

CHAPTER 3

Statistical Mechanics

We now turn our attention to the molecular foundation of thermodynamics, or more generally, the answer to the following question: If particles (atoms, molecules, or electrons and nuclei, ...) obey certain microscopic laws with specified interparticle interactions, what are the observable properties of a system containing a very large number of such particles? That is, we want to discuss the relationship between the microscopic dynamics or fluctuations (as governed by Schrödinger's equation or Newton's laws of motion) and the observed properties of a large system (such as the heat capacity or equation of state).

The task of solving the equations of motion for a many-body system (say N = number of particles $\sim 10^{23}$) is so complicated that even modern day computers find the problem intractable. (Though scientists do use computers to follow the motion of thousands of particles for times often long enough to simulate condensed phases for times of the order of 10^{-10} or 10^{-9} sec.) At first you might think that as the number of particles increases, the complexity and obscurity of the properties of a mechanical system should increase tremendously, and that you would be unable to find any regularity in the behavior of a macroscopic body. But as you know from thermodynamics, large systems are, in a sense, quite orderly. An example is the fact that at thermodynamic equilibrium one can characterize observations of a macroscopic system with only a handful of variables. The attitude we shall take is that these distinctive regularities are consequences of statistical laws governing the behavior of systems composed of very many particles. We will thereby avoid the need to directly evaluate the precise N -particle

dynamics, and assume that probability statistics provides the correct description of what we see during a macroscopic measurement.

The word "measurement" is important in these remarks. If we imagined, for example, observing the time evolution of one particular particle in a many-body system, its energy, its momentum, and its position would all fluctuate widely, and the precise behavior of any of these properties would change drastically with the application of the slightest perturbation. One cannot imagine a reproducible measurement of such chaotic properties since even the act of observation involves a perturbation. Further, to reproduce the precise time evolution of a many-body system, one must specify at some initial time a macroscopic number ($\sim 10^{23}$) of variables. These variables are initial coordinates and momenta of all the particles if they are classical, or an equally cumbersome list of numbers if they are quantal. If we would fail to list just one of these 10^{23} variables, the time evolution of the system would no longer be deterministic, and an observation that depended upon the precise time evolution would no longer be reproducible. It is beyond our capacity to control 10^{23} variables. As a result, we confine our attention to simpler properties, those controlled by only a few variables. In some areas of physical and biological science, it might not be easy to identify those variables. But as a philosophical point, scientists approach most observations with an eye to discovering which small number of variables guarantees the reproducibility of phenomena.

The use of statistics for reproducible phenomena does not imply that our description will be entirely undeterministic or vague. To the contrary, we will be able to predict that the observed values of many physical quantities remain practically constant and equal to their average values, and only very rarely show any detectable deviations. (For example, if one isolates a small volume of gas containing, say, only 0.01 moles of gas, then the average relative deviation of the energy of this quantity from its mean value is $\sim 10^{-11}$. The probability of finding in a single measurement a relative deviation of 10^{-6} is $\sim 10^{-3 \times 10^{15}}$.) As a rough rule of thumb: If an observable of a many particle system can be specified by a small number of other macroscopic properties, we assume that the observable can be described with statistical mechanics. For this reason, statistical mechanics is often illustrated by applying it to equilibrium thermodynamic quantities.

3.1 The Statistical Method and Ensembles

While it is not possible in practice, let us imagine that we could observe a many-body system in a particular microscopic state. Its

characterization would require an enormous number of variables. For example, suppose the system was quantal obeying Schrödinger's equation

$$i\hbar \frac{\partial}{\partial t} |\psi\rangle = \mathcal{H} |\psi\rangle.$$

Here, as always, $2\pi\hbar$ is Planck's constant, \mathcal{H} is the Hamiltonian operating on the state vector $|\psi\rangle$, and t is the time. To specify the state $|\psi\rangle$ at a particular time, we need a number of variables of the order of N , the number of particles in the system.

Consider, for example, stationary solutions

$$\mathcal{H} |\psi_v\rangle = E_v |\psi_v\rangle,$$

and some simple and familiar quantum mechanical systems such as the hydrogen atom, or non-interacting particles in a box. The index v is then the collection of $D \cdot N$ quantum numbers, where D is the dimensionality.

Once the initial state is specified, if it could be, the state at all future times is determined by the time integration of Schrödinger's equation. The analogous statement for classical systems considers points in *phase space*

$$(r^N, p^N) \equiv (r_1, r_2, \dots, r_N; p_1, \dots, p_N),$$

where r_i and p_i are the coordinates and conjugate momenta, respectively, for particle i . Points in phase space characterize completely the mechanical (i.e., microscopic) state of a classical system, and flow in this space is determined by the time integration of Newton's equation of motion, $F = ma$, with the initial phase space point providing the initial conditions.

Exercise 3.1 Write down the differential equations corresponding to Newton's laws when the total potential energy is the function $U(r_1, r_2, \dots, r_N)$.

Now try to think about this time evolution—the trajectory—of a many-body system. As illustrated in Fig. 3.1, we might picture the evolution as a line in “state space” (phase space in the classical case, or Hilbert space spanned by all the state vectors $|\psi\rangle$ in the quantal case). In preparing the system for this trajectory a certain small number of variables is controlled. For example, we might fix the total energy, E , the total number of particles, N , and the volume, V . These constraints cause the trajectory to move on a “surface” of state space—though the dimensionality of the surface is still enormously high.

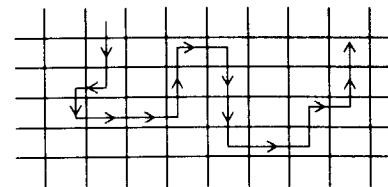


Fig. 3.1. Trajectory in state space with each box representing a different state.

A basic concept in statistical mechanics is that if we wait long enough, the system will eventually flow through (or arbitrarily close to) all the microscopic states consistent with the constraints we have imposed to control the system. Suppose this is the case, and imagine that the system is constantly flowing through the state space as we perform a multitude \mathcal{N} of independent measurements on the system. The observed value ascertained from these measurements for some property G is

$$G_{\text{obs}} = \frac{1}{\mathcal{N}} \sum_{a=1}^{\mathcal{N}} G_a,$$

where G_a is the value during the a th measurement whose time duration is very short—so short, in fact, that during the a th measurement the system can be considered to be in only one microscopic state. Then we can partition the sum as

$$G_{\text{obs}} = \sum_v \left[\frac{1}{\mathcal{N}} \left(\frac{\text{number of times state } v \text{ is observed in the } \mathcal{N} \text{ observations}}{\mathcal{N}} \right) \right] G_v,$$

where $G_v = \langle v | G | v \rangle$ is the expectation value for G when the system is in state v . The term in square brackets is the probability or weight for finding the system during the course of the measurements in state v . Remember, we believe that after a long enough time, all states are visited. We give the probability or fraction of time spent in state v the symbol P_v and write

$$G_{\text{obs}} = \sum_v P_v G_v \equiv \langle G \rangle.$$

The averaging operation (i.e., the weighted summation over G_v) indicated by the pointed brackets, $\langle G \rangle$, is called an *ensemble average*. An “ensemble” is the assembly of all possible microstates—all states consistent with the constraints with which we characterize the system macroscopically. For example, the *microcanonical ensemble* is the assembly of all states with fixed total energy E , and

fixed size (usually specified by number of molecules, N , and volume, V). The *canonical ensemble*, another example, considers all states with fixed size, but the energy can fluctuate. The former is appropriate to a closed isolated system; the latter is appropriate for a closed system in contact with a heat bath. There will be much more said about these ensembles later.

