kT}{\epsilon_F} \right)^2 \right] \\ \chi &\approx \frac{3\mu_0^2}{2\epsilon_F v} \left[1 - \frac{\pi^2}{12} \left(\frac{kT}{\epsilon_F} \right)^2 \right] \end{aligned} \quad (11.123)$$

For high temperatures ($kT \gg \epsilon_F$) we use (11.12):

$$\nu(N) \approx \log \left(\frac{N\lambda^3}{V} \right)$$

Hence (11.115) gives

$$\begin{aligned} \log \left[\frac{\lambda^3(1+r)}{v} \right] - \log \left[\frac{\lambda^3(1-r)}{v} \right] &= \frac{2\mu_0 H}{kT} \\ r &= \tanh \frac{\mu_0 H}{kT} \approx \frac{\mu_0 H}{kT} \end{aligned} \quad (11.124)$$

The magnetic susceptibility per unit volume is then given by

$$\chi \approx \frac{\mu_0^2}{kTv} \quad (11.125)$$

A qualitative plot of $kT\chi$ is shown in Fig. 11.17.

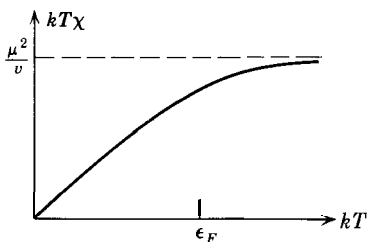


Fig. 11.17 Pauli paramagnetism.

11.7 MAGNETIC PROPERTIES OF AN IMPERFECT GAS

How would interparticle interaction affect the magnetic properties of an electron gas? Qualitatively speaking, the effect of a repulsive interaction will enhance the paramagnetism, because two electrons would prefer to be in a spatially antisymmetric wave function to minimize the repulsive energy. But an antisymmetric spatial wave function requires a symmetric spin wave function, which is a triplet state. Thus the repulsion tends to align the spins of the electrons.

We demonstrate this effect by considering an imperfect gas of spin- $\frac{1}{2}$ fermions at very low temperatures, so that the repulsive interaction can be characterized by a single parameter, the scattering length a , or effective hard-sphere diameter. To first order in a , the energy for an N -particle system is given by (A.41) in the Appendix. We take our model to be defined by (11.102) plus the interaction energy:

$$E_n = \sum_{\mathbf{p}} (n_{\mathbf{p}}^+ + n_{\mathbf{p}}^-) \frac{p^2}{2m} + \frac{4\pi a \hbar^2}{mV} N_+ N_- - (N_+ - N_-) \mu_0 H \quad (11.126)$$

The condition for the validity of this formula is that $k_F |a| \ll 1$ where k_F is the wave number of a particle at the Fermi level: $k_F^2 = (3\pi^2 n)^{2/3}$. Thus, the condition for validity is low density, i.e., $na^3 \ll 1$.

The partition function is

$$Q_N = \sum'_{\{n_{\mathbf{p}}^+\}, \{n_{\mathbf{p}}^-\}} \exp \left\{ -\beta \left[\sum_{\mathbf{p}} (n_{\mathbf{p}}^+ + n_{\mathbf{p}}^-) \frac{p^2}{2m} - \mu_0 H (N_+ - N_-) + \frac{4\pi a \hbar^2 N_+ N_-}{mV} \right] \right\} \quad (11.127)$$

The notation is the same as that of (11.106). Proceeding in the same way as in the evaluation of (11.106), we obtain

$$\frac{1}{N} \log Q_N = \beta g(\bar{N}_+) \quad (11.128)$$

where

$$\begin{aligned} g(\bar{N}_+) &= \max [g(N_+)] \\ g(N_+) &\equiv \mu_0 H \left(\frac{2N_+}{N} - 1 \right) - \frac{4\pi a \hbar^2}{mV} N_+ (N - N_+) \\ &\quad - \frac{1}{N} [A(N_+) + A(N - N_+)] \end{aligned} \quad (11.129)$$

Thus \bar{N}_+ is the root of the equations

$$\begin{aligned} \left[\frac{\partial g(N_+)}{\partial N_+} \right]_{N_+ = \bar{N}_+} &= 0 \\ \left[\frac{\partial^2 g(N_+)}{\partial N_+^2} \right]_{N_+ = \bar{N}_+} &< 0 \end{aligned} \quad (11.130)$$

It must be noted that (11.130) locates the point at which the curve $g(N_+)$ passes through a maximum. It is conceivable (and in fact true) that \bar{N}_+ may occur not at a maximum of the curve $g(N_+)$ but at the boundary of the range of N_+ , i.e., at $\bar{N}_+ = 0$ or $\bar{N}_+ = N$. We keep this in mind as we proceed. With $\nu(N)$ defined as in (11.114), we rewrite (11.130) as

$$kT [\nu(\bar{N}_+) - \nu(N - \bar{N}_+)] = 2\mu H + \frac{4\pi a \hbar^2}{mV} (2\bar{N}_+ - N) \quad (11.131)$$

$$kT [\nu'(\bar{N}_+) + \nu'(N - \bar{N}_+)] - \frac{8\pi a \hbar^2}{mV} > 0$$

where $\nu'(N) \equiv \partial \nu(N) / \partial N$. Let

$$r \equiv \frac{2\bar{N}_+}{N} - 1 \quad (-1 \leq r \leq +1) \quad (11.132)$$

Then (11.131) becomes

$$kT \left\{ \nu \left[\frac{N}{2}(1+r) \right] - \nu \left[\frac{N}{2}(1-r) \right] \right\} = 2\mu_0 H + \frac{a\lambda^2}{\nu} 2kTr \quad (11.133)$$

$$\frac{\partial}{\partial r} \left\{ N \left[\frac{N}{2}(1+r) \right] - \nu \left[\frac{N}{2}(1-r) \right] \right\} - \frac{2a\lambda^2}{\nu} > 0$$

where $\nu[x] \equiv \nu(x)$ and $\lambda = \sqrt{2\pi \hbar^2 / mkT}$, the thermal wavelength. The low-temperature and high-temperature approximations for $\nu(Nx/2)$ are obtainable from (11.24) and (11.12), respectively. They are

$$kT \nu \left(\frac{N}{2} x \right) \approx x^{2/3} \epsilon_F \left[1 - \frac{\pi^2}{12} \left(\frac{kT}{\epsilon_F} \right)^2 \frac{1}{x^{4/3}} \right] \quad \left(\frac{kT}{\epsilon_F} \ll 1 \right) \quad (11.134)$$

$$kT \nu \left(\frac{N}{2} x \right) \approx \log \frac{\lambda^2}{2\nu} \quad \left(\frac{kT}{\epsilon_F} \gg 1 \right)$$

Spontaneous Magnetization

We first consider the case $H = 0$. At absolute zero, (11.134) reduces to

$$(1+r)^{2/3} - (1-r)^{2/3} = \zeta r \quad (11.135)$$

$$\frac{1}{2} \left[\frac{1}{(1+r)^{1/3}} + \frac{1}{(1-r)^{1/3}} \right] > \frac{3}{4} \zeta$$

where

$$\zeta \equiv \frac{8}{3\pi} k_F a \quad (11.136)$$

Equation (11.135) is invariant under a change of sign of r . This is to be expected; in the absence of field, no absolute meaning can be attached to "up" or "down."

Thus it is sufficient consider $r \geq 0$. We may solve (11.135) graphically by referring to Fig. 11.16, where $(1+r)^{2/3} - (1-r)^{2/3}$ is plotted against r . We need only obtain the intersection between the curve in Fig. 11.16 and the straight line ζr . It is seen that for $\zeta < \frac{4}{3}$, $r = 0$ is the only intersection. If ζ is such that

$$\frac{4}{3} < \zeta < 2^{2/3} \quad (11.137)$$

then there is an additional intersection $r > 0$, and the value $r > 0$ corresponds to a maximum, whereas the value $r = 0$ corresponds to a minimum. If $\zeta > 2^{2/3}$, then (11.135) has no solution. In this case the maximum of $g(N_+)$ must occur either at $N_+ = 0$ or at $N_+ = N$, unless $g(N_+)$ is a constant. Since $g(N_+)$ is not a constant, and since there is no distinction between $N_+ = 0$ and $N_+ = N$, we can choose to let $\bar{N}_+ = N$, or $r = 1$. The value of r at absolute zero as a function of the repulsive strength ζ , is summarized as follows:

$$\begin{aligned} r = 0 & \quad (\zeta < \frac{4}{3}) & \quad \text{(no spontaneous magnetization)} \\ 0 < r < 1 & \quad (\frac{4}{3} < \zeta < 2^{2/3}) & \quad \text{(partial spontaneous magnetization)} \\ r = 1 & \quad (\zeta > 2^{2/3}) & \quad \text{(saturated spontaneous magnetization)} \end{aligned} \quad (11.138)$$

That is, if the repulsive strength is sufficiently strong, the system becomes ferromagnetic. The critical value of a at which ferromagnetism first sets in ($\zeta = \frac{4}{3}$) corresponds to

$$k_F a = \frac{\pi}{2} \quad (11.139)$$

The foregoing results hold at absolute zero. At a finite but small temperature we have, instead of (11.135),

$$(1+r)^{2/3} - (1-r)^{2/3} - \frac{\pi^2}{12} \left(\frac{kT}{\epsilon_F} \right)^2 \left[\frac{1}{(1+r)^{2/3}} - \frac{1}{(1-r)^{2/3}} \right] = \zeta r$$

$$\frac{1}{2} \left\{ \frac{1}{(1+r)^{1/3}} + \frac{1}{(1-r)^{1/3}} + \frac{\pi^2}{12} \left(\frac{kT}{\epsilon_F} \right)^2 \left[\frac{1}{(1+r)^{5/3}} + \frac{1}{(1-r)^{5/3}} \right] \right\} > \frac{3}{4} \zeta \quad (11.140)$$

Let $r(T)$ be the solution at absolute temperature T . It is easily seen that if $r(0) = 0$, then $r(T) = 0$; if $r(0) > 0$, then $r(T) < r(0)$. Thus, if there is spontaneous magnetization at absolute zero, the magnetization decreases with temperature. The spontaneous magnetization vanishes above a critical temperature T_c (the Curie temperature), which is the value of T at which both equations in (13.45) are satisfied for $\zeta > \frac{4}{3}$ and $r = 0$. We find that

$$\frac{kT_c}{\epsilon_F} = \frac{3}{\pi} \sqrt{\zeta - \frac{4}{3}} = \frac{2}{\pi\sqrt{3}} \sqrt{\frac{2}{\pi} k_F a - 1} \quad (11.141)$$

A qualitative plot of the magnetization $\mu r / \nu$ is shown in Fig. 11.18.

