

## Chapter 2

# Statistical Mechanics

The topic of this book is computer simulation. Computer simulation allows us to study properties of many-particle systems. However, not all properties can be directly measured in a simulation. Conversely, most of the quantities that can be measured in a simulation do not correspond to properties that are measured in real experiments. To give a specific example: in a Molecular Dynamics simulation of liquid water, we could measure the instantaneous positions and velocities of all molecules in the liquid. However, this kind of information cannot be compared to experimental data, because no real experiment provides us with such detailed information. Rather, a typical experiment measures an average property, averaged over a large number of particles and, usually, also averaged over the time of the measurement. If we wish to use computer simulation as the numerical counterpart of experiments, we must know what kind of averages we should aim to compute. In order to explain this, we need to introduce the language of statistical mechanics. This we shall do here. We provide the reader with a quick (and slightly dirty) derivation of the basic expressions of statistical mechanics. The aim of these derivations is only to show that there is nothing mysterious about concepts such as phase space, temperature and entropy and many of the other statistical mechanical objects that will appear time and again in the remainder of this book.

### 2.1 Entropy and Temperature

Most of the computer simulations that we discuss are based on the assumption that classical mechanics can be used to describe the motions of atoms and molecules. This assumption leads to a great simplification in almost all calculations, and it is therefore most fortunate that it is justified in many cases of practical interest. Surprisingly, it turns out to be easier to derive the

basic laws of statistical mechanics using the language of quantum mechanics. We will follow this route of least resistance. In fact, for our derivation, we need only little quantum mechanics. Specifically, we need the fact that a quantum mechanical system can be found in different states. For the time being, we limit ourselves to quantum states that are eigenvectors of the Hamiltonian  $\mathcal{H}$  of the system (i.e., energy eigenstates). For any such state  $|i\rangle$ , we have that  $\mathcal{H}|i\rangle = E_i|i\rangle$ , where  $E_i$  is the energy of state  $|i\rangle$ . Most examples discussed in quantum mechanics textbooks concern systems with only a few degrees of freedom (e.g., the one-dimensional harmonic oscillator or a particle in a box). For such systems, the degeneracy of energy levels will be small. However, for the systems that are of interest to statistical mechanics (i.e., systems with  $\mathcal{O}(10^{23})$  particles), the degeneracy of energy levels is astronomically large. In what follows, we denote by  $\Omega(E, V, N)$  the number of eigenstates with energy  $E$  of a system of  $N$  particles in a volume  $V$ . We now express the basic assumption of statistical mechanics as follows: a system with fixed  $N, V$ , and  $E$  is equally likely to be found in any of its  $\Omega(E)$  eigenstates. Much of statistical mechanics follows from this simple (but highly nontrivial) assumption.

To see this, let us first consider a system with total energy  $E$  that consists of two weakly interacting subsystems. In this context, *weakly interacting* means that the subsystems can exchange energy but that we can write the total energy of the system as the sum of the energies  $E_1$  and  $E_2$  of the subsystems. There are many ways in which we can distribute the total energy over the two subsystems such that  $E_1 + E_2 = E$ . For a given choice of  $E_1$ , the total number of degenerate states of the system is  $\Omega_1(E_1) \times \Omega_2(E_2)$ . Note that the total number of states is not the sum but the product of the number of states in the individual systems. In what follows, it is convenient to have a measure of the degeneracy of the subsystems that is additive. A logical choice is to take the (natural) logarithm of the degeneracy. Hence:

$$\ln \Omega(E_1, E - E_1) = \ln \Omega_1(E_1) + \ln \Omega_2(E - E_1). \quad (2.1.1)$$