The idea that we observe the ensemble average, $\langle G \rangle$, arises from the view in which measurements are performed over a long time, and that due to the flow of the system through state space, the time average is the same as the ensemble average. The equivalence of a time average and an ensemble average, while sounding reasonable, is not at all trivial. Dynamical systems that obey this equivalence are said to be *ergodic*. It is difficult, in general, to establish the principle of ergodicity, though we believe it holds for all many-body systems encountered in nature. (It is often true for very small systems too, such as polyatomic molecules. Indeed, the basis of the standard theories of unimolecular kinetics rests on the assumed ergodic nature of intramolecular dynamics.)

Exercise 3.2 Give some examples of *non-ergodic* systems. That is, describe systems that do not sample all possible states even after a very long time.

Incidentally, suppose you thought of employing stationary solutions of Schrödinger's equation to specify microscopic states. If truly in a stationary state at some point in time, the system will remain there for all time, and the behavior cannot be ergodic. But in a many-body system, where the spacing between energy levels is so small as to form a continuum, there are always sources of perturbation or randomness (the walls of the container, for example) that make moot the chance of the system ever settling into a stationary state.

The primary assumption of statistical mechanics—that the observed value of a property corresponds to the ensemble average of that property—seems reasonable, therefore, if the observation is carried out over a very long time or if the observation is actually the average over very many independent observations. The two situations are actually the same if “long time” refers to a duration much longer than any *relaxation time* for the system. The idea that the system is chaotic at a molecular level leads to the concept that after some period of time—a relaxation time, τ_{relax} —the system will lose all memory of (i.e., correlation with) its initial conditions. Therefore, if a measurement is performed over a period τ_{measure} that is $\mathcal{N}\tau_{\text{relax}}$,

the measurement actually corresponds to \mathcal{N} independent observations.

In practice, we often consider measurements on macroscopic systems that are performed for rather short periods of time, and the concept of ensemble averages is applicable for these situations, too. This can be understood by imagining a division of the observed macroscopic system into an assembly of many macroscopic subsystems. If the subsystems are large enough, we expect that the precise molecular behavior in one subsystem is uncorrelated with that in any of the neighboring subsystems. The distance across one of these subsystems is then said to be much larger than the *correlation length* or *range of correlations*. When subsystems are this large they behave as if they are macroscopic. Under these conditions, one instantaneous measurement of the total macroscopic system is equivalent to many independent measurements of the macroscopic subsystems. The many independent measurements should correspond to an ensemble average.

3.2 Microcanonical Ensemble and the Rational Foundation of Thermodynamics

The basic idea of statistical mechanics is, therefore, that during a measurement, every microscopic state or fluctuation that is possible does in fact occur, and observed properties are actually the averages from all the microscopic states. To quantify this idea, we need to know something about the probability or distribution of the various microscopic states. This information is obtained from an assumption about the behavior of many-body systems:

For an isolated system with fixed total energy E , and fixed size (perhaps specified by the volume V and numbers of particles N_1, N_2, \dots) all microscopic states are equally likely at thermodynamic equilibrium.

In other words, the macroscopic equilibrium state corresponds to the most random situation—the distribution of microscopic states with the same energy and system size is entirely uniform.

Exercise 3.3 List several everyday examples supporting this statistical characterization of the terminal state of a macroscopic system (e.g., the behavior of a drop of ink in a glass of water).

To examine the implications of this reasonable assumption, let us define

$\Omega(N, V, E)$ = number of microscopic states with N and V ,
and energy between E and $E + \delta E$.

For notational and perhaps conceptual simplicity, we often omit subscripts and simply write N to refer to the number of particles, and we use the volume V to specify the spatial extent of the system. Our remarks, however, are not confined to one-component three-dimensional systems. The width δE is some energy interval characteristic of the limitation in our ability to specify absolutely precisely the energy of a macroscopic system. If δE was zero, the quantity $\Omega(N, V, E)$ would be a wildly varying discontinuous function, and when it was non-zero, its value would be the degeneracy of the energy level E . With a finite δE , $\Omega(N, V, E)$ is a relatively continuous function for which standard mathematical analysis is permissible. It will turn out that the thermodynamic consequences are extraordinarily insensitive to the size of δE . The reason for the insensitivity, we will see, is that $\Omega(N, V, E)$ is typically such a rapidly increasing function of E , that any choice of $\delta E \leq E$ will usually give the same answer for the thermodynamic consequences examined below. Due to this insensitivity, we adopt the shorthand where the symbol δE is not included in our formulas.

For macroscopic systems, energy levels will often be spaced so closely as to approach a continuum. In the continuum limit it can be convenient to adopt the notation

$\bar{\Omega}(N, V, E) dE$ = number of states with energy
between E and $E + dE$,

where $\bar{\Omega}(N, V, E)$, defined by this equation, is called the *density of states*. In the applications we pursue, however, we will have little need to employ this notation.

Exercise 3.4 For a system with discrete energy levels, give a formula for the density of states, $\bar{\Omega}(N, V, E)$.
[Hint: You will need to use the Dirac delta function.]

According to the statistical assumption, the probability of a macroscopic state v for an equilibrium system is

$$P_v = 1/\Omega(N, V, E)$$

for all states in the ensemble. For states outside the ensemble, for example those for which $E_v \neq E$, P_v is zero. This ensemble, which is

appropriate to a system with fixed energy, volume, and particle number—the assembly of all microstates with these constraints—is called a microcanonical ensemble.

We can also consider as a *definition of entropy* the quantity

$$S = k_B \ln \Omega(N, V, E),$$

where k_B is an arbitrary constant. (It's called *Boltzmann's constant* and we shall find that from comparison with experiment that it has the value

$$k_B = 1.380 \times 10^{-16} \text{ erg/deg.})$$

Notice that S defined in this way is extensive since if the total system were composed of two independent subsystems, A and B , with number of states Ω_A and Ω_B separately, then the total number would be $\Omega_A \Omega_B$. That is, $S_{A+B} = k_B \ln (\Omega_A \Omega_B) = S_A + S_B$.

The definition is also consistent with the variational statements of the second law of thermodynamics. To see why, imagine dividing the system with fixed total N , V , and E into two subsystems and constraining the partitioning of N , V , and E to be $N^{(1)}, N^{(2)}; V^{(1)}, V^{(2)}; E^{(1)}, E^{(2)}$, respectively. Any specific partitioning is a subset of all the allowed states, and therefore the number of states with this partitioning, $\Omega(N, V, E; \text{internal constraint})$ is less than the total number $\Omega(N, V, E)$. As a result,

$$S(N, V, E) > S(N, V, E; \text{internal constraint}).$$

This inequality is the second law, and we now see its statistical meaning: the maximization of entropy coinciding with the attainment of equilibrium corresponds to the maximization of disorder or molecular randomness. The greater the microscopic disorder, the larger the entropy.

The temperature T is determined from the derivative $(\partial S / \partial E)_{N,V} = 1/T$. Therefore,

$$\beta = (k_B T)^{-1} = (\partial \ln \Omega / \partial E)_{N,V}.$$

The thermodynamic condition that temperature is positive requires that $\Omega(N, V, E)$ be a monotonic increasing function of E . For macroscopic systems encountered in nature, this will always be the case.

Before accepting this fact as an obvious one, however, consider the following puzzle: Suppose a system of N non-interacting spins in a magnetic field H has the energy

$$- \sum_{j=1}^N \mu_j H, \quad \mu_j = \pm \mu.$$

In the ground state, all the spins are lined up with the field, and $\Omega = 1$. In a first excited state, one spin is flipped and $\Omega = N$. The next excitation has two spins flipped and $\Omega = N(N-1)/2$. Everything looks fine until we realize that the degeneracy of the most excited state is 1. Thus, at some point, $\Omega(E, N, V)$ becomes a decreasing function of E , which implies a negative temperature. How could this be?

Exercise 3.5* Answer this question.