It must be pointed out that the model we have used is a physical model only if $k_F a \ll 1$. Therefore the case of ferromagnetism, which requires $k_F a > \pi/2$, is beyond the domain of validity of the model. It is instructive, however, to see how

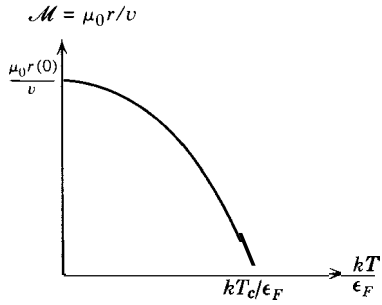


Fig. 11.18 Spontaneous magnetization of an imperfect Fermi gas with repulsive interactions.

the spatial repulsion between the fermions can enhance the spin alignment to such an extent that, if we are willing to extrapolate the results of a weak interaction model, ferromagnetism results.

Paramagnetic Susceptibility

We now consider the case of $H > 0$. Let $r_0(T)$ be the value of r for $H = 0$, but for an arbitrary temperature. Putting

$$r = r_0(T) + \frac{\chi v}{p_0} H \quad (11.142)$$

and treating $\chi v H / \mu_0$ as a small quantity, we can solve (11.134) and obtain

$$\chi = \frac{2\mu_0^2/\epsilon_F v}{\frac{NkT}{2\epsilon_F} \left\{ \nu' \left[\frac{N}{2}(1+r_0) \right] + \nu' \left[\frac{N}{2}(1-r_0) \right] \right\} - \frac{8}{3a} k_F a} \quad (11.143)$$

The low- and high-temperature limits are

$$\chi \xrightarrow{T \rightarrow 0} \frac{3\mu_0^2/\epsilon_F v}{(1+r_0)^{-1/3} + (1-r_0)^{-1/3} - (4/\pi)k_F a} \quad (11.144)$$

$$\chi \xrightarrow{T \rightarrow \infty} \frac{\mu_0^2}{kTv} \quad (11.145)$$

Hence Curie's constant is

$$C = \frac{\mu_0^2}{kv} \quad (11.146)$$

Note that r_0 depends on $k_F a$. It approaches unity when $k_F a$ exceeds a certain value. Thus it can be seen from (11.143) that in general $\chi > 0$. The system is either ferromagnetic or paramagnetic, never diamagnetic.

Consider now the case of paramagnetism, for which we require $r_0 = 0$ for all temperatures. This means that

$$k_F a < \frac{\pi}{2} \quad (11.147)$$

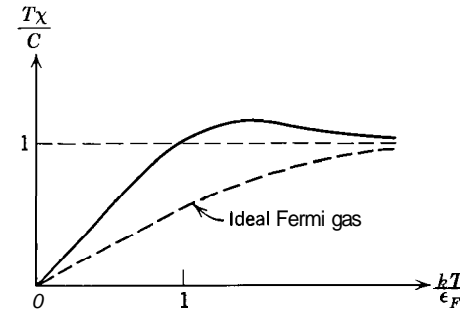


Fig. 11.19 Paramagnetic susceptibility of an imperfect Fermi gas with repulsive interactions. The model used is well founded only for $kT/\epsilon_F \ll 1$.

Here (11.143) becomes

$$\frac{T\chi}{C} = \frac{3kT}{2\epsilon_F} \frac{1}{f - (2/\pi)k_F a} \quad (11.148)$$

where

$$f \equiv \frac{3kT}{2\epsilon_F} \frac{N}{2} \nu' \left(\frac{N}{2} \right) \quad (11.149)$$

The function $(T\chi/C)$ rises linearly at $T = 0$, with a slope given by

$$\frac{\partial}{\partial T} \left(\frac{T\chi}{C} \right)_{T=0} = \frac{3}{2} \frac{1}{1 - (2/\pi)k_F a} \quad (11.150)$$

It reaches a maximum value, which is greater than unity, at $kT/\epsilon_F \approx 1$. Then it approaches unity as $T \rightarrow \infty$. A qualitative plot of $T\chi/C$ is shown in Fig. 11.19. If we calculate χ for an ideal Fermi gas endowed with the same magnetic moment, we find the slope

$$\frac{\partial}{\partial T} \left(\frac{T\chi}{C} \right)_{T=0} = \frac{3}{2} \quad (\text{ideal Fermi gas}) \quad (11.151)$$

The imperfect gas has a steeper slope, as (11.150) shows, which is again a reflection of the enhancement of spin alignment by the repulsive interaction. The result is sometimes described by saying that imperfect gas behaves like an ideal gas with a higher Fermi energy.*

PROBLEMS

11.1 Give numerical estimates for the Fermi energy of

(a) electrons in a typical metal;

(b) nucleons in a heavy nucleus;

(c) He^3 atoms in liquid He^3 (atomic volume = $46.2 \text{ \AA}^3/\text{atom}$). Treat all the mentioned particles as free particles.

*See, however, Problem 11.7.

11.2 Show that for the ideal Fermi gas the Helmholtz free energy per particle at low temperatures is given by

$$\frac{A}{N} = \frac{3}{5} \epsilon_F \left[1 - \frac{5\pi^2}{12} \left(\frac{kT}{\epsilon_F} \right)^2 + \dots \right]$$

11.3 A collection of free nucleons is enclosed in a box of volume V . The energy of a single nucleon of momentum p is

$$\epsilon_p = \frac{p^2}{2m} + mc^2$$

where $mc^2 = 1000$ MeV.

- (a) Pretending that there is no conservation law for the number of nucleons, calculate the partition function of a system of nucleons (which obey Fermi statistics) at temperature T .
- (b) Calculate the average energy density.
- (c) Calculate the average particle density.
- (d) Discuss the necessity for a conservation law for the number of nucleons, in the light of the foregoing calculations.

11.4 (a) What is the heat capacity C_V of a three-dimensional cubic lattice of atoms at room temperature? Assume each atom to be bound to its equilibrium position by Hooke's law forces.

(b) Assuming that a metal can be represented by such a lattice of atoms plus freely moving electrons, compare the specific heat due to the electrons with that due to the lattice, at room temperature.

11.15 A cylinder is separated into two compartments by a free sliding piston. Two ideal Fermi gases are placed into the two compartments, numbered 1 and 2. The particles in compartment 1 have spin $\frac{1}{2}$, while those in compartment 2 have spin $\frac{3}{2}$. They all have the same mass. Find the equilibrium relative density of the two gases at $T = 0$ and at $T \rightarrow \infty$.

11.6 Consider a two-dimensional electron gas in a magnetic field strong enough so that all particles can be accommodated in the lowest Landau level. Taking into account both orbital and spin paramagnetism, find the magnetization at absolute zero.

11.7 (a) Show that for the imperfect Fermi gas discussed in Section 11.6 the specific heat at constant volume is given by

$$\frac{C_V}{N} = -2k \frac{\partial}{\partial T} \left[I(r) T^2 \frac{\partial r}{\partial T} \right] + \frac{32\pi a \hbar^2}{mv} T \left[\left(\frac{\partial r}{\partial T} \right)^2 + r \frac{\partial^2 r}{\partial T^2} \right]$$

where

$$I(r) \equiv \nu \left[\frac{N}{2}(1+r) \right] - \nu \left[\frac{N}{2}(1-r) \right]$$

(b) Show that when there is no spontaneous magnetization

$$C_V = (C_V)_{\text{ideal gas}}$$

and hence the interpretation that the imperfect gas behaves like an ideal gas with a higher Fermi energy cannot be consistently maintained.

CHAPTER

12

BOSE SYSTEMS

The dominant characteristic of a system of bosons is a "statistical" attraction between the particles. In contradistinction to the case of **fermions**, the particles like to have the same quantum numbers. When the particle number is conserved, this attraction leads to the Bose-Einstein condensation, which is the basis of **superfluidity**. In this chapter we illustrate various Bose systems, discuss the Bose-Einstein condensation, and introduce the notion of the superfluid order parameter.

12.1 PHOTONS

Consider the equilibrium properties of electromagnetic radiation enclosed in a volume V at temperature T , a system known as a "blackbody cavity." It can be experimentally produced by making a cavity in any material, evacuating the cavity completely, and then heating the material to a given temperature. The atoms in the walls of this cavity will constantly emit and absorb electromagnetic radiation, so that in equilibrium there will be a certain amount of electromagnetic radiation in the cavity, and nothing else. If the cavity is sufficiently large, the thermodynamic properties of the radiation in the cavity should be independent of the nature of the wall. Accordingly we can impose on the radiation field any boundary condition that is convenient.

The Hamiltonian for a free electromagnetic field can be written as a sum of terms, each having the form of a **Hamiltonian** for a harmonic oscillator of some frequency. This corresponds to the possibility of regarding any radiation field as a linear superposition of plane waves of various frequencies. In quantum theory each harmonic oscillator of frequency ω can only have the energies $(n + \frac{1}{2})\hbar\omega$, where $n = 0, 1, 2, \dots$. This fact leads to the concept of photons as quanta of the electromagnetic field. A state of the free electromagnetic field is specified by the number n for each of the oscillators. In other words, it is specified by enumerating the number of photons present for each frequency.

According to the quantum theory of radiation, photons are massless bosons of spin A . The masslessness implies that a photon always moves with the velocity of light c in free space, and that its spin can have only two independent orientations: parallel and antiparallel to the momentum. A photon in a definite spin state corresponds to a plane electromagnetic wave that is either right- or left-circularly polarized. We may, however, superimpose two photon states with definite spins and obtain a photon state that is linearly polarized but that is not an eigenstate of spin. In the following we consider linearly polarized photons.

For our purpose it is sufficient to know that a photon of frequency ω has the following properties:

$$\begin{aligned} \text{Energy} &= \hbar\omega \\ \text{Momentum} &= \hbar\mathbf{k}, \quad |\mathbf{k}| = \frac{\omega}{c} \\ \text{Polarization vector} &= \mathbf{r}, \quad |\boldsymbol{\epsilon}| = 1, \quad \mathbf{k} \cdot \boldsymbol{\epsilon} = 0 \end{aligned} \quad (12.1)$$

Such a photon corresponds* to a plane wave of electromagnetic radiation whose electric field vector is

$$\mathbf{E}(\mathbf{r}, t) = \boldsymbol{\epsilon} e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} \quad (12.2)$$

The direction of $\boldsymbol{\epsilon}$ is the direction of the electric field. The condition $\boldsymbol{\epsilon} \cdot \mathbf{k} = 0$ is a consequence of the transversality of the electric field, i.e., $\mathbf{V} \cdot \mathbf{E} = 0$. Thus for given \mathbf{k} there are two and only two independent polarization vectors \mathbf{r} . If we impose periodic boundary conditions on $\mathbf{E}(\mathbf{r}, t)$ in a cube of volume $V = L^3$, we obtain the following allowed values of \mathbf{k} :

$$\mathbf{k} = \frac{2\pi\mathbf{n}}{L} \quad (12.3)$$

$\mathbf{n} = a$ vector whose components are $0, \pm 1, \pm 2, \dots$

Thus the number of allowed momentum values between \mathbf{k} and $\mathbf{k} + d\mathbf{k}$ is

$$\frac{V}{(2\pi)^3} 4\pi k^2 dk \quad (12.4)$$

Since atoms can emit and absorb photons, the total number of photons is not a conserved quantity.