We assume that subsystems 1 and 2 can exchange energy. What is the most likely distribution of the energy? We know that *every* energy state of the total system is equally likely. But the number of eigenstates that correspond to a given distribution of the energy over the subsystems depends very strongly on the value of  $E_1$ . We wish to know the most likely value of  $E_1$ , that is, the one that maximizes  $\ln \Omega(E_1, E - E_1)$ . The condition for this maximum is that

$$\left( \frac{\partial \ln \Omega(E_1, E - E_1)}{\partial E_1} \right)_{N, V, E} = 0 \quad (2.1.2)$$

or, in other words,

$$\left( \frac{\partial \ln \Omega_1(E_1)}{\partial E_1} \right)_{N_1, V_1} = \left( \frac{\partial \ln \Omega_2(E_2)}{\partial E_2} \right)_{N_2, V_2}. \quad (2.1.3)$$

We introduce the shorthand notation

$$\beta(E, V, N) \equiv \left( \frac{\partial \ln \Omega(E, V, N)}{\partial E} \right)_{N, V}. \quad (2.1.4)$$

With this definition, we can write equation (2.1.3) as

$$\beta(E_1, V_1, N_1) = \beta(E_2, V_2, N_2). \quad (2.1.5)$$

Clearly, if initially we put all energy in system 1 (say), there will be energy transfer from system 1 to system 2 until equation (2.1.3) is satisfied. From that moment on, no net energy flows from one subsystem to the other, and we say that the two subsystems are in (thermal) equilibrium. When this equilibrium is reached,  $\ln \Omega$  of the total system is at a maximum. This suggests that  $\ln \Omega$  is somehow related to the thermodynamic entropy  $S$  of the system. After all, the second law of thermodynamics states that the entropy of a system  $N, V$ , and  $E$  is at its maximum when the system is in thermal equilibrium. There are many ways in which the relation between  $\ln \Omega$  and entropy can be established. Here we take the simplest route; we simply define the entropy to be equal to  $\ln \Omega$ . In fact, for (unfortunate) historical reasons, entropy is not simply equal to  $\ln \Omega$ ; rather we have

$$S(N, V, E) \equiv k_B \ln \Omega(N, V, E), \quad (2.1.6)$$

where  $k_B$  is Boltzmann's constant, which in S.I. units has the value  $1.38066 \times 10^{-23}$  J/K. With this identification, we see that our assumption that all degenerate eigenstates of a quantum system are equally likely immediately implies that, in thermal equilibrium, the entropy of a composite system is at a maximum. It would be a bit premature to refer to this statement as the second law of thermodynamics, as we have not yet demonstrated that the present definition of entropy is, indeed, equivalent to the thermodynamic definition. We simply take an advance on this result.

The next thing to note is that thermal equilibrium between subsystems 1 and 2 implies that  $\beta_1 = \beta_2$ . In everyday life, we have another way to express the same thing: we say that two bodies brought into thermal contact are in equilibrium if their temperatures are the same. This suggests that  $\beta$  must be related to the absolute temperature. The thermodynamic definition of temperature is

$$1/T = \left( \frac{\partial S}{\partial E} \right)_{N, V}. \quad (2.1.7)$$

If we use the same definition here, we find that

$$\beta = 1/(k_B T). \quad (2.1.8)$$

Now that we have defined *temperature*, we can consider what happens if we have a system (denoted by  $A$ ) that is in thermal equilibrium with a large heat

bath (B). The total system is closed; that is, the total energy  $E = E_B + E_A$  is fixed (we assume that the system and the bath are weakly coupled, so that we may ignore their interaction energy). Now suppose that the system A is prepared in one specific quantum state  $i$  with energy  $E_i$ . The bath then has an energy  $E_B = E - E_i$  and the degeneracy of the bath is given by  $\Omega_B(E - E_i)$ . Clearly, the degeneracy of the bath determines the probability  $P_i$  to find system A in state  $i$ :

$$P_i = \frac{\Omega_B(E - E_i)}{\sum_j \Omega_B(E - E_j)}. \quad (2.1.9)$$

To compute  $\Omega_B(E - E_i)$ , we expand  $\ln \Omega_B(E - E_i)$  around  $E_i = 0$ :

$$\ln \Omega_B(E - E_i) = \ln \Omega_B(E) - E_i \frac{\partial \ln \Omega_B(E)}{\partial E} + \mathcal{O}(1/E) \quad (2.1.10)$$

or, using equations (2.1.6) and (2.1.7),

$$\ln \Omega_B(E - E_i) = \ln \Omega_B(E) - E_i/k_B T + \mathcal{O}(1/E). \quad (2.1.11)$$