Assuming $(\partial\Omega/\partial E)_{N,V,E}$ is positive, the statistical postulate that at fixed N, V , and E all microstates are equally likely provides a molecular foundation for the theory of thermodynamics. The many results derived during our discussion of that topic (concerning stability, phase equilibria, Maxwell relations, etc.) are all consequences of this single fundamental law of nature

3.3 Canonical Ensemble

When applying the microcanonical ensemble, the natural variables characterizing the macroscopic state of the system are E, V , and N . As we have seen in the context of thermodynamics, it is often convenient to employ other variables, and various representations of thermodynamics are found by applying Legendre transforms. In statistical mechanics, these manipulations are related to changes in ensembles. As an important example, we consider now the *canonical ensemble*—the assembly of all microstates with fixed N and V . The energy can fluctuate, however, and the system is kept at equilibrium by being in contact with a heat bath at temperature T (or inverse temperature β).

Schematically, we might picture the ensemble as in Fig. 3.2. The states we refer to here with the label ν are states of definite energy—eigenfunctions of Schrödinger's equation, $\mathcal{H}\psi_\nu = E_\nu\psi_\nu$.

A system for which the canonical ensemble is appropriate can be viewed as a subsystem of one for which the microcanonical is applicable. See Fig. 3.3. This observation allows us to derive the distribution law for states in the canonical ensemble.

To begin, consider the case where the bath is so large that the energy of the bath, E_B , is overwhelmingly larger than the energy of the system, E_ν . Further, the bath is so large that the energy levels of the bath are a continuum and $d\Omega/dE$ is well defined. The energy in the system fluctuates because the system is in contact with the bath, but

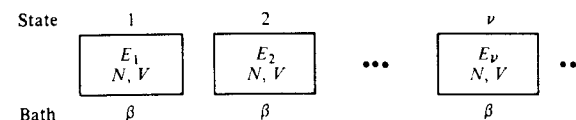


Fig. 3.2. Assembly of states for a closed system in a heat bath.

the sum $E = E_B + E_\nu$ is a constant. If the system is in one definite state ν , the number of states accessible to the system plus bath is $\Omega(E_B) = \Omega(E - E_\nu)$. Therefore, according to the statistical postulate—the principle of equal weights—the equilibrium probability for observing the system in state ν obeys

$$P_\nu \propto \Omega(E - E_\nu) = \exp [\ln \Omega(E - E_\nu)].$$

Since $E_\nu \ll E$, we can expand $\ln \Omega(E - E_\nu)$ in the Taylor series

$$\ln \Omega(E - E_\nu) = \ln \Omega(E) - E_\nu (d \ln \Omega / dE) + \dots$$

We choose to expand $\ln \Omega(E)$ itself because the latter is a much more rapidly varying function of E than the former. We believe this because the formula $S = k_B \ln \Omega$ suggests that $\ln \Omega$ is relatively well behaved.

By retaining only those terms exhibited explicitly in the expansion (which is valid because the bath is considered to be an infinite thermal reservoir), and noting $(\partial \ln \Omega / \partial E)_{N,V} = \beta$, we obtain

$$P_\nu \propto \exp(-\beta E_\nu),$$

which is the canonical (or Boltzmann) distribution law. The constant of proportionality is independent of the specific state of the system and is determined by the normalization requirement

$$\sum_\nu P_\nu = 1.$$

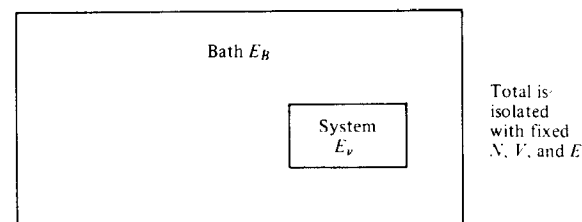


Fig. 3.3. A canonical ensemble system as a subsystem to microcanonical subsystem.

Hence,

$$P_v = Q^{-1} \exp(-\beta E_v),$$

where

$$Q(\beta, N, V) = \sum_v e^{-\beta E_v}.$$

The function $Q(\beta, N, V)$ is called the *canonical partition function*. It depends upon N and V through the dependence of the E_v 's on these variables.

As an instructive example of its use, consider the calculation of the internal energy $E(\beta, N, V)$, which is $\langle E \rangle$ in the canonical distribution:

$$\begin{aligned} \langle E \rangle &= \langle E_v \rangle = \sum_v P_v E_v \\ &= \sum_v E_v e^{-\beta E_v} / \sum_v e^{-\beta E_v} \\ &= -(\partial Q / \partial \beta)_{N,V} / Q \\ &= -(\partial \ln Q / \partial \beta)_{N,V}, \end{aligned}$$

which suggests that $\ln Q$ is a familiar thermodynamic function. In fact, we will soon show that $-\beta^{-1} \ln Q$ is the Helmholtz free energy. For the next few pages, however, let us take this fact as simply given.

Exercise 3.6 Show that $(\partial \beta A / \partial \beta)_{N,V} = E$, where $A = E - TS$ is the Helmholtz free energy.

The energies E_v refer to the eigen values of Schrödinger's equation for the system of interest. In general, these energies are difficult, if not impossible, to obtain. It is significant, therefore, that a canonical ensemble calculation can be carried out independent of the exact solutions to Schrödinger's equation. This fact is understood as follows:

$$\begin{aligned} Q &= \sum_v e^{-\beta E_v} = \sum_v \langle v | e^{-\beta \mathcal{H}} | v \rangle \\ &= \text{Tr } e^{-\beta \mathcal{H}}, \end{aligned}$$

where "Tr" denotes the trace of a matrix (in this case, the trace of the Boltzmann operator matrix). It is a remarkable property of traces that they are independent of the representation of a matrix. (Proof: $\text{Tr } A = \text{Tr } S S^{-1} A = \text{Tr } S^{-1} A S$.) Thus, once we know \mathcal{H} we can use any complete set of wavefunctions to compute Q . In other words, one may calculate $Q = \exp(-\beta A)$ without actually solving Schrödinger's equation with the Hamiltonian \mathcal{H} .

Exercise 3.7 Show that the internal energy, which is the average of E_v , can be expressed as $\text{Tr}(\mathcal{H} e^{-\beta \mathcal{H}}) / \text{Tr}(e^{-\beta \mathcal{H}})$.

When calculating properties like the internal energy from the canonical ensemble, we expect that values so obtained should be the same as those found from the microcanonical ensemble. Indeed, as the derivation given above indicates, the two ensembles will be equivalent when the system is large. This point can be illustrated in two ways. First, imagine partitioning the sum over states in Q into groups of states with the same energy levels, that is

$$\begin{aligned} Q &= \sum_{v(\text{states})} e^{-\beta E_v} \\ &= \sum_{l(\text{levels})} \Omega(E_l) e^{-\beta E_l}, \end{aligned}$$

where we have noted that the number of states, $\Omega(E_l)$, is the degeneracy of the l th energy level. For a very large system, the spacing between levels is very small, and it seems most natural to pass to the continuum limit

$$Q \rightarrow \int_0^\infty dE \bar{\Omega}(E) e^{-\beta E},$$

where $\bar{\Omega}(E)$ is the density of states. In other words, for large systems, the canonical partition function is the *Laplace transform* of the microcanonical $\Omega(E)$. An important theorem of mathematics is that Laplace transforms are unique. Due to this uniqueness, the two functions contain the identical information.