The total energy of the state of the electromagnetic field in which there are $n_{\mathbf{k}, \boldsymbol{\epsilon}}$ photons of momentum \mathbf{k} and polarization $\boldsymbol{\epsilon}$ is given by

$$E\{n_{\mathbf{k}, \boldsymbol{\epsilon}}\} = \sum_{\mathbf{k}, \boldsymbol{\epsilon}} \hbar\omega n_{\mathbf{k}, \boldsymbol{\epsilon}} \quad (12.5)$$

where

$$\begin{aligned} \omega &= c|\mathbf{k}| \\ n_{\mathbf{k}, \boldsymbol{\epsilon}} &= 0, 1, 2, \dots \end{aligned} \quad (12.6)$$

*For a precise meaning of this statement, we refer the reader to any book on the quantum theory of radiation.

Since the number of photons is indefinite, the partition function is

$$Q = \sum_{\{n_{\mathbf{k}, \boldsymbol{\epsilon}}\}} e^{-\beta E\{n_{\mathbf{k}, \boldsymbol{\epsilon}}\}} \quad (12.7)$$

with no restriction on $\{n_{\mathbf{k}, \boldsymbol{\epsilon}}\}$.* The calculation of Q is trivial:

$$\begin{aligned} Q &= \sum_{\{n_{\mathbf{k}, \boldsymbol{\epsilon}}\}} \exp\left(-\beta \sum_{\mathbf{k}, \boldsymbol{\epsilon}} \hbar\omega n_{\mathbf{k}, \boldsymbol{\epsilon}}\right) = \prod_{\mathbf{k}, \boldsymbol{\epsilon}} \sum_{n=0}^{\infty} e^{-\beta\hbar\omega n} = \prod_{\mathbf{k}, \boldsymbol{\epsilon}} \frac{1}{1 - e^{-\beta\hbar\omega}} \\ \log Q &= - \sum_{\mathbf{k}, \boldsymbol{\epsilon}} \log(1 - e^{-\beta\hbar\omega}) = -2 \sum_{\mathbf{k}} \log(1 - e^{-\beta\hbar\omega}) \end{aligned} \quad (12.8)$$

The average occupation number for photons of momentum \mathbf{k} , regardless of polarization, is

$$\langle n_{\mathbf{k}} \rangle = - \frac{1}{\beta} \frac{\partial}{\partial (\hbar\omega)} \log Q = \frac{2}{e^{\beta\hbar\omega} - 1} \quad (12.9)$$

where the factor 2 comes from the two possible polarizations.

The internal energy is

$$U = - \frac{\partial}{\partial \beta} \log Q = \sum_{\mathbf{k}} \hbar\omega \langle n_{\mathbf{k}} \rangle \quad (12.10)$$

To find the pressure, we express Q in the form

$$\log Q = -2 \sum_{\mathbf{n}} \log(1 - e^{-\beta\hbar c 2\pi|\mathbf{n}|V^{-1/3}}) \quad (12.11)$$

from which we obtain

$$P = \frac{1}{\beta} \frac{\partial}{\partial V} \log Q = \frac{1}{3V} \sum_{\mathbf{k}} \hbar\omega \langle n_{\mathbf{k}} \rangle$$

Comparison between this equation and (12.10) leads to the equation of state

$$PV = \frac{1}{3} U \quad (12.12)$$

We now calculate U in the limit as $V \rightarrow \infty$. From (12.10), (12.9), and (12.3) we have

$$U = \frac{2V}{(2\pi)^3} \int_0^{\infty} dk 4\pi k^2 \frac{\hbar ck}{e^{\beta\hbar ck} - 1} = \frac{V\hbar}{\pi^2 c^3} \int_0^{\infty} d\omega \frac{\omega^3}{e^{\beta\hbar\omega} - 1}$$

Hence the internal energy per unit volume is

$$\frac{U}{V} = \int_0^{\infty} d\omega u(\omega, T) \quad (12.13)$$

where

$$u(\omega, T) = \frac{A}{\pi^2 c^3} \frac{\omega^3}{e^{\beta\hbar\omega} - 1} \quad (12.14)$$

*One could say that the chemical potential is 0, because a photon can disappear into the vacuum.

This is Planck's radiation law, which gives the energy density due to photons of frequency ω , regardless of polarization and direction of momentum. The integral (12.13) can be explicitly evaluated to give

$$\frac{U}{V} = \frac{\pi^2 (kT)^4}{15 (\hbar c)^3} \quad (12.15)$$

It follows that the specific heat per unit volume is

$$c_V = \frac{4\pi^2 k^4 T^3}{15 (\hbar c)^3} \quad (12.16)$$

The specific heat is not bounded as $T \rightarrow \infty$, because the number of photons in the cavity is not bounded.

Both (12.14) and (12.15) can be verified experimentally by opening the black-body cavity to the external world through a small window. Radiation would then escape from the cavity with the velocity c . The amount of energy radiated per second per unit area of the opening, in the form of photons of frequency ω , is

$$I(\omega, T) = c \int \frac{d\Omega}{4\pi} u(\omega, T) \cos \theta = \frac{c}{4} u(\omega, T) \quad (12.17)$$

where the angular integration extends only over a hemisphere. Integrating over the frequency, we obtain

$$I(T) = \int_0^\infty d\omega I(\omega, T) = \sigma T^4 \quad (12.18)$$

$$\sigma = \frac{\pi^2 k^4}{60 \hbar^3 c^3}$$

This is known as Stefan's law, and σ is Stefan's constant. The function $I(\omega, T)$ is shown in Fig. 12.1, showing that the radiation peaks at a frequency that is an increasing function of T . The area under the curves shown in Fig. 12.1 increases like T^4 . All these conclusions are in excellent agreement with experiments.

It should be noted that although the form of $u(\omega, T)$ can be arrived at only through quantum theory the equation of state $PV = U/3$ and the fact that $U \propto T^4$ can be derived in classical physics.

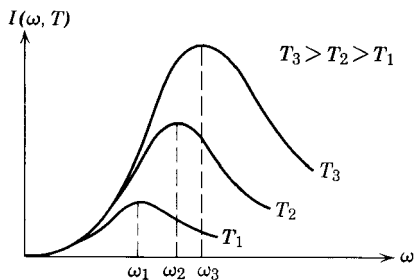


Fig. 12.1 Planck's radiation law.

The equation of state may be derived as follows. Consider first a plane wave whose electric and magnetic field vectors are \mathbf{E} and \mathbf{B} . The average energy density is

$$\frac{1}{2} \overline{(\mathbf{E}^2 + \mathbf{B}^2)} = \overline{E^2}$$

The radiation pressure, which is equal to the average momentum flux, is

$$\overline{|\mathbf{E} \times \mathbf{B}|} = \overline{E^2}$$

Thus the energy density is numerically equal to the radiation pressure. Now consider an amount of isotropic radiation contained in a cubical box. The radiation field in the box may be considered an *incoherent* superposition of plane waves propagating in all directions. The relative intensities of the plane waves depend only on the temperature as determined by the walls of the box. The radiation pressure on any wall of the box is one-third of the energy density in the box, because, whereas all the plane waves contribute to the energy density, only one-third of the plane waves contribute to the radiation pressure on any wall of the box.

To derive $U \propto T^4$, recall that the second law of thermodynamics implies the following relation, which holds for all systems:

$$\left(\frac{\partial U}{\partial V} \right)_T = T \left(\frac{\partial P}{\partial T} \right)_V - P \quad (12.19)$$

From $PV = U/3$ and the fact that P depends on temperature alone we have

$$\left(\frac{\partial U}{\partial V} \right)_T = 3P = \frac{U}{V} \equiv u(T) \quad (12.20)$$

Using (12.19) we have

$$u = \frac{T}{3} \frac{du}{dT} - \frac{1}{3}u$$

$$\frac{du}{u} = 4 \frac{dT}{T}$$

Hence

$$u = CT^4 \quad (12.21)$$

The constant C cannot be obtained through classical considerations.

If the photon had a finite rest mass, no matter how small, then it would have three independent polarizations instead of two.* There would be, in addition to transverse photons, longitudinal photons. If this were so, Planck's radiation formula (12.14) would be altered by a factor of $\frac{3}{2}$. The fact that (12.14) has been experimentally verified means that either the photon has no rest mass, or if it does the coupling between longitudinal photons and matter is so small that

*If the photon had a finite rest mass, it could be transformed to rest by a Lorentz transformation. We could then make a second Lorentz transformation in an arbitrary direction, so that the spin would lie neither parallel nor antiparallel to the momentum.

thermal equilibrium between longitudinal photons and matter cannot be established during the course of any of our experiments concerned with Planck's radiation law.

12.2 PHONONS IN SOLIDS

Phonons are quanta of sound waves in a macroscopic body. Mathematically they emerge in a similar way that photons arise from the quantization of the electromagnetic field. For low-lying excitations, the Hamiltonian for a solid, which is made up of atoms arranged in a crystal lattice, may be approximated by a sum of terms, each representing a harmonic oscillator, corresponding to a normal mode of lattice oscillation.* Each normal mode is classically a wave of distortion of the lattice planes—a sound wave. In quantum theory these normal modes give rise to quanta called phonons. A quantum state of a crystal lattice near its ground state may be specified by enumerating all the phonons present. Therefore at a very low temperature a solid can be regarded as a volume containing a gas of noninteracting phonons.

Since a phonon is a quantum of a certain harmonic oscillator, it has a characteristic frequency ω_i and an energy $A\omega_i$. The state of the lattice in which one phonon is present corresponds to a sound wave of the form

$$\epsilon e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \quad (12.22)$$

where the propagation vector \mathbf{k} has the magnitude

$$|\mathbf{k}| = \frac{\omega}{c} \quad (12.23)$$

in which c is the velocity of sound.† The polarization vector ϵ can have three independent directions, corresponding to one longitudinal mode of compression wave and two transverse modes of shear wave. Since an excited state of a harmonic oscillator may contain any number of quanta, the phonons obey Bose statistics, with no conservation of their total number.

If a solid has N atoms, it has $3N$ normal modes. Therefore there will be $3N$ different types of phonon with the characteristic frequencies

$$\omega_1, \omega_2, \dots, \omega_{3N} \quad (12.24)$$

The values of these frequencies depend on the nature of the lattice. In the Einstein model of a lattice they are taken to be equal to one another. An improved model is that of Debye, who assumed that for the purpose of finding the frequencies (12.24), one may consider the solid as an elastic continuum of volume V . The frequencies (12.24) are then taken to be the lowest $3N$ normal frequencies of such a system. Since an elastic continuum has a continuous

*In as much as anharmonic forces between atoms, which at high temperatures allow the lattice to melt, can be neglected.