If we insert this result in equation (2.1.9), we get

$$P_i = \frac{\exp(-E_i/k_B T)}{\sum_j \exp(-E_j/k_B T)}. \quad (2.1.12)$$

This is the well-known Boltzmann distribution for a system at temperature  $T$ . Knowledge of the energy distribution allows us to compute the average energy  $\langle E \rangle$  of the system at the given temperature  $T$ :

$$\langle E \rangle = \sum_i E_i P_i \quad (2.1.13)$$

$$\begin{aligned} &= \frac{\sum_i E_i \exp(-E_i/k_B T)}{\sum_j \exp(-E_j/k_B T)} \\ &= - \frac{\partial \ln \sum_i \exp(-E_i/k_B T)}{\partial (1/k_B T)} \\ &= - \frac{\partial \ln Q}{\partial (1/k_B T)}, \end{aligned} \quad (2.1.14)$$

where, in the last line, we have defined the partition function  $Q$ . If we compare equation (2.1.13) with the thermodynamic relation

$$E = \frac{\partial F/T}{\partial (1/T)},$$

where  $F$  is the Helmholtz free energy, we see that  $F$  is related to the partition function  $Q$ :

$$F = -k_B T \ln Q = -k_B T \ln \left( \sum_i \exp(-E_i/k_B T) \right). \quad (2.1.15)$$

Strictly speaking,  $F$  is fixed only up to a constant. Or, what amounts to the same thing, the reference point of the energy can be chosen arbitrarily. In what follows, we can use equation (2.1.15) without loss of generality. The relation between the Helmholtz free energy and the partition function is often more convenient to use than the relation between  $\ln \Omega$  and the entropy. As a consequence, equation (2.1.15) is the workhorse of equilibrium statistical mechanics.

## 2.2 Classical Statistical Mechanics

Thus far, we have formulated statistical mechanics in purely quantum mechanical terms. The entropy is related to the density of states of a system with energy  $E$ , volume  $V$ , and number of particles  $N$ . Similarly, the Helmholtz free energy is related to the partition function  $Q$ , a sum over all quantum states  $i$  of the Boltzmann factor  $\exp(-E_i/k_B T)$ . To be specific, let us consider the average value of some observable  $A$ . We know the probability that a system at temperature  $T$  will be found in an energy eigenstate with energy  $E_i$  and we can therefore compute the thermal average of  $A$  as

$$\langle A \rangle = \frac{\sum_i \exp(-E_i/k_B T) \langle i|A|i \rangle}{\sum_j \exp(-E_j/k_B T)}, \quad (2.2.1)$$

where  $\langle i|A|i \rangle$  denotes the expectation value of the operator  $A$  in quantum state  $i$ . This equation suggests how we should go about computing thermal averages: first we solve the Schrödinger equation for the (many-body) system of interest, and next we compute the expectation value of the operator  $A$  for all those quantum states that have a nonnegligible statistical weight. Unfortunately, this approach is doomed for all but the simplest systems. First of all, we cannot hope to solve the Schrödinger equation for an arbitrary many-body system. And second, even if we could, the number of quantum states that contribute to the average in equation (2.2.1) would be so astronomically large ( $\mathcal{O}(10^{10^{25}})$ ) that a numerical evaluation of all expectation values would be unfeasible. Fortunately, equation (2.2.1) can be simplified to a more workable expression in the classical limit. To this end, we first rewrite equation (2.2.1) in a form that is independent of the specific basis set. We note that  $\exp(-E_i/k_B T) = \langle i|\exp(-\mathcal{H}/k_B T)|i \rangle$ , where  $\mathcal{H}$  is the Hamiltonian of the system. Using this relation, we can write