Nevertheless, energy fluctuates in the canonical ensemble while energy is fixed in the microcanonical ensemble. This inherent difference between the two does not contradict the equivalency of ensembles, however, because the relative size of the fluctuations becomes vanishingly small in the limit of large systems. To see why, let us compute the averaged square fluctuation in the canonical ensemble:

$$\begin{aligned} \langle (\delta E)^2 \rangle &= \langle (E - \langle E \rangle)^2 \rangle \\ &= \langle E^2 \rangle - \langle E \rangle^2 \\ &= \sum_v P_v E_v^2 - \left(\sum_v P_v E_v \right)^2 \\ &= Q^{-1} (\partial^2 Q / \partial \beta^2)_{N,V} - Q^{-2} (\partial Q / \partial \beta)_{N,V}^2 \\ &= (\partial^2 \ln Q / \partial \beta^2)_{N,V} \\ &= -(\partial \langle E \rangle / \partial \beta)_{N,V}. \end{aligned}$$

Noting the definition of heat capacity, $C_v = (\partial E / \partial T)_{N, V}$, we have

$$\langle (\delta E)^2 \rangle = k_B T^2 C_v,$$

which is a remarkable result in its own right since it relates the size of spontaneous fluctuations, $\langle (\delta E)^2 \rangle$, to the rate at which energy will change due to alterations in the temperature. (The result foreshadows the topics of linear response theory and the fluctuation-dissipation theorem, which we will discuss in Chapter 8.) In the present context, we use the fluctuation formula to estimate the relative r.m.s. value of the fluctuations. Since the heat capacity is extensive, it is of order N (where N is the number of particles in the system). Furthermore $\langle E \rangle$ is also of order N . Hence the ratio of the dispersion to the average value is of order $N^{-1/2}$; that is,

$$\frac{\sqrt{\langle [E - \langle E \rangle]^2 \rangle}}{\langle E \rangle} = \frac{\sqrt{k_B T^2 C_v}}{\langle E \rangle} \sim 0 \left(\frac{1}{\sqrt{N}} \right).$$

For a large system ($N \sim 10^{23}$) this is a very small number and we may thus regard the average value, $\langle E \rangle$, as a meaningful prediction of the experimental internal energy. (For an ideal gas of structureless particles, $C_v = \frac{3}{2} N k_B$, $\langle E \rangle = \frac{3}{2} N k_B T$. Suppose $N \sim 10^{22}$, then the ratio above is numerically $\sim 10^{-11}$.) Furthermore, the microcanonical E , when written as a function of β , N , V by inverting $(\partial \ln \Omega / \partial E)_{N, V} = \beta(E, N, V)$, will be indistinguishable from the canonical internal energy $\langle E \rangle$ provided the system is large.

Exercise 3.8 Note that the probability for observing a closed thermally equilibrated system with a given energy E is $P(E) \propto \Omega(E) e^{-\beta E} = \exp [\ln \Omega(E) - \beta E]$. Both $\ln \Omega(E)$ and $-\beta E$ are of the order of N , which suggests that $P(E)$ is a very narrow distribution centered on the most probable value of E . Verify this suggestion by performing a steepest descent calculation with $P(E)$. That is, expand $\ln P(E)$ in powers of $\delta E = E - \langle E \rangle$, and truncate the expansion after the quadratic term. Use this expansion to estimate for 0.001 moles of gas the probability for observing a spontaneous fluctuation in E of the size of $10^{-6} \langle E \rangle$.

3.4 A Simple Example

To illustrate the theory we have been describing, consider a system of N distinguishable independent particles each of which can exist in

one of two states separated by an energy ϵ . We can specify the state of a system, v , by listing

$$v = (n_1, n_2, \dots, n_j, \dots, n_N), \quad n_j = 0 \text{ or } 1,$$

where n_j gives the state of particle j . The energy of the system for a given state is

$$E_v = \sum_{j=1}^N n_j \epsilon,$$

where we have chosen the ground state as the zero of energy.

To calculate the thermodynamic properties of this model, we first apply the microcanonical ensemble. The degeneracy of the m th energy level is the number of ways one may choose m objects from a total of N . That is,

$$\Omega(E, N) = N! / (N - m)! m!,$$

where

$$m = E / \epsilon.$$

The entropy and temperature are given by

$$S / k_B = \ln \Omega(E, N)$$

and

$$\begin{aligned} \beta &= 1 / k_B T = (\partial \ln \Omega / \partial E)_N \\ &= \epsilon^{-1} (\partial \ln \Omega / \partial m)_N. \end{aligned}$$

For the last equality to have meaning, N must be large enough that $\Omega(E, N)$ will depend upon m in a continuous way. The continuum limit of factorials is *Stirling's approximation*: $\ln M! \approx M \ln M - M$, which becomes exact in the limit of large M . With that approximation

$$\begin{aligned} \frac{\partial}{\partial m} \ln \frac{N!}{(N - m)! m!} &= - \frac{\partial}{\partial m} [(N - m) \ln (N - m) \\ &\quad - (N - m) + m \ln m - m] \\ &= \ln \left(\frac{N}{m} - 1 \right). \end{aligned}$$

Combining this result with the formula for β yields

$$\beta \epsilon = \ln \left(\frac{N}{m} - 1 \right)$$

or

$$\frac{m}{N} = \frac{1}{1 + e^{\beta \epsilon}}.$$

As a result, the energy $E = m\epsilon$ as a function of temperature is

$$E = N\epsilon \frac{1}{1 + e^{\beta\epsilon}},$$

which is 0 at $T = 0$ (i.e., only the ground state is populated), and it is $N\epsilon/2$ as $T \rightarrow \infty$ (i.e., all states are then equally likely).

Exercise 3.9 Use Stirling's approximation together with the formula for m/N to derive an expression for $S(\beta, N)$. Show that as $\beta \rightarrow \infty$ (i.e., $T \rightarrow 0$), S tends to zero. Find $S(E, N)$ and examine the behavior of $1/T$ as a function of E/N . Show that for some values of E/N , $1/T$ can be negative.

Of course, we could also study this model system with the canonical ensemble. In that case, the link to thermodynamics is

$$-\beta A = \ln Q = \ln \sum_v e^{-\beta E_v}.$$

Use of the formula for E_v gives

$$Q(\beta, N) = \sum_{n_1, n_2, \dots, n_N=0,1} \exp \left[-\beta \sum_{j=1}^N \epsilon n_j \right]$$

since the exponential factors into an uncoupled product,

$$\begin{aligned} Q(\beta, N) &= \prod_{j=1}^N \sum_{n_j=0,1} e^{-\beta \epsilon n_j} \\ &= (1 + e^{-\beta \epsilon})^N. \end{aligned}$$

As a result

$$-\beta A = N \ln (1 + e^{-\beta \epsilon}).$$

The internal energy is

$$\begin{aligned} \langle E \rangle &= \left(\frac{\partial(-\beta A)}{\partial(-\beta)} \right)_N = N \frac{\epsilon e^{-\beta \epsilon}}{1 + e^{-\beta \epsilon}} \\ &= N\epsilon (1 + e^{\beta \epsilon})^{-1}, \end{aligned}$$

in precise agreement with the result obtained with the microcanonical ensemble.

Exercise 3.10 Determine the entropy using

$$-\beta(A - \langle E \rangle) = S/k_B$$

and show the result is the same as that obtained for large N from the microcanonical ensemble.

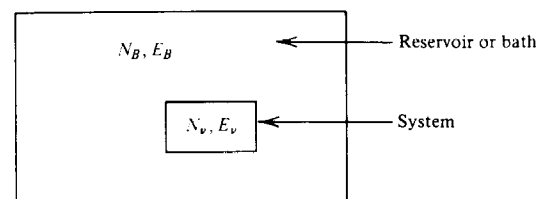


Fig. 3.4. System immersed in a bath.

3.5 Generalized Ensembles and the Gibbs Entropy Formula

Let us consider now, in a rather general way, why changes in ensembles correspond thermodynamically to performing Legendre transforms of the entropy. To begin, consider a system with X denoting the mechanical extensive variables. That is, $S = k_B \ln \Omega(E, X)$, and

$$k_B^{-1} dS = \beta dE + \xi dX.$$

For example, if $X = N$, then $\xi = -\beta\mu$. Or if X was the set of variables V, N_1, N_2, \dots , then ξ would correspond to the conjugate set $\beta p, -\beta\mu_1, -\beta\mu_2, \dots$, respectively. The quantity $-\xi/\beta$ therefore corresponds to f of Chapters 1 and 2.