†We assume an isotropic solid, for which c is independent of the polarization vector ϵ .

distribution of normal frequencies we shall be interested in the number of normal modes whose frequency lies between ω and $\omega + d\omega$. To find this number we must know the boundary conditions on a sound wave in the elastic medium. Taking periodic boundary conditions, we find as usual that $\mathbf{k} = (2\pi/L)\mathbf{n}$, where $L = V^{1/3}$ and \mathbf{n} has the components $0, \pm 1, \pm 2, \dots$. The number we seek is then

$$f(\omega) d\omega = \frac{\text{no. of normal modes with frequency between } \omega \text{ and } \omega + d\omega}{(2\pi)^3} 4\pi k^2 dk \quad (12.25)$$

where the factor 3 comes from the three possible polarizations. Since $k = \omega/c$ we have

$$f(\omega) d\omega = V \frac{3\omega^2}{2\pi^2 c^3} d\omega \quad (12.26)$$

The maximum frequency ω_m is obtained by the requirement that

$$\int_0^{\omega_m} f(\omega) d\omega = 3N \quad (12.27)$$

which gives, with $v = V/N$,

$$\omega_m = c \left(\frac{6\pi^2}{v} \right)^{1/3} \quad (12.28)$$

The wavelength corresponding to ω_m is

$$\lambda_m = \frac{2\pi c}{\omega_m} = \left(\frac{4}{3}\pi v \right)^{1/3} \approx \text{interparticle distance} \quad (12.29)$$

This is a reasonable criterion because for wavelengths shorter than λ_m a wave of displacements of atoms becomes meaningless.

We now calculate the equilibrium properties of a solid at low temperatures by calculating the partition function for an appropriate gas of phonons. The energy of the state in which there are n_i phonons of the i th type is*

$$E\{n_i\} = \sum_{i=1}^{3N} n_i \hbar \omega_i \quad (12.30)$$

The partition function is

$$Q = \sum_{\{n_i\}} e^{-\beta E\{n_i\}} = \prod_{i=1}^{3N} \frac{1}{1 - e^{-\beta \hbar \omega_i}}$$

Hence

$$\log Q = - \sum_{i=1}^{3N} \log(1 - e^{-\beta \hbar \omega_i}) \quad (12.31)$$

*We should add to (12.30) an unknown constant representing the ground state energy of the solid, but this constant does not affect any subsequent results and hence can be ignored.

The average occupation number is

$$\langle n_i \rangle = -\frac{1}{\beta} \frac{d}{d(\hbar\omega_i)} \log Q = \frac{1}{e^{\beta\hbar\omega_i} - 1} \quad (12.32)$$

The integral energy is

$$U = -\frac{\partial}{\partial\beta} \log Q = \sum_{i=1}^{3N} \hbar\omega_i \langle n_i \rangle = \sum_{i=1}^{3N} \frac{\hbar\omega_i}{e^{\beta\hbar\omega_i} - 1} \quad (12.33)$$

Passing to the limit $V \rightarrow \infty$ we obtain, with the help of (12.26),

$$U = \frac{3V}{2\pi^2c^3} \int_0^{\omega_m} d\omega \omega^2 \frac{\hbar\omega}{e^{\beta\hbar\omega} - 1} \quad (12.34)$$

$$\frac{U}{N} = \frac{9(kT)^4}{(\hbar\omega_m)^3} \int_0^{\beta\hbar\omega_m} dt \frac{t^3}{e^t - 1} \quad (12.35)$$

We define the Debye function $D(x)$ by

$$D(x) = \frac{3}{x} \int_0^x dt \frac{t^3}{e^t - 1} = \begin{cases} 1 - \frac{3}{8}x + \frac{1}{20}x^2 + \dots & (x \ll 1) \\ \frac{\pi^4}{5x^3} + O(e^{-x}) & (x \gg 1) \end{cases} \quad (12.36)$$

and the Debye temperature T_D by

$$kT_D \equiv \hbar\omega_m = \hbar c \left(\frac{6\pi^2}{v} \right)^{1/3} \quad (12.37)$$

Then

$$\frac{U}{N} = 3kTD(\lambda) = \begin{cases} 3kT \left(1 - \frac{3}{8} \frac{T_D}{T} + \dots \right) & (T \gg T_D) \\ 3kT \left[\frac{\pi^4}{5} \left(\frac{T}{T_D} \right)^3 + O(e^{-T_D/T}) \right] & (T \ll T_D) \end{cases} \quad (12.38)$$

where $\lambda \equiv T_D/T$. Then the specific heat is given by

$$\frac{C_V}{Nk} = 3D(\lambda) + 3T \frac{dD(\lambda)}{dT} = 3 \left[4D(\lambda) - \frac{3\lambda}{e^\lambda - 1} \right] \quad (12.39)$$

The high- and low-temperature behaviors of C_V are as follows:

$$\frac{C_V}{Nk} = \begin{cases} 3 \left[1 - \frac{1}{20} \left(\frac{T_D}{T} \right)^2 + \dots \right] & (T \gg T_D) \\ \frac{12\pi^4}{5} \left(\frac{T}{T_D} \right)^3 + O(e^{-T_D/T}) & (T \ll T_D) \end{cases} \quad (12.40)$$

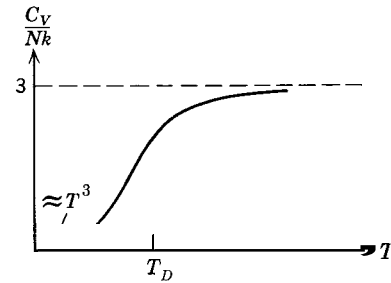


Fig. 12.2 Specific heat of a crystal lattice in Debye's theory.

A plot of the specific heat is shown in Fig. 12.2, which agrees quite well with experimental findings.

At low temperatures C_V vanishes like T^3 , verifying the third law of thermodynamics. When the temperature is much greater than the Debye temperature the lattice behaves classically, as indicated by the fact that $C_V = 3Nk$. For most solids the Debye temperature is of the order of 200 K. This is why the Dulong-Petit law $C_V \approx 3Nk$ holds at room temperatures. At extremely high temperatures the model of noninteracting phonons breaks down because the lattice eventually melts. The melting of the lattice is made possible by the fact that the forces between the atoms in the lattice are not strictly harmonic forces. In the phonon language the phonons are not strictly free. They must interact with each other, and this interaction becomes strong at very high temperatures.

12.3 BOSE-EINSTEIN CONDENSATION

Equation (8.71) gives the equation of state for the ideal Bose gas of N particles of mass m contained in a volume V . To study in detail the properties of the equation of state we must find the fugacity z as a function of temperature and specific volume by solving the second equation of (8.71), namely

$$\frac{1}{v} = \frac{1}{\lambda^3} g_{3/2}(z) + \frac{1}{V} \frac{z}{1-z} \quad (12.41)$$

where $v = V/N$, and $\lambda = \sqrt{2\pi\hbar^2/mkT}$, the thermal wavelength. To do this, we must first study the properties of the function $g_{3/2}(z)$, which is a special case of a more general class of functions

$$g_n(z) \equiv \sum_{l=1}^{\infty} \frac{z^l}{l^n} \quad (12.42)$$

These functions have been studied* and tabulated† in the literature.

It is obvious that for real values of z between 0 and 1, $g_{3/2}(z)$ is a bounded, positive, monotonically increasing function of z . To satisfy (12.41) it is necessary

*J. E. Robinson, *Phys. Rev.* **83**, 678 (1951).

†F. London, *Superfluids*, Vol. II (Wiley, New York, 1954), Appendix.

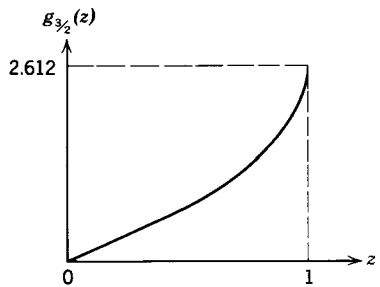


Fig. 12.3 The function $g_{3/2}(z)$.

that

$$0 \leq z \leq 1$$

For comparison we recall that $0 \leq z < \infty$ in the case of Fermi statistics. For small z , the power series (12.42) furnishes a practical way to calculate $g_{3/2}(z)$:

$$g_{3/2}(z) = z + \frac{z^2}{2^{3/2}} + \frac{z^3}{3^{3/2}} + \dots \quad (12.43)$$

At $z = 1$ its derivative diverges, but its value is finite:

$$g_{3/2}(1) = \sum_{l=1}^{\infty} \frac{1}{l^{3/2}} = \zeta\left(\frac{3}{2}\right) = 2.612\dots \quad (12.44)$$

where $\zeta(x)$ is the Riemann zeta function of x . Thus for all z between 0 and 1,

$$g_{3/2}(z) \leq 2.612\dots \quad (12.45)$$

A graph of $g_{3/2}(z)$ is shown in Fig. 12.3.

Let us rewrite (12.41) in the form

$$\lambda^3 \frac{\langle n_0 \rangle}{V} = \frac{\lambda^3}{v} - g_{3/2}(z) \quad (12.46)$$

This implies that $\langle n_0 \rangle/V > 0$ when the temperature and the specific volume are such that

$$\frac{\lambda^3}{v} > g_{3/2}(1) \quad (12.47)$$

This means that a finite fraction of the particles occupies the level with $p = 0$. This phenomenon is known as the **Bose-Einstein condensation**. The condition (12.47) defines a subspace of the thermodynamic P-u-T space of the ideal Bose gas, which corresponds to the transition region of the Bose-Einstein condensation. As we see later, in this region the system can be considered to be a mixture of two thermodynamic phases, one phase being composed of particles with $p = 0$ and the other with $p \neq 0$. We refer to the region (12.47) as the condensation region. It is separated from the rest of the P-u-T space by the two-dimensional surface

$$\frac{\lambda^3}{v} = g_{3/2}(1) \quad (12.48)$$

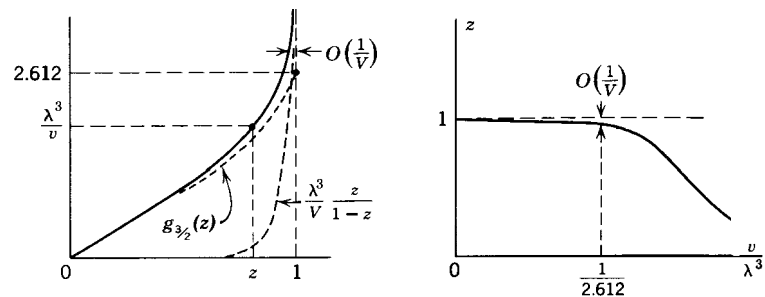


Fig. 12.4 (a) Graphical solution of (12.41); (b) the fugacity for an ideal Bose gas contained in a finite volume V .