Imagine an equilibrated system in which E and X can fluctuate. It can be viewed as a part of an isolated composite system in which the other part is a huge reservoir for E and X . An example could be an open system in contact with a bath with particles and energy flowing between the two. This example is pictured in Fig. 3.4.

The probability for microstates in the system can be derived in the same way we established the canonical distribution law. The result is

$$P_v = \exp(-\beta E_v - \xi X_v) / \Xi,$$

with

$$\Xi = \sum_v \exp(-\beta E_v - \xi X_v).$$

Exercise 3.11 Verify this result.

The thermodynamic E and X are given by the averages

$$\langle E \rangle = \sum_v P_v E_v = \left[\frac{\partial \ln \Xi}{\partial(-\beta)} \right]_{\xi, X}$$

and

$$\langle X \rangle = \sum_{\nu} P_{\nu} X_{\nu} = \left[\frac{\partial \ln \Xi}{\partial (-\xi)} \right]_{\beta, Y},$$

where Y refers to all the extensive variables that are not fluctuating in the system. In view of the derivative relationships,

$$d \ln \Xi = -\langle E \rangle d\beta - \langle X \rangle d\xi.$$

Now consider the quantity

$$\mathcal{S} = -k_B \sum_{\nu} P_{\nu} \ln P_{\nu}.$$

We have

$$\begin{aligned} \mathcal{S} &= -k_B \sum_{\nu} P_{\nu} [-\ln \Xi - \beta E_{\nu} - \xi X_{\nu}] \\ &= k_B \{ \ln \Xi + \beta \langle E \rangle + \xi \langle X \rangle \}. \end{aligned}$$

Therefore, \mathcal{S}/k_B is the Legendre transform that converts $\ln \Xi$ to a function of $\langle E \rangle$ and $\langle X \rangle$; that is,

$$d\mathcal{S} = \beta k_B d\langle E \rangle + \xi k_B d\langle X \rangle,$$

which implies the \mathcal{S} is, in fact, the entropy S . Thus in general

$$S = -k_B \sum_{\nu} P_{\nu} \ln P_{\nu}.$$

This result for the entropy is a famous one. It is called the Gibbs entropy formula.

Exercise 3.12 Verify that the microcanonical $S = k_B \ln \Omega(N, V, E)$ is consistent with the Gibbs formula.

The most important example of these formulas is that of the *grand canonical ensemble*. This ensemble is the assembly of all states appropriate to an *open* system of volume V . Both energy and particle number can fluctuate from state to state, and the conjugate fields that control the size of these fluctuations are β and $-\beta\mu$, respectively. Thus, letting ν denote the state with N_{ν} particles and energy E_{ν} , we have

$$P_{\nu} = \Xi^{-1} \exp(-\beta E_{\nu} + \beta\mu N_{\nu}),$$

and the Gibbs entropy formula yields

$$\begin{aligned} S &= -k_B \sum_{\nu} P_{\nu} [-\ln \Xi - \beta E_{\nu} + \beta\mu N_{\nu}] \\ &= -k_B [-\ln \Xi - \beta \langle E \rangle + \beta\mu \langle N \rangle], \end{aligned}$$

or, on rearranging terms

$$\ln \Xi = \beta p V,$$

where p is the thermodynamic pressure. Note that

$$\Xi = \sum_{\nu} \exp(-\beta E_{\nu} + \beta\mu N_{\nu})$$

is a function of β , $\beta\mu$, and the volume. (It depends upon volume because the energies E_{ν} depend upon the size of the system.) Hence, the “free energy” for an open system, $\beta p V$, is a natural function of β , $\beta\mu$, and V .

Fluctuation formulas in the grand canonical ensemble are analyzed in the same fashion as in the canonical ensemble. For example,

$$\begin{aligned} \langle (\delta N)^2 \rangle &= \langle (N - \langle N \rangle)^2 \rangle = \langle N^2 \rangle - \langle N \rangle^2 \\ &= \sum_{\nu} N_{\nu}^2 P_{\nu} - \sum_{\nu} \sum_{\nu'} N_{\nu} N_{\nu'} P_{\nu} P_{\nu'} \\ &= [\partial^2 \ln \Xi / \partial (\beta\mu)^2]_{\beta, V}, \end{aligned}$$

or

$$\langle (\partial N)^2 \rangle = (\partial \langle N \rangle / \partial \beta\mu)_{\beta, V}.$$

Generalizations to multicomponent systems can also be worked out in the same way, and they are left for the Exercises.

Recall that in our study of thermodynamic stability (i.e., the convexity of free energies) we found that $(\partial n / \partial \mu) \geq 0$. Now we see the same result in a different context. In particular, note that $\langle N \rangle = n N_0$, where N_0 is Avogadro's number, and since $\delta N = N - \langle N \rangle$ is a real number its square is positive. Hence, $\partial \langle N \rangle / \partial \beta\mu = \langle (\delta N)^2 \rangle \geq 0$. Similarly, in Chapter 2, we found from thermodynamic stability that $C_v \geq 0$, and in this chapter we learn that $k_B T^2 C_v = \langle (\delta E)^2 \rangle \geq 0$. In general, statistical mechanics will always give

$$-(\partial \langle X \rangle / \partial \xi) = \langle (\delta X)^2 \rangle.$$

The right-hand side is manifestly positive, and the left-hand side determines the curvature or convexity of a thermodynamic free energy.

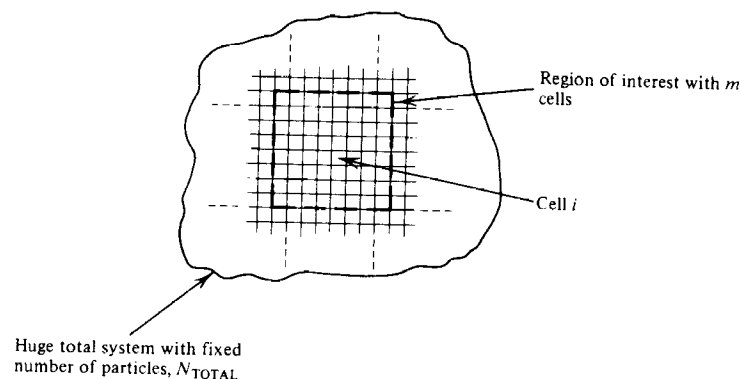


Fig. 3.5. Partitioning into cells.

3.6 Fluctuations Involving Uncorrelated Particles

In this section we will illustrate how the nature of spontaneous microscopic fluctuations governs the macroscopic observable behavior of a system. In the illustration, we consider concentration or density fluctuations in a system of uncorrelated particles, and we show that the ideal gas law (i.e., $pV = nRT$) follows from the assumption of no interparticle correlations. We will return to the ideal gas in Chapter 4 where we will derive its thermodynamic properties from detailed considerations of its energy levels. The following analysis, however, is of interest due to its generality being applicable even to large polymers at low dilution in a solvent.

To begin we imagine partitioning the volume of a system with cells as pictured in Fig. 3.5. Fluctuations in the region of interest follow the grand canonical distribution law described in Sec. 3.5. We will assume that the cells are constructed to be so small that there is a negligible chance for more than one particle to be in the same cell at the same time. Therefore, we can characterize any statistically likely configurations by listing the numbers (n_1, n_2, \dots, n_m) , where

$$\begin{aligned} n_i &= 1, \text{ if a particle is in cell } i \\ &= 0, \text{ otherwise.} \end{aligned}$$

In terms of these numbers, the instantaneous total number of particles in the region of interest is

$$N = \sum_{i=1}^m n_i,$$

and the mean square fluctuation in this number is

$$\begin{aligned} \langle (\delta N)^2 \rangle &= \langle [N - \langle N \rangle]^2 \rangle = \langle N^2 \rangle - \langle N \rangle^2 \\ &= \sum_{i,j=1}^m [\langle n_i n_j \rangle - \langle n_i \rangle \langle n_j \rangle]. \end{aligned} \quad (a)$$

These relations are completely general. A simplification is found by considering the case in which different particles are uncorrelated with each other and this lack of correlation is due to a very low concentration of particles. These two physical considerations imply

$$\langle n_i n_j \rangle = \langle n_i \rangle \langle n_j \rangle \quad \text{for } i \neq j \quad (b)$$

(see Exercise 3.17), and

$$\langle n_i \rangle \ll 1, \quad (c)$$

respectively. Further, since n_i is either zero or one, $n_i^2 = n_i$ and hence

$$\langle n_i^2 \rangle = \langle n_i \rangle = \langle n_1 \rangle, \quad (d)$$

where the last equality follows from the assumption each cell is of the same size or type. Hence, on the average, each cell behaves identically to every other cell.

The insertion of (b) into (a) yields

$$\langle (\delta N)^2 \rangle = \sum_{i=1}^m [\langle n_i^2 \rangle - \langle n_i \rangle^2],$$

and the application of (d) gives

$$\langle (\delta N)^2 \rangle = m \langle n_1 \rangle (1 - \langle n_1 \rangle).$$

Finally, from (c) we arrive at

$$\langle (\delta N)^2 \rangle \approx m \langle n_1 \rangle = \langle N \rangle.$$

By itself, this relationship is already a remarkable result, but its thermodynamic ramification is even more impressive.

In particular, since the region of interest is described by the grand canonical ensemble, we know that (see Sec. 3.5 and Exercise 3.15)

$$\langle (\delta N)^2 \rangle = (\partial \langle N \rangle / \partial \beta \mu)_{\beta, V}.$$

Hence, for a system of uncorrelated particles, we have

$$(\partial \langle N \rangle / \partial \beta \mu)_{\beta, V} = \langle N \rangle,$$

or dividing by V and taking the reciprocal

$$(\partial \beta \mu / \partial \rho)_{\beta} = \rho^{-1},$$

where $\rho = \langle N \rangle / V$. Thus, by integration we have

$$\beta\mu = \text{constant} + \ln \rho.$$

Further, from standard manipulations (see Exercise 1.14)

$$(\partial\beta p / \partial\rho)_\beta = \rho(\partial\beta\mu / \partial\rho)_\beta = 1,$$

where the first equality is a general thermodynamic relation and the second applies what we have discovered for uncorrelated particles. Integration yields

$$\beta p = \rho,$$

where we have set the constant of integration to zero since the pressure should vanish as the density ρ goes to zero. This equation is the celebrated ideal gas law, $pV = nRT$, where we identify the gas constant, R , with Boltzmann's constant times Avogadro's number, N_0 :

$$R = k_B N_0.$$

In summary, we have shown that the assumption of uncorrelated statistical behavior implies that for a one-component system

$$\rho \propto e^{\beta\mu}$$

and

$$\beta p / \rho = 1.$$

Generalizations to multicomponent systems are straightforward and left for Exercises.

3.7 Alternative Development of Equilibrium Distribution Functions

The approach we have followed thus far begins with a statistical characterization of equilibrium states and then arrives at the inequalities and distribution laws we regard as the foundation of thermodynamics. Alternatively, we could begin with the second law and the Gibbs entropy formula rather than deducing them from the principle of equal weights. In the next few pages we follow this alternative development.

Extensivity of Entropy

Since the Gibbs entropy formula is our starting point, let's check that it satisfies the additivity property (extensivity) that we associate with



Fig. 3.6. Two independent subsystems A and B.

entropy. Consider a system contained in two boxes, A and B (see Fig. 3.6). Denote the total entropy of the system by S_{AB} . If the entropy is extensive, $S_{AB} = S_A + S_B$. From the Gibbs formula

$$S_{AB} = -k_B \sum_{v_A} \sum_{v_B} P_{AB}(v_A, v_B) \ln P_{AB}(v_A, v_B),$$

where v_A and v_B denote states of the subsystems A and B, respectively. Since the subsystems are uncoupled,

$$P_{AB}(v_A, v_B) = P_A(v_A)P_B(v_B).$$

Thus

$$\begin{aligned} S_{AB} &= -k_B \sum_{v_A} \sum_{v_B} P_{AB}(v_A, v_B) [\ln P_{AB}(v_A, v_B)] \\ &= -k_B \sum_{v_B} P_B(v_B) \sum_{v_A} P_A(v_A) \ln P_A(v_A) \\ &\quad - k_A \sum_{v_A} P_A(v_A) \sum_{v_B} P_B(v_B) \ln P_B(v_B) \\ &= -k_B \sum_{v_A} P_A(v_A) \ln P_A(v_A) - k_B \sum_{v_B} P_B(v_B) \ln P_B(v_B) \\ &= S_A + S_B, \end{aligned}$$

where the second to last equality is obtained from the normalization condition. This simple calculation shows that the Gibbs entropy exhibits the thermodynamic property $S_{AB} = S_A + S_B$.

Exercise 3.13 Show that if one assumes the functional form

$$S = \sum_v P_v f(P_v),$$

where $f(x)$ is some function of x , then the requirement that S is extensive implies that $f(x) = c \ln x$, where c is an arbitrary constant.

The Microcanonical Ensemble

For an isolated system, the energy, E , the number of particles, N , and the volume, V , are fixed. The ensemble appropriate to such a system is the *microcanonical ensemble*: the assembly of all states with E , N , and V fixed.

To derive the equilibrium probability for state j , P_j , we require that the condition for thermodynamic equilibrium be satisfied. According to the second law,

$$(\delta S)_{E,V,N} = 0.$$

In other words, the partitioning of microscopic states at equilibrium is the partitioning that maximizes the entropy. We use this principle and carry out a maximization procedure with the constraints

$$\langle E \rangle = \sum_j E_j P_j, \quad (a)$$

$$\langle N \rangle = \sum_j N_j P_j, \quad (b)$$

and

$$1 = \sum_j P_j. \quad (c)$$

In the microcanonical ensemble where $E_j = E = \text{constant}$, and $N_j = N = \text{constant}$, conditions (a), (b), and (c) are all the same.

Using the Lagrange multiplier γ , we seek a P_j for which

$$\delta(S + \gamma 1) = 0,$$

or, inserting Eq. (c) and the Gibbs entropy formula,

$$\begin{aligned} 0 &= \delta \left\{ -k_B \sum_j P_j \ln P_j + \gamma \sum_j P_j \right\} \\ &= \sum_j \delta P_j [-k_B \ln P_j - k_B + \gamma]. \end{aligned}$$

For this equation to be valid for all δP_j , the quantity in [] must equal zero. Thus

$$\ln P_j = \frac{\gamma - k_B}{k_B} = \text{constant}.$$

The constant can be determined from the normalization condition

$$1 = \sum_j P_j = \sum_j e^{\text{constant}} = \sum_j \frac{1}{\Omega} = \frac{1}{\Omega} \left(\sum_j 1 \right).$$

Thus

$$\Omega = \text{the number of states with energy } E.$$

In summary, for the microcanonical ensemble

$$\begin{aligned} P_j &= \frac{1}{\Omega}, \quad \text{for } E_j = E \\ &= 0, \quad \text{for } E_j \neq E \end{aligned}$$

and the entropy is

$$S = +k_B \sum_j \frac{1}{\Omega} \ln \Omega = k_B \ln \sum_j \frac{1}{\Omega} = k_B \ln \Omega.$$