For a given specific volume v , (12.48) defines a critical temperature T_c :

$$\lambda_c^3 = v g_{3/2}(1) \quad (12.49)$$

or

$$kT_c = \frac{2\pi\hbar^2/m}{[v g_{3/2}(1)]^{2/3}} \quad (12.50)$$

As indicated by (12.49), T_c is the temperature at which the thermal wavelength is of the same order of magnitude as the average interparticle separation. For a given temperature T , (12.48) defines a critical volume v_c :

$$v_c = \frac{\lambda^3}{g_{3/2}(1)} \quad (12.51)$$

In terms of T_c and v_c the region of condensation is the region in which $T < T_c$ or $v < v_c$.

To find z as a function of T and v we solve (12.41) graphically. For a large but finite value of the total volume V the graphical construction in Fig. 12.4a yields the curve for z shown in Fig. 12.4b. In the limit as $V \rightarrow \infty$ we obtain

$$z = \begin{cases} 1 & \left(\frac{\lambda^3}{v} \geq g_{3/2}(1) \right) \\ \text{the root of } g_{3/2}(z) = \lambda^3/v & \left(\frac{\lambda^3}{v} \leq g_{3/2}(1) \right) \end{cases} \quad (12.52)$$

For $(\lambda^3/v) \leq g_{3/2}(1)$, the value of z must be found by numerical methods. A graph of z is given in Fig. 12.5.

To make these considerations more rigorous the following point must be noted. It is recalled that (12.41) is derived from the condition

$$\frac{N}{V} = \frac{1}{V} \sum_{\mathbf{p} \neq 0} \langle n_{\mathbf{p}} \rangle + \frac{\langle n_0 \rangle}{V}$$

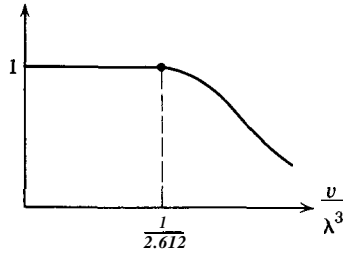


Fig. 12.5 The fugacity for an ideal Bose gas of infinite volume.

by replacing the sum on the right side by an integral. It is clear that this integral is unchanged if we subtract from the sum any *finite* number of terms. More generally, (12.41) should be replaced by the equation

$$\frac{N}{V} = \frac{1}{\lambda^3} g_{3/2}(z) + \frac{\langle n_0 \rangle}{V} + \left(\frac{\langle n_1 \rangle}{V} + \frac{\langle n_2 \rangle}{V} + \dots \right)$$

where, in the parentheses, there appear any finite number of terms. Every term in the parentheses, however, approaches zero as $V \rightarrow \infty$. For example,

$$\frac{\langle n_1 \rangle}{V} = \frac{1}{V} \frac{1}{z^{-1} e^{\beta \epsilon_1} - 1} \leq \frac{1}{V} \frac{1}{e^{\beta \epsilon_1} - 1}$$

where

$$2m\epsilon_1 = (2\pi\hbar)^2 \frac{l_1}{V^{2/3}}$$

l_1 = sum of the squares of three integers not all zero

Hence

$$\frac{\langle n_1 \rangle}{V} \leq \frac{1}{V} \frac{2m\beta V^{2/3}}{(2\pi\hbar)^2 \beta^2 l_1} \underset{V \rightarrow \infty}{\rightarrow} 0 \quad (12.53)$$

This shows that (12.41) is valid.

By (12.52) and the fact that $\langle n_0 \rangle = z/(1-z)$ we can write

$$\frac{\langle n_0 \rangle}{N} = \begin{cases} 0 & \left(\frac{\lambda^3}{v} \leq g_{3/2}(1) \right) \\ 1 - \left(\frac{T}{T_c} \right)^{3/2} = 1 - \frac{v}{v_c} & \left(\frac{\lambda^3}{v} \geq g_{3/2}(1) \right) \end{cases} \quad (12.54)$$

A plot of $\langle n_0 \rangle/N$ is shown in Fig. 12.6. It is seen that when $T < T_c$, a finite fraction of the particles in the system occupy the single level with $p = 0$. On the other hand (12.53) shows that $\langle n_p \rangle/N$ is always zero for $p \neq 0$. Therefore we have the following situation: For $T > T_c$ no single level is occupied by a finite fraction of all the particles. The particles "spread thinly" over all levels. For $T < T_c$ a finite fraction $1 - (T/T_c)^{3/2}$ occupies the level with $p = 0$ while the rest of the particles "spread thinly" over the levels with $p \neq 0$. At absolute zero all particles occupy the level with $p = 0$.

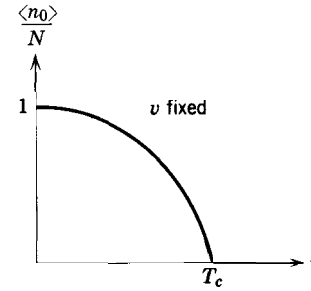


Fig. 12.6 Average occupation number of the level with $p = 0$.

The Bose-Einstein condensation is sometimes described as a "condensation in momentum space." We shall see, however, that its thermodynamic manifestations are those of a first-order phase transition. If we examine the equation of state alone, we discern no difference between the Bose-Einstein condensation and an ordinary gas-liquid condensation. If the particles of the ideal Bose gas are placed in a gravitational field, then in the condensation region there will be a spatial separation of the two phases, just as in a gas-liquid condensation.* The term "momentum-space condensation" merely serves to emphasize the fact that the cause of the Bose-Einstein condensation lies in the symmetry of the wave function and not in any interparticle interaction.

By virtue of (12.52) all thermodynamic functions of the ideal Bose gas will be given by different analytical expressions for the region of condensation and for the complement of that region. Only in the condensation region will these analytical expressions be simple. In the other region numerical computations would be necessary to obtain explicit formulas.

Throughout the remainder of this section let z be defined only for the region $(A^3/u) \leq g_{3/2}(1)$. Some equivalent definitions of z are

$$\begin{aligned} g_{3/2}(z) &= \frac{A^3}{v} \\ \frac{g_{3/2}(z)}{g_{3/2}(1)} &= \frac{v_c}{v} \\ \frac{g_{3/2}(z)}{g_{3/2}(1)} &= \left(\frac{T_c}{T} \right)^{3/2} \end{aligned} \quad (12.55)$$

In the region $(A^3/u) > g_{3/2}(1)$, z need not be mentioned because $z = 1$.

The equation of state can be obtained from (8.71):

$$\frac{P}{kT} = \begin{cases} \frac{1}{\lambda^3} g_{5/2}(z) & (v > v_c) \\ \frac{1}{\lambda^3} g_{5/2}(1) & (v < v_c) \end{cases} \quad (12.56)$$

*W. Lamb and A. Nordsieck, *Phys. Rev.* **59**, 677 (1941).

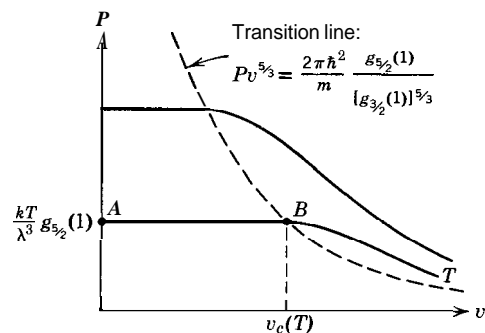


Fig. 12.7 Isotherms of the ideal Bose gas.

where

$$g_{5/2}(1) = \zeta\left(\frac{5}{2}\right) = 1.342\dots \quad (12.57)$$

The term $V^{-1} \log(1 - z)$ in (8.71) is zero as $V \rightarrow \infty$. For $v > v_c$ this is obvious. For $v < v_c$, it is also true, because $(1 - z) \propto V^{-1}$. It is immediately seen that for $v < v_c$, P is independent of u . The isotherms are shown in Fig. 12.7, and the P - T diagram is shown in Fig. 12.8. We may, as in the case of a gas-liquid condensation, interpret the horizontal portion of an isotherm to mean that in that region the system is a mixture of two phases. In the present case these two phases correspond to the two points labeled A and B in Fig. 12.7. We refer to these respectively as the condensed phase and the gas phase. The horizontal portion of the isotherm is the region of phase transition between the two phases. The vapor pressure is

$$P_0(T) = \frac{kT}{\lambda^3} g_{5/2}(1) \quad (12.58)$$

Differentiation of this equation leads to

$$\frac{dP_0(T)}{dT} = \frac{5}{2} \frac{kg_{5/2}(1)}{\lambda^3} = \frac{1}{Tv_c} \left[\frac{5}{2} kT \frac{g_{5/2}(1)}{g_{3/2}(1)} \right] \quad (12.59)$$

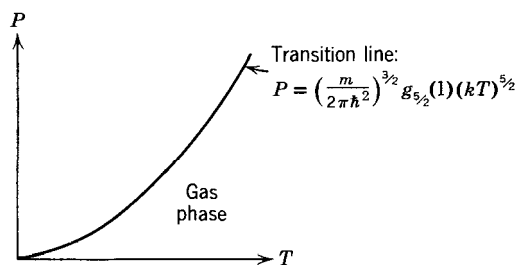


Fig. 12.8 P - T diagram of the ideal Bose gas. Note that the space above the transition curve does not correspond to anything. The condensed phase lies on the transition line itself.

When the two phases coexist the gas phase has the specific volume v_c , whereas the condensed phase has the specific volume 0. Hence the difference in specific volume between the two phases is

$$\Delta v = v_c \quad (12.60)$$

In fact (12.59) is the Clapeyron equation, and the latent heat of transition per particle is

$$L = \frac{g_{5/2}(1)}{g_{3/2}(1)} \frac{5}{2} kT \quad (12.61)$$

Therefore the Bose-Einstein condensation is a first-order phase transition.

Other thermodynamic functions for the ideal Bose gas are given in the following. For each thermodynamic function the upper equation refers to the region $v > v_c$ (or $T > T_c$) and the lower equation refers to the region $v < v_c$ (or $T < T_c$):

$$\frac{U}{N} = \frac{3}{2} P v = \begin{cases} \frac{3}{2} \frac{kT v}{\lambda^3} g_{5/2}(z) \\ \frac{3}{2} \frac{kT v}{\lambda^3} g_{5/2}(1) \end{cases} \quad (12.62)$$

$$-\frac{A}{NkT} = \begin{cases} \frac{v}{\lambda^3} g_{5/2}(z) - \log z \\ \frac{v}{\lambda^3} g_{5/2}(1) \end{cases} \quad (12.63)$$

$$\frac{G}{NkT} = \begin{cases} \log z \\ 0 \end{cases} \quad (12.64)$$

$$\frac{S}{Nk} = \begin{cases} \frac{5}{2} \frac{v}{\lambda^3} g_{5/2}(z) - \log z \\ \frac{5}{2} \frac{v}{\lambda^3} g_{5/2}(1) \end{cases} \quad (12.65)$$

$$\frac{C_V}{Nk} = \begin{cases} \frac{15}{4} \frac{v}{\lambda^3} g_{5/2}(z) - \frac{9}{4} \frac{g_{3/2}(z)}{g_{1/2}(z)} \\ \frac{15}{4} \frac{v}{\lambda^3} g_{5/2}(1) \end{cases} \quad (12.66)$$

The specific heat is shown in Fig. 12.9. Near absolute zero, C_V vanishes like $T^{3/2}$. This behavior is to be contrasted with a photon gas or a phonon gas, for which C_V vanishes like T^3 near absolute zero. The reason for this difference lies in the difference between the particle spectrum $\epsilon_p = p^2/2m$ and the photon or phonon spectrum $\epsilon_p = cp$. At the same energy the particle spectrum has a higher density of states than the photon or phonon spectrum. Consequently there are more modes of excitation available for a particle, and the specific heat is greater.