The Canonical Ensemble

This ensemble is appropriate to a closed system in a temperature bath. N , V , and T are fixed, but the energy is not. Thermodynamic equilibrium now gives

$$\delta(S + \alpha \langle E \rangle + \gamma 1) = 0,$$

where α and γ are Lagrange multipliers. By combining Eqs. (a), (c), and the Gibbs entropy formula with the above, we obtain

$$\sum_j [-k_B \ln P_j - k_B + \alpha E_j + \gamma] \delta P_j = 0.$$

For this expression to hold for all δP_j ,

$$[-k_B \ln P_j - k_B + \alpha E_j + \gamma] = 0$$

or

$$\ln P_j = \frac{\alpha E_j - k_B + \gamma}{k_B}. \quad (d)$$

To determine α and γ , we use the thermodynamic identity

$$\left[\frac{\delta \langle E \rangle}{\delta S} \right]_{V,N} = T = \text{temperature}.$$

With Eq. (a) we find

$$(\delta \langle E \rangle)_{V,N} = \sum_j E_j \delta P_j,$$

and from the Gibbs entropy formula and (d) we find

$$\begin{aligned} (\delta S)_{V,N} &= -k_B \sum_j \delta P_j \left[\frac{\alpha E_j - k_B + \gamma}{k_B} \right] \\ &= -k_B \sum_j \delta P_j E_j \alpha / k_B, \end{aligned}$$

where the last equality follows from the fact that $\sum_j \delta P_j = \delta 1 = 0$. Note that in the variation of $\langle E \rangle$, we do not alter E_j since the variation refers to changes in P_j (i.e., partitioning of states) with the energies of the states fixed. Dividing $(\delta \langle E \rangle)_{V,N}$ by $(\delta S)_{V,N}$ yields

$$T = \left[\frac{\delta \langle E \rangle}{\delta S} \right]_{V,N} = -\frac{1}{\alpha}.$$

Combining this result with (d) and the Gibbs entropy formula gives

$$\begin{aligned} S &= \sum_j P_j \left[\frac{E_j + k_B T - \gamma T}{T} \right] \\ &= \frac{\langle E \rangle + k_B T - \gamma T}{T}. \end{aligned}$$

Thus

$$\gamma T = A + k_B T,$$

where

$$A = \langle E \rangle - TS = \text{Helmholtz free energy.}$$

In summary, the canonical ensemble has

$$P_j = e^{-\beta(E_j - A)},$$

where

$$\beta = \frac{1}{k_B T}.$$

Since P_j is normalized,

$$\sum_j P_j = 1 = e^{\beta A} \sum_j e^{-\beta E_j}.$$

Thus, the *partition function*, Q ,

$$Q = \sum_j e^{-\beta E_j}$$

is also given by

$$Q = e^{-\beta A}.$$

From thermodynamic considerations alone, it is clear that the knowledge of Q tells us everything about the thermodynamics of our system. For example,

$$\left[\frac{\partial \ln Q}{\partial V} \right]_{T,N} = \left[\frac{\partial(-\beta A)}{\partial V} \right]_{T,N} = \beta p,$$

where p is the pressure, and

$$\left[\frac{\partial \ln Q}{\partial \beta} \right]_{V,N} = \left[\frac{\partial(-\beta A)}{\partial \beta} \right]_{V,N} = -\langle E \rangle,$$

where $\langle E \rangle$ is the internal energy.

Similar analysis can be applied to other ensembles, too. In general, therefore, the principle of equal weights is equivalent to the Gibbs entropy formula and the variational statement of the second law of thermodynamics.

Additional Exercises

3.14. By applying Gibbs entropy formula and the equilibrium condition

$$(\delta S)_{\langle E \rangle, V, \langle N \rangle} = 0,$$

derive the probability distribution for the grand canonical ensemble—the ensemble in which N and E can vary. Your result should be

$$P_v = \Xi^{-1} \exp[-\beta E_v + \beta \mu N_v],$$

where v labels that state of the system (including the number of particles) and

$$\Xi = \exp(\beta p V).$$

3.15. For an open multicomponent system, show that

$$\langle \delta N_i \delta N_j \rangle = (\partial \langle N_i \rangle / \partial \beta \mu_j)_{\beta, \beta \mu_i, V},$$

where $\delta N_i = N_i - \langle N_i \rangle$ is the fluctuation from the average of the number of particles of type i , and μ_i is the chemical potential for that type of particle. Similarly, relate $\langle \delta N_i \delta N_i \delta N_j \rangle$ to a thermodynamic derivative. Finally, for a one-component system in the grand canonical ensemble, evaluate $\langle (\delta E)^2 \rangle$ and relate this quantity to the constant volume heat capacity and the compressibility. The former determines the size of the mean square energy fluctuations in the canonical ensemble where density does not fluctuate, and the latter determines the size of the mean square density fluctuations.

3.16. For 0.01 moles of ideal gas in an open thermally equilibrated system, evaluate numerically the relative root mean square deviation of the energy from its mean value and the relative root mean square deviation of the density from its mean value.

- 3.17. (a) Consider a random variable x that can take on any value in the interval $a \leq x \leq b$. Let $g(x)$ and $f(x)$ be any functions of x and let $\langle \cdots \rangle$ denote the average over the distribution for x , $p(x)$ —that is,

$$\langle g \rangle = \int_a^b dx g(x)p(x).$$

Show that

$$\langle gf \rangle = \langle g \rangle \langle f \rangle$$

for arbitrary $g(x)$ and $f(x)$ if and only if

$$p(x) = \delta(x - x_0),$$

where x_0 is a point between a and b , and $\delta(x - x_0)$ is the Dirac delta function,

$$\delta(y) = 0, \quad y \neq 0$$

and

$$\int_{-\epsilon}^{\epsilon} dy \delta(y) = 1.$$

Note that according to this definition, $\delta(x - x_0)$ is a normalized distribution of zero (or infinitesimal) width located at $x = x_0$.

- (b) Consider two random variables x and y with the joint probability distribution $p(x, y)$. Prove that

$$\langle f(x)g(y) \rangle = \langle f \rangle \langle g \rangle$$

for all functions $f(x)$ and $g(y)$, if and only if

$$p(x, y) = p_1(x)p_2(y),$$

where $p_1(x)$ and $p_2(y)$ are the distributions for x and y , respectively.

- 3.18. Consider a system of N distinguishable non-interacting spins in a magnetic field H . Each spin has a magnetic moment of size μ , and each can point either parallel or antiparallel to the field. Thus, the energy of a particular state is

$$\sum_{i=1}^N -n_i \mu H, \quad n_i = \pm 1,$$

where $n_i \mu$ is the magnetic moment in the direction of the field.

- (a) Determine the internal energy of this system as a function

of β , H , and N by employing an ensemble characterized by these variables.

- (b) Determine the entropy of this system as a function of β , H , and N .
 (c) Determine the behavior of the energy and entropy for this system as $T \rightarrow 0$.

- 3.19. (a) For the system described in Exercise 3.18, derive the average total magnetization,

$$\langle M \rangle = \left\langle \sum_{i=1}^N \mu n_i \right\rangle,$$

as a function of β , H , and N .

- (b) Similarly, determine $\langle (\delta M)^2 \rangle$, where

$$\delta M = M - \langle M \rangle,$$

and compare your result with the susceptibility

$$(\partial \langle M \rangle / \partial H)_{\beta, N},$$

- (c) Derive the behavior of $\langle M \rangle$ and $\langle (\delta M)^2 \rangle$ in the limit $T \rightarrow 0$.

- 3.20. Consider the system studied in Exercises 3.18 and 3.19. Use an ensemble in which the total magnetization is fixed, and determine the magnetic field over temperature, βH , as a function of the natural variables for that ensemble. Show that in the limit of large N , the result obtained in this way is equivalent to that obtained in Exercise 3.19.

- 3.21.* In this problem you consider the behavior of mixed valence compounds solvated at very low concentrations in a crystal. Figure 3.7 is a schematic depiction of such a compound. We shall assume the compound has only two configurational states as illustrated in Fig. 3.8. The two states correspond to having the electron localized on the left or right iron atoms, respectively. This type of two-state model is like the LCAO

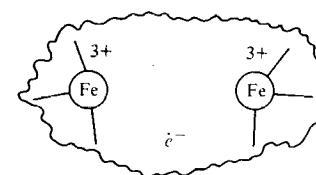


Fig. 3.7. A mixed valence compound conceived of as two cations plus an electron.

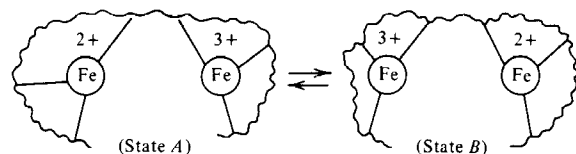


Fig. 3.8. Two-state model of a mixed valence compound.

treatment of the H_2^+ molecule in elementary quantum chemistry. In the solid state physics literature, the model is called the “tight binding” approximation.

In the absence of the surrounding crystal, the Hamiltonian for a compound is \mathcal{H}_0 with matrix elements

$$\langle A | \mathcal{H}_0 | A \rangle = \langle B | \mathcal{H}_0 | B \rangle = 0 \quad (\text{our choice for the zero of energy}),$$

$$\langle A | \mathcal{H}_0 | B \rangle = -\Delta.$$

The dipole moment of one of the compounds for states A or B is given by

$$\mu = \langle A | m | A \rangle = -\langle B | m | B \rangle,$$

where m denotes the electronic dipole operator. For further simplicity, imagine that there is negligible spatial overlap between states A and B ; that is,

$$\langle A | B \rangle = 0 \quad \text{and} \quad \langle A | m | B \rangle = 0.$$

The solvent crystal couples to the impurity mixed valence compounds through the electric crystal field, \mathcal{E} . The Hamiltonian for each compound is

$$\mathcal{H} = \mathcal{H}_0 - m\mathcal{E}.$$

- (a) Show that when $\mathcal{E} = 0$, the eigenstates of the single compound Hamiltonian are

$$|\pm\rangle = \frac{1}{\sqrt{2}}[|A\rangle \pm |B\rangle],$$

and the energy levels are $\pm\Delta$.

- (b) Compute the canonical partition function for the system of mixed valence compounds when $\mathcal{E} = 0$, by (i) performing the Boltzmann weighted sum with energy eigenvalues, and (ii) performing the matrix trace of $e^{-\beta\mathcal{H}_0}$ employing the configurational states $|A\rangle$ and $|B\rangle$. The latter states

diagonalize m but not \mathcal{H}_0 . Nevertheless, the two computations should yield the same result. Why?

- (c) When \mathcal{E} is zero, determine the averages (i) $\langle m \rangle$, (ii) $\langle |m| \rangle$, and (iii) $\langle (\delta m)^2 \rangle$, where $\delta m = m - \langle m \rangle$.
- (d) When $\mathcal{E} \neq 0$, the crystal couples to the impurity compounds, and there is a free energy of solvation, $[A(\mathcal{E}) - A(0)]/N$, where N is the number of compounds. Compute this free energy of solvation by (i) first determining the eigen energies as a function of \mathcal{E} , then performing the appropriate Boltzmann weighted sum, and (ii) performing the appropriate matrix trace employing configuration states $|A\rangle$ and $|B\rangle$. The two calculations yield the same result, though the second is algebraically more tedious. (You might find it useful to organize the algebra in the second case by exploring the properties of Pauli spin matrices.)
- (e) When $\mathcal{E} \neq 0$, compute $\langle |m| \rangle$ and $\langle m \rangle$. Compare its value with what is found when $\mathcal{E} = 0$. Why does $\langle m \rangle$ increase with increasing \mathcal{E} ?

- 3.22. (a) Consider a region within a fluid described by the van der Waals equation $\beta p = \rho/(1 - b\rho) - \beta a\rho^2$, where $\rho = \langle N \rangle/V$. The volume of the region is L^3 . Due to the spontaneous fluctuations in the system, the instantaneous value of the density in that region can differ from its average by an amount $\delta\rho$. Determine, as a function of β , ρ , a , b , and L^3 , the typical relative size of these fluctuations; that is, evaluate $\langle (\delta\rho)^2 \rangle^{1/2}/\rho$. Demonstrate that when one considers observations of a macroscopic system (i.e., the size of the region becomes macroscopic, $L^3 \rightarrow \infty$) the relative fluctuations become negligible.
- (b) A fluid is at its “critical point” when

$$(\partial\beta p/\partial\rho)_\beta = (\partial^2\beta p/\partial\rho^2)_\beta = 0.$$

Determine the critical point density and temperature for the fluid obeying the van der Waals equation. That is, compute β_c and ρ_c as a function of a and b .

- (c) Focus attention on a subvolume of size L^3 in the fluid. Suppose L^3 is 100 times the space filling volume of a molecule—that is, $L^3 \approx 100b$. For this region in the fluid, compute the relative size of the density fluctuations when $\rho = \rho_c$, and the temperature is 10% above the critical temperature. Repeat this calculation for temperatures 0.1% and 0.001% from the critical temperature.

- (d) Light that we can observe with our eyes has wavelengths of the order of 1000 \AA . Fluctuations in density cause changes in the index of refraction, and those changes produce scattering of light. Therefore, if a region of fluid 1000 \AA across contains significant density fluctuations, we will visually observe these fluctuations. On the basis of the type of calculation performed in part (b), determine how close to the critical point a system must be before critical fluctuations become optically observable. The phenomenon of long wavelength density fluctuations in a fluid approaching the critical point is known as critical opalescence. (Note: You will need to estimate the size of b , and to do this you should note that the typical diameter of a small molecule is around 5 \AA .)

- 3.23. Consider a solution containing a solute species at very low concentrations. The solute molecules undergo conformational transitions between two isomers, A and B . Let N_A and N_B denote the numbers of A and B isomers, respectively. While the total number of solute molecules $N = N_A + N_B$ remain constant, at any instant the values of N_A and N_B differ from their mean values of $\langle N_A \rangle$ and $\langle N_B \rangle$. Show that mean square fluctuations are given by

$$\langle (N_A - \langle N_A \rangle)^2 \rangle = x_A x_B N,$$

where x_A and x_B are the average mole fractions of A and B species; that is,

$$x_A = \langle N_A \rangle / N.$$

[Hint: You will need to assume that the solutes are at such a low concentration that each solute molecule is uncorrelated from every other solute molecule. See Sec. 3.6.]

Bibliography

There are many excellent texts treating the basic elementary principles of statistical mechanics. For example:

- T. L. Hill, *Introduction to Statistical Thermodynamics* (Addison-Wesley, Reading, Mass., 1960).
 G. S. Rushbrooke, *Introduction to Statistical Mechanics* (Oxford University Press, Oxford, 1951).
 D. McQuarrie, *Statistical Mechanics* (Harper & Row, N.Y., 1976).
 F. Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, N.Y., 1965).

The following are somewhat more advanced than this text:

S. K. Ma, *Statistical Mechanics* (World Scientific, Philadelphia, 1985).

R. Balescu, *Equilibrium and Nonequilibrium Statistical Mechanics* (John Wiley, N.Y., 1975).

M. Toda, R. Kubo, and N. Saito, *Statistical Physics I* (Springer-Verlag, N.Y., 1983).

Each contains chapters that introduce modern ideas about ergodicity and chaos.

A colorful discussion of the history and principles of statistical mechanics is given by

P. W. Atkins, *The Second Law* (Scientific American Books and W. H. Freeman and Co., N.Y., 1984).

This book also contains instructive illustrations of concepts concerning chaos and chaotic structures including codes to run on microcomputers.