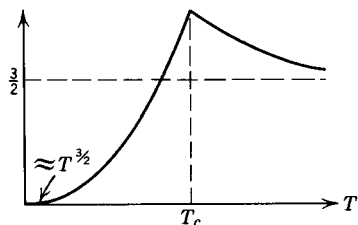


Fig. 12.9 Specific heat of the ideal Bose gas.

From (12.65) we see that $S = 0$ at $T = 0$, in accordance with the third law of thermodynamics. This means that the condensed phase (which exists at $T = 0$) has no entropy. At any finite temperature the total entropy is entirely due to the gas phase. The fraction of particles in the gas phase in the transition region is v/v_c , or $(T/T_c)^{3/2}$. If we rewrite S in the transition region in the form

$$\frac{S}{N} = \left(\frac{T}{T_c}\right)^{3/2} s = \left(\frac{v}{v_c}\right) s \quad (12.67)$$

we find that

$$s = \frac{g_{5/2}(1)}{g_{3/2}(1)} \frac{5}{2} k \quad (12.68)$$

which is the entropy per particle of the gas phase. The difference in specific entropy between the gas phase and the condensed phase is

$$\Delta s = s = \frac{g_{5/2}(1)}{g_{3/2}(1)} \frac{5}{2} k \quad (12.69)$$

Comparing this with (12.61), we find that

$$L = T\Delta s \quad (12.70)$$

This shows that the interpretation of the Bose-Einstein condensation as a first-order phase transition is self-consistent.

The only Bose system known to exist at low temperatures is liquid He^4 . At a temperature of 2.18 K, He^4 exhibits the remarkable λ transition, at which the specific heat becomes logarithmically infinite. Since He^4 atoms obey Bose statistics, it is natural to suppose that this transition is the Bose-Einstein condensation modified by intermolecular interactions. This is supported by the fact that no such transition occurs in liquid He^3 , whose atoms obey Fermi statistics. Furthermore, substituting the mass of He^4 and the density of liquid helium into (12.50) leads to the transition temperature $T_c = 3.14$ K, which is of the right order of magnitude.

Finally, we must re-emphasize that Bose-Einstein condensation can occur only when the particle number is conserved. For example, photons do not condense. They have a simpler alternative, namely, to disappear into the vacuum.

We have pointed out in Section 7.5 that heavy-particle conservation as physically observed is a low-energy approximation to the real conservation law, which says that the conserved quantity is the number of particles minus the number of antiparticles. Thus, any discussion of the Bose-Einstein condensation for a relativistic Bose gas must take antiparticles into account.*

12.4 AN IMPERFECT BOSE GAS

The ideal Bose gas is an artificial example in that the particles condense into a highly idealized phase with infinite compressibility. That is, the Bose-Einstein condensate is unphysical and uninteresting. We now study an interacting Bose gas in a crude approximation, to see how the nature of the Bose-Einstein condensation changes.

The Energy Levels

We consider a dilute system of N identical spinless bosons of mass m , contained in a box of volume V , at very low temperatures. The bosons interact with one another through binary collisions characterized by the scattering length a which is assumed to be positive. The energy levels to the first order in a may be obtained from (10.124) through the use of first-order perturbation theory.

Let the unperturbed wave functions be free-particle wave functions Φ_n , labeled by the occupation numbers $\{\dots, n_{\mathbf{p}}, \dots\}$, where $n_{\mathbf{p}}$ is the number of bosons with momentum \mathbf{p} . The energy levels to the first order in a are

$$E_n \equiv (\Phi_n, \mathcal{H}' \Phi_n) = \sum_{\mathbf{p}} \frac{p^2}{2m} n_{\mathbf{p}} + \frac{4\pi a \hbar^2}{m} \left(\Phi_n, \sum_{i < j} \delta(\mathbf{r}_i - \mathbf{r}_j) \Phi_n \right) \quad (12.71)$$

The second term is calculated in (A.36) of the Appendix. With that, we have

$$E_n = \sum_{\mathbf{p}} \frac{p^2}{2m} n_{\mathbf{p}} + \frac{4\pi a \hbar^2}{mV} \left(N^2 - \frac{1}{2} \sum_{\mathbf{p}} n_{\mathbf{p}}^2 \right) \quad (12.72)$$

This formula is valid only under the conditions

$$\frac{a}{v^{1/3}} \ll 1 \quad (12.73)$$

$$ka \ll 1$$

where k is the relative wave number of any pair of particles. Thus (12.72) becomes invalid if there are excited particles of high momentum.

Let us first study the implications of (12.72). The ground state energy per particle is obtained from (12.72) by setting all $n_{\mathbf{p}} = 0$ for $\mathbf{p} \neq 0$, and $n_0 = N$:

$$\frac{E_0}{N} = \frac{2\pi a \hbar^2}{mv} = \left(\frac{\hbar}{m} \right)^2 2\pi a \rho \quad (12.74)$$

*H. E. Haber and H. A. Weldon, *Phys. Rev. Lett.* **46**, 1497 (1981)

where p is the mass density. It is proportional to the scattering length a and to the mass density, and it may be interpreted to be the energy shift of an average particle in the "optical approximation," whereby the effect of the rest of the system is replaced by a medium having an index of refraction. This interpretation can be justified as follows. In the shape-independent approximation we may replace a scattering potential by one of any shape, provided it gives the same scattering length. Let us replace the interparticle potential by a very shallow but very long-ranged square well such that the scattering length is still a . Now a particle moving through the system essentially "sees" a uniform potential of an appropriate depth. This gives (12.74).

For an excited state in which the particles have vanishingly small momenta the energy per particle is

$$\frac{E_n}{N} = \left(\frac{\hbar}{m}\right)^2 4\pi a \rho \left[1 - \frac{1}{2} \sum_{\mathbf{p}} \left(\frac{n_{\mathbf{p}}}{N}\right)^2\right] \quad (12.75)$$

The second term is most negative when all the excited particles are in the same momentum state. Thus we may say that "spatial repulsion leads to momentum space attraction." This is a consequence of the symmetry of the wave function.*

The "momentum space attraction" just mentioned also leads to an "energy gap" in the spectrum (12.72). This may be seen as follows. The energies of the very low excited states of the system are approximately given by

$$E_n \approx \sum_{\mathbf{p}} \frac{p^2}{2m} n_{\mathbf{p}} + N \left(\frac{\hbar}{m}\right)^2 4\pi a \rho \left[1 - \frac{1}{2} \left(\frac{n_0}{N}\right)^2\right] \quad (12.76)$$

According to this formula, the excitation of one particle from the momentum state $p = 0$ to a state of infinitesimal momentum changes the energy by the *finite* amount

$$\Delta = \left(\frac{\hbar}{m}\right)^2 2\pi a \rho \quad (12.77)$$

Thus the single-particle energy spectrum is separated from the zero point of energy by the amount A . This "energy gap," however, is a feature only of the lowest-order formula. When the energy levels are calculated to higher orders in perturbation theory,[†] the energy gap disappears. Instead, there is only a decrease of level density just above the ground state, changing the single-particle spectrum $p^2/2m$ into a phonon spectrum $\hbar c p/2m$, where c is a constant. The "energy gap," which implies that the level density is strictly zero just above the ground state, is a crude approximation to the actual state of affairs.

The foregoing discussions make it clear that the energy levels (12.72), although not exact, possess many qualitative features of the effect of a repulsive interaction among bosons. We use them to calculate the partition function. The

validity of this calculation is discussed as we proceed. We introduce a further simplification, namely, we take the energy levels to be

$$E_n = \sum_{\mathbf{p}} \frac{p^2}{2m} n_{\mathbf{p}} + \frac{4\pi a \hbar^2}{mV} \left(N^2 - \frac{1}{2} n_0^2\right) \quad (12.78)$$

The behavior of the model defined by (12.78) should be qualitatively the same as that by (12.72) when the temperature is so low that few particles are excited.*

The Equation of State

For the calculation of the partition function, we confine our considerations to the region in which

$$a/\lambda \ll 1, \quad a\lambda^2/v \ll 1 \quad (12.79)$$

because these are the only dimensionless parameters in the problem involving a , and our model is valid only to the first order in a .

Let n be an abbreviation for $\{n_i\}$, and let ϵ_n denote the first term of (12.78), the unperturbed energy. Introducing the parameter

$$\xi \equiv \frac{n_0}{N} \quad (12.80)$$

and denoting, as usual, the thermal wavelength by $A = \sqrt{2\pi\hbar^2/mkT}$, we can write the partition function in the form

$$\mathcal{Q}_n = \sum_n e^{-\beta\epsilon_n} e^{-N(a\lambda^2/v)(2-\xi^2)} = \mathcal{Q}_N^{(0)} \langle e^{-N(a\lambda^2/v)(2-\xi^2)} \rangle_0 \quad (12.81)$$

where $\mathcal{Q}_N^{(0)}$ is the partition function of the ideal Bose gas, and $\langle \rangle_0$ denotes the thermodynamic average with respect to the ideal Bose gas. Hence the free energy per particle is

$$\begin{aligned} \frac{A}{N} &= \frac{A^{(0)}}{N} + \frac{kT}{N} \log \langle e^{-N(a\lambda^2/v)(2-\xi^2)} \rangle_0 \\ &\approx \frac{A^{(0)}}{N} + \frac{kT}{N} \frac{a\lambda^2}{v} \langle (2-\xi^2) \rangle_0 \end{aligned} \quad (12.82)$$

It can be easily verified that the fluctuations of $\langle n_0 \rangle$ are small. In fact, for any k

$$\langle n_k^2 \rangle_0 - \langle n_k \rangle_0^2 = \langle n_k \rangle_0 \quad (12.83)$$

Hence the mean-square fluctuation of $\langle \xi \rangle_0$ is of the order $1/N$. Hence

$$\frac{A}{N} = \frac{A^{(0)}}{N} + \frac{\hbar^2}{m} \frac{4\pi a}{v} \left(1 - \frac{1}{2} \bar{\xi}^2\right) \quad (12.84)$$

*For a detailed study of the equation of state based on the more accurate energy levels (12.72), see K. Huang, C. N. Yang, and J. M. Luttinger, *Phys. Rev.* 105, 776 (1957).

*See Problem 12.7.

†See Section 13.8.

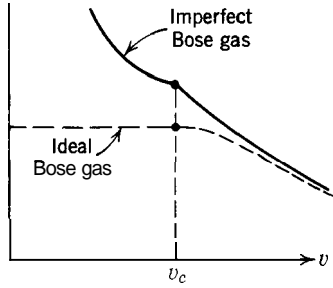


Fig. 12.10 Isotherm of an imperfect Bose gas with repulsive interactions.

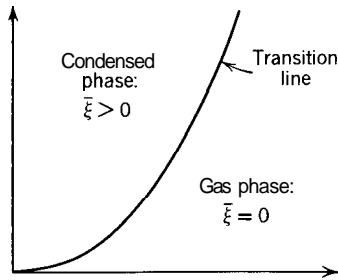


Fig. 12.11 P-T diagram of the imperfect Bose gas. In contradistinction to that of the diagram of the ideal Bose gas, Fig. 12.8, the space above the transition now corresponds to the condensed phase.

where $\bar{\xi} \equiv \langle \xi \rangle_0 = \langle n_0/N \rangle_0$ is given by (12.54). This result is extremely simple, being the free energy of the ideal gas plus the interaction term in (12.78), except that the quantum number n_0 is replaced by its thermodynamic average with respect to the ideal gas, thereby turning it into a thermodynamic parameter.

The pressure can be immediately obtained:

$$P = P^{(0)} + \frac{4\pi a \hbar^2}{m} \left[\frac{1}{v^2} \left(1 - \frac{1}{2} \bar{\xi}^2 \right) + \frac{1}{v} \bar{\xi} \frac{\partial \bar{\xi}}{\partial v} \right] \quad (12.85)$$

where $P^{(0)}$ is the pressure of the ideal Bose gas. Using (12.54) to evaluate $\bar{\xi}$ and $\partial \bar{\xi} / \partial v$, we obtain

$$P = \begin{cases} P^{(0)} + \frac{4\pi a \hbar^2}{m v^2} & (v > v_c, T > T_c) \\ P^{(0)} + \frac{2\pi a \hbar^2}{m} \left(\frac{1}{v^2} + \frac{1}{v_c^2} \right) & (v < v_c, T < T_c) \end{cases} \quad (12.86)$$

An isotherm is shown in Fig. 12.10, and the P-T diagram is shown in Fig. 12.11. The Bose-Einstein condensation is here a second-order transition. The specific heat decreases across the transition point by the amount

$$\frac{\Delta C_v}{Nk} = \frac{9a}{2\lambda_c} g_{3/2}(1) \quad (12.87)$$

We cannot deduce from these results that an imperfect Bose gas with repulsive interactions generally exhibits a second-order transition. The present model merely shows that the transition appears to be a second-order transition if higher-order effects in a/λ and $a\lambda^2/v$ are neglected.

The model here is more realistic than the ideal Bose gas in that the condensed phase now has finite compressibility. We see from (12.86) that the isothermal compressibility increases discontinuously by a factor of 2 when we go from the gas to the condensed phase.

In a nonequilibrium situation, the condensed phase presumably can flow like a real substance. Since the system has a paucity of energy levels just above the ground state, we expect it to flow with little or no dissipation. This is the phenomenon of **superfluidity** that we shall discuss at greater length in the next chapter.

12.5 THE SUPERFLUID ORDER PARAMETER

We have suggested that the Bose-Einstein condensate is a "superfluid" in more realistic systems than the ideal Bose gas. This idea will be taken up in the next chapter. Here we shall analyze in greater depth the implications of the existence of a condensate. It will be necessary to use the quantized-field description of a many-body system (sometimes called "second quantization"), as discussed in the appendix.

Identifying the Order Parameter

Let us begin by examining the one-particle density matrix

$$\rho_1(\mathbf{x}, \mathbf{y}) = \langle \psi^\dagger(\mathbf{x}) \psi(\mathbf{y}) \rangle = \frac{1}{V} \sum_{\mathbf{k}, \mathbf{q}} e^{i(\mathbf{k} \cdot \mathbf{x} - \mathbf{q} \cdot \mathbf{y})} \langle a_{\mathbf{q}}^\dagger a_{\mathbf{k}} \rangle \quad (12.88)$$

where $\langle \rangle$ denotes ensemble average, and $\psi(\mathbf{x})$ is the quantized boson field operator, which is expanded in terms of annihilation operators $a_{\mathbf{k}}$ for a plane-wave state of wave vector \mathbf{k} (see (A.65)). Roughly speaking, this is the probability that, having lost a particle at \mathbf{x} , you will find one at \mathbf{y} .

Consider first a translationally invariant system. Using the fact that the total momentum operator \mathbf{P} commutes with the Hamiltonian, we can verify

$$\langle [\mathbf{P}, a_{\mathbf{q}}^\dagger a_{\mathbf{k}}] \rangle = 0$$

by writing out the trace and using $\text{Tr}(AB) = \text{Tr}(BA)$. On the other hand, a direct calculation gives

$$[\mathbf{P}, a_{\mathbf{q}}^\dagger a_{\mathbf{k}}] = \hbar(\mathbf{k} - \mathbf{q}) a_{\mathbf{q}}^\dagger a_{\mathbf{k}}$$

Hence for a translationally invariant system

$$\langle a_{\mathbf{q}}^\dagger a_{\mathbf{k}} \rangle = \delta_{\mathbf{q}, \mathbf{k}} \langle n_{\mathbf{k}} \rangle \quad (12.89)$$

where $n_{\mathbf{k}} = a_{\mathbf{k}}^\dagger a_{\mathbf{k}}$. Thus

$$\begin{aligned} \rho_1(\mathbf{x}, \mathbf{y}) &= \frac{1}{V} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})} \langle n_{\mathbf{k}} \rangle \\ &= \frac{\langle n_0 \rangle}{V} + \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})} \langle n_{\mathbf{k}} \rangle \end{aligned} \quad (12.90)$$

where we have separated out the $\mathbf{k} = 0$ term before passing to the limit $V \rightarrow \infty$. The second term in (12.90) vanishes when $|\mathbf{x} - \mathbf{y}| \rightarrow \infty$, because in that limit the integral gets contributions only from the neighborhood of $\mathbf{k} = 0$ (see Problem 12.9). Thus

$$\rho_1(\mathbf{x}, \mathbf{y}) \xrightarrow{|\mathbf{x} - \mathbf{y}| \rightarrow \infty} \frac{\langle n_0 \rangle}{V} \quad (12.91)$$

This does not imply a positional correlation over infinite distances, as it might seem at first sight. It says only that there is a constant density of zero-momentum particles over the entire system.

Note that in a general interacting system, the single-particle momenta are not good quantum numbers. In particular n_0 does not commute with the Hamiltonian. But $\langle n_0 \rangle / N$ can still be used as a characterization of a Bose-Einstein condensate.

In analogy with (12.91), Penrose and Onsager* proposed the following general criterion for Bose-Einstein condensation:

$$\langle \psi^\dagger(\mathbf{x}) \psi(\mathbf{y}) \rangle \xrightarrow{|\mathbf{x} - \mathbf{y}| \rightarrow \infty} f^*(\mathbf{x}) f(\mathbf{y}) \quad (12.92)$$

To be of practical value, this criterion must be applicable to real systems with nontranslationally invariant geometry, and under nonequilibrium situations. It is then not obvious (in fact, it is somewhat of a mystery) how the criterion can be satisfied. Consider, for example, a Bose fluid contained in two separate tanks connected by a pipe a mile long. (For added realism, have an experimentalist kick the apparatus from time to time.) Suppose \mathbf{x} lies in one tank, and \mathbf{y} the other. It is physically absurd to suppose that there can be any correlation between \mathbf{x} and \mathbf{y} . But then how does it come about that these separate points are characterized by the same function f ?

The way out is to make f dependent only on local dynamical variables. It is now generally accepted that the correct choice is $f = \langle \psi \rangle$.[†] Thus the Penrose-Onsager criterion takes the more specific form

$$\langle \psi^\dagger(\mathbf{x}) \psi(\mathbf{y}) \rangle \xrightarrow{|\mathbf{x} - \mathbf{y}| \rightarrow \infty} \langle \psi^\dagger(\mathbf{x}) \rangle \langle \psi(\mathbf{y}) \rangle \quad (12.93)$$

*O. Penrose, *Philos. Mag.* 42, 1373 (1951); O. Penrose and L. Onsager, *Phys. Rev.* 104, 576 (1956).

[†]The first use of this was implicit in N. N. Bogoliubov, *J. Phys. USSR* 11, 23 (1947). J. Goldstone, *N. Cim.*, 19, 154 (1961), clarified its meaning in terms of "broken symmetry." For a review see P. W. Anderson, *Rev. Mod. Phys.* 38, 298 (1966).

We call the complex number

$$\langle \psi(\mathbf{x}) \rangle = r(\mathbf{x}) e^{i\phi(\mathbf{x})} \quad (12.94)$$

the superfluid order parameter. The fact that $r(\mathbf{x}) > 0$ implies the existence of momentum-space order, i.e., a Bose-Einstein condensate. As we shall see in the next chapter, the phase $\phi(\mathbf{x})$ is the velocity potential for superfluid flow.

Spontaneous Symmetry Breaking

The ensemble average (I) should be taken in a grand canonical ensemble, because we are interested in open systems, where the number of particles is not definite. Thus,

$$\begin{aligned} \langle \psi(x) \rangle &= \frac{\text{Tr} [e^{-\beta \mathcal{E}[\psi]} \psi(x)]}{\text{Tr} e^{-\beta \mathcal{E}[\psi]}} \\ \mathcal{E}[\psi] &= \mathcal{H} - \mu \mathcal{N} \end{aligned} \quad (12.95)$$

where \mathcal{H} is the Hamiltonian and \mathcal{N} is the number operator. The notation $\mathcal{E}[\psi]$ indicates that \mathcal{E} is a functional of ψ . The immediate question is why the ensemble average above should not be always zero. There is a trivial and a subtle aspect to this question.

First, the trivial aspect. Since ψ annihilates a particle, its expectation value with respect to any eigenstate of \mathcal{N} is zero. This makes one feel uneasy about taking its ensemble average. In the grand canonical ensemble, however, the relevant basis are not the simultaneous eigenstates of \mathcal{H} and \mathcal{N} , but those of $\mathcal{H} - \mu \mathcal{N}$, and the latter need not be eigenstates of \mathcal{N} . In the infinite-volume limit, the eigenvalues of $\mathcal{H} - \mu \mathcal{N}$ are highly degenerate: systems with different particle numbers can maintain the same eigenvalue by going into different energy levels. One can form a new basis by superimposing these degenerate states (of different particle numbers), with respect to which the expectation value of ψ will have the form (12.94). The phase ϕ labels the degenerate states.

But, one argues, the ensemble of ψ is still zero, because one has the freedom to calculate the trace using a basis with definite particle numbers. This is true, and is a reflection of the fact that particle number is conserved, which can be expressed formally by saying that the Hamiltonian has a "global gauge invariance"—an invariance under the transformation

$$\psi(\mathbf{x}) \rightarrow e^{i\alpha} \psi(\mathbf{x}) \quad (12.96)$$

where α is an arbitrary real number. Thus, in the ensemble average, every value $\psi = r \exp(i\alpha)$ will be canceled by a value $\psi = r \exp(i\alpha + i\pi)$ of equal weight. This argument is technically correct, and illustrates the need to redefine the ensemble average more carefully. This is the subtle part of the problem.

There is a parallel between (I) and the spontaneous magnetization of a ferromagnet:

$$\langle M \rangle = \frac{\text{Tr} (e^{-\beta \mathcal{H}} M)}{\text{Tr} e^{-\beta \mathcal{H}}} \quad (12.97)$$

where \mathbf{M} is the total magnetic moment. Since the Hamiltonian \mathcal{H} in the absence of external field is invariant under rotations, the ensemble average of \mathbf{M} is always zero because \mathbf{M} and $-\mathbf{M}$ occur with equal probability. The mathematical correctness of this statement is irrefutable. But we know that it is physically the wrong answer, for we do have ferromagnets in nature.

The resolution of the apparent paradox lies in the recognition that the symmetry of a system may be "spontaneously broken," in that the ground state of a Hamiltonian does not possess the symmetry of the Hamiltonian. This requires that the ground state be degenerate. The symmetry is realized by the fact that any one of the degenerate ground states is equally as good as the physical ground state, and by the existence of characteristic "Goldstone excitations."* For a ferromagnet, the ground state is not rotationally invariant, because the magnetization points along a definite axis in space. The Goldstone excitations in this case are the spin waves.

The essential point in the present context is that, once the system magnetizes along a certain direction it cannot make a transition to another direction, even though doing so requires no expenditure in energy. For to do so requires that all the atomic magnetic moments in the system spontaneously and simultaneously rotate through exactly the same angle. The probability for this to happen is essentially zero for a macroscopic system. (One would have to wait for a time of the order of a Poincaré cycle to see this happen.) The ensemble average has physical significance only if it corresponds to time averages over microscopic relaxation times. One must therefore redefine it in such a way that \mathbf{M} and $-\mathbf{M}$ are not both included among the configurations. This can be done most simply by placing the system in an external field \mathbf{H} pointing along an arbitrary but fixed direction, and calculating the ensemble average in the limit $\mathbf{H} \rightarrow 0$. To emphasize the importance of various limiting processes, we explicitly indicate the infinite-volume limit:

$$\frac{\langle \mathbf{M} \rangle}{V} \equiv \lim_{H \rightarrow 0} \lim_{V \rightarrow \infty} \frac{1}{V} \frac{\text{Tr} [e^{-\beta(\mathcal{H} - \mathbf{M}\mathbf{H})} \mathbf{M}]}{\text{Tr} e^{-\beta(\mathcal{H} - \mathbf{M}\mathbf{H})}} \quad (12.98)$$

The thermodynamic limit of (12.97), which is not physically relevant, corresponding a reversal of the limiting process above[†]:

$$\lim_{V \rightarrow \infty} \lim_{H \rightarrow 0} \frac{\text{Tr} [e^{-\beta(\mathcal{H} - \mathbf{M}\mathbf{H})} \mathbf{M}]}{\text{Tr} e^{-\beta(\mathcal{H} - \mathbf{M}\mathbf{H})}} = 0 \quad (12.99)$$

Returning to the Bose system, we see that Bose-Einstein condensation corresponds to a spontaneous breaking of the global gauge invariance. In analogy with ferromagnetism, we imagine subjecting the system to an external field coupled to $\psi(\mathbf{x})$, calculate the ensemble average of $\psi(\mathbf{x})$ in the thermodynamic

*J. Goldstone, *op. cit.* A brief discussion of this phenomenon will be given in Section 16.6.

[†]Note that in calculating the spontaneous magnetization in the model in Section 11.6, we in effect used the correct average (12.98) instead of (12.99), because we ignored the $-\mathbf{M}$ solution (by common sense).

limit, and then let the external field go to zero:

$$\langle \psi(\mathbf{x}) \rangle = \lim_{\eta \rightarrow 0} \lim_{V \rightarrow \infty} \frac{\text{Tr} [e^{-\beta \mathcal{E}[\psi, \eta]} \psi(\mathbf{x})]}{\text{Tr} e^{-\beta \mathcal{E}[\psi, \eta]}} \quad (12.100)$$

where

$$\mathcal{E}[\psi, \eta] = \mathcal{H} - \mu \mathcal{N} - \int d^3x [\psi(\mathbf{x})\eta(\mathbf{x}) + \psi^\dagger(\mathbf{x})\eta^\dagger(\mathbf{x})] \quad (12.101)$$

The only essential difference with the ferromagnetic case is that, unlike the magnetic field, the external field $\eta(\mathbf{x})$ here is a mathematical device that cannot be realized experimentally.*

PROBLEMS

12.1 (a) Show that the entropy per photon in blackbody radiation is independent of the temperature, and in d spatial dimensions is given by

$$s = (d + 1) \frac{\sum_{n=1}^{\infty} n^{-d-1}}{\sum_{n=1}^{\infty} n^{-d}}$$

(b) Show that the answer would have been $d + 1$ if the photons obeyed Boltzmann statistics.

12.2 Some experimental values[†] for the specific heat of liquid He^4 are given in the accompanying table. The values are obtained along the vapor pressure curve of liquid He^4 , but we may assume that they are not very different from the values of c_V at the same temperatures.

Temperature (K)	Specific Heat (joule/g-deg)
0.60	0.0051
0.65	0.0068
0.70	0.0098
0.75	0.0146
0.80	0.0222
0.85	0.0343
0.90	0.0510
0.95	0.0743
1.00	0.1042

*The Bose-Einstein condensation of the ideal gas has been reanalyzed in terms of the superfluid order parameter by J. D. Gunton and M. J. Buckingham, *Phys. Rev.* 166, 152 (1968).

[†]Taken from H. C. Kramers, "Some Properties of Liquid Helium below 1°K," Dissertation, Leiden (1955).

(a) Show that the behavior of the specific heat at very low temperatures is characteristic of that of a gas of phonons.

(b) Find the velocity of sound in liquid He⁴ at low temperature.

12.3 Equation (12.64) states that $G = 0$ for $v < v_c$. Using the formula $S = -(\partial G/\partial T)_p$, we would obtain $S = 0$ for $v < v_c$, in contradiction to (12.65). What is wrong with the previous statement?

12.4 In the neighborhood of $z = 1$ the following expansion may be obtained (F. London, *loc. cit.*):

$$g_{5/2}(z) = 2.363v^{3/2} + 1.342 - 2.612v - 0.730v^2 + \dots$$

where $v \equiv -\log z$. From this the corresponding expansions for $g_{3/2}$, $g_{1/2}$, and $g_{-1/2}$ may be obtained by the recursion formula $g_{n-1} = -\partial g_n/\partial v$. Using this expansion show that for the ideal Bose gas the discontinuity of $\partial C_V/\partial T$ at $T = T_c$ is given by

$$\left(\frac{\partial C_V}{\partial T Nk}\right)_{T \rightarrow T_c^+} - \left(\frac{\partial C_V}{\partial T Nk}\right)_{T \rightarrow T_c^-} = \frac{3.66}{T_c}$$

12.5 Show that the equation of state of the ideal Bose gas in the gas phase can be written in the form of a virial expansion, i.e.,

$$\frac{Pv}{kT} = 1 - \frac{1}{4\sqrt{2}} \left(\frac{\lambda^3}{v}\right) + \left(\frac{1}{8} - \frac{2}{9\sqrt{3}}\right) \left(\frac{\lambda^3}{v}\right)^2 - \dots$$

12.6 (a) Calculate the grand partition function $\mathcal{Q}(z, V, T)$ for a two-dimensional ideal Bose gas and obtain the limit

$$\lim_{V \rightarrow \infty} \frac{1}{V} \log \mathcal{Q}(z, V, T)$$

where $V = L^2$ is the area available to the system.

(b) Find the average number of particles per unit area as a function of z and T .

(c) Show that there is no Bose-Einstein condensation for a two-dimensional ideal Bose gas.

12.7 Consider two free bosons contained in a box of volume V with periodic boundary conditions. Let the momenta of the two particles be p and q .

(a) Write down the normalized wave function $\psi_{pq}(\mathbf{r}_1, \mathbf{r}_2)$ for both $p \neq q$ and $p = q$.

(b) Show that for $p \neq q$

$$|\psi_{pq}(\mathbf{r}, \mathbf{r})|^2 > |\psi_{pp}(\mathbf{r}, \mathbf{r})|^2$$

(c) Explain the meaning of the statement "spatial repulsion leads to momentum space attraction."

12.8 For the imperfect Bose gas discussed in Section 12.4, show that in the gas phase

$$\frac{Pv}{kT} = 1 + \left(-\frac{1}{4\sqrt{2}} + \frac{2a}{\lambda}\right) \frac{\lambda^3}{v} + \left(\frac{1}{8} - \frac{1}{3\sqrt{3}}\right) \left(\frac{\lambda^3}{v}\right)^2 - \dots$$

Thus we can conclude that the third and higher virial coefficients, if they depend on a , must involve orders of a^2 or higher.

12.9 Consider an ideal Bose gas. Let $\psi(\mathbf{x})$ be the boson field operator.

(a) Show

$$\langle \psi^\dagger(\mathbf{x})\psi(\mathbf{y}) \rangle = \frac{\langle n_0 \rangle}{V} + f(|\mathbf{x} - \mathbf{y}|)$$

where

$$f(r) = \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{r}} \langle n_k \rangle = \frac{mkT}{2\hbar^2} \frac{e^{-r/r_0}}{r}$$

with $r_0 = \hbar/\sqrt{2mkT|\log z|}$

(b) Let $T \rightarrow T_c$ from the high-temperature side. Find r_0 as a function of $t = (T - T_c)/T_c$ as $t \rightarrow 0$.

(c) The density-density correlation function is defined as

$$\Gamma(\mathbf{x}) \equiv \langle \rho(\mathbf{x})\rho(0) \rangle - (N/V)^2$$

where $\rho(\mathbf{x}) = \psi^\dagger(\mathbf{x})\psi(\mathbf{x})$ is the density operator. Show

$$\Gamma(\mathbf{x}) = \frac{1}{V^2} \sum_{\mathbf{k} \neq \mathbf{q}} e^{i(\mathbf{k}-\mathbf{q})\cdot\mathbf{x}} \langle n_{\mathbf{q}}(n_{\mathbf{k}} + 1) \rangle \xrightarrow{V \rightarrow \infty} \left| \frac{1}{V} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}} \langle n_{\mathbf{k}} \rangle \right|^2$$

Work out $\Gamma(\mathbf{x})$ more explicitly, using the results of (a) and (b).