$$\begin{aligned} \langle A \rangle &= \frac{\sum_i \langle i|\exp(-\mathcal{H}/k_B T)A|i \rangle}{\sum_j \langle j|\exp(-\mathcal{H}/k_B T)|j \rangle} \\ &= \frac{\text{Tr} \exp(-\mathcal{H}/k_B T)A}{\text{Tr} \exp(-\mathcal{H}/k_B T)}, \end{aligned} \quad (2.2.2)$$

where  $\text{Tr}$  denotes the trace of the operator. As the value of the trace of an operator does not depend on the choice of the basis set, we can compute thermal averages using any basis set we like. Preferably, we use simple basis sets, such as the set of eigenfunctions of the position or the momentum operator. Next, we use the fact that the Hamiltonian  $\mathcal{H}$  is the sum of a kinetic part  $\mathcal{K}$  and a potential part  $\mathcal{U}$ . The kinetic energy operator is a quadratic function of the momenta of all particles. As a consequence, momentum eigenstates are also eigenfunctions of the kinetic energy operator. Similarly, the potential energy operator is a function of the particle coordinates. Matrix elements of  $\mathcal{U}$  therefore are most conveniently computed in a basis set of position eigenfunctions. However,  $\mathcal{H} = \mathcal{K} + \mathcal{U}$  itself is not diagonal in either basis set nor is  $\exp[-\beta(\mathcal{K} + \mathcal{U})]$ . However, if we could replace  $\exp(-\beta\mathcal{H})$  by  $\exp(-\beta\mathcal{K})\exp(-\beta\mathcal{U})$ , then we could simplify equation (2.2.2) considerably. In general, we cannot make this replacement because

$$\exp(-\beta\mathcal{K})\exp(-\beta\mathcal{U}) = \exp\{-\beta[\mathcal{K} + \mathcal{U} + \mathcal{O}([\mathcal{K}, \mathcal{U}])]\},$$

where  $[\mathcal{K}, \mathcal{U}]$  is the commutator of the kinetic and potential energy operators while  $\mathcal{O}([\mathcal{K}, \mathcal{U}])$  is meant to note all terms containing commutators and higher-order commutators of  $\mathcal{K}$  and  $\mathcal{U}$ . It is easy to verify that the commutator  $[\mathcal{K}, \mathcal{U}]$  is of order  $\hbar$  ( $\hbar = h/(2\pi)$ , where  $h$  is Planck's constant). Hence, in the limit  $\hbar \rightarrow 0$ , we may ignore the terms of order  $\mathcal{O}([\mathcal{K}, \mathcal{U}])$ . In that case, we can write

$$\text{Tr} \exp(-\beta\mathcal{H}) \approx \text{Tr} \exp(-\beta\mathcal{U}) \exp(-\beta\mathcal{K}). \quad (2.2.3)$$

If we use the notation  $|r\rangle$  for eigenvectors of the position operator and  $|k\rangle$  for eigenvectors of the momentum operator, we can express equation (2.2.3) as

$$\text{Tr} \exp(-\beta\mathcal{H}) = \sum_{r,k} \langle r|e^{-\beta\mathcal{U}}|r\rangle \langle r|k\rangle \langle k|e^{-\beta\mathcal{K}}|k\rangle \langle k|r\rangle. \quad (2.2.4)$$

All matrix elements can be evaluated directly:

$$\langle r|\exp(-\beta\mathcal{U})|r\rangle = \exp[-\beta\mathcal{U}(\mathbf{r}^N)],$$

where  $\mathcal{U}(\mathbf{r}^N)$  on the right-hand side is no longer an operator but a function of the coordinates of all  $N$  particles. Similarly,

$$\langle k|\exp(-\beta\mathcal{K})|k\rangle = \exp\left[-\beta \sum_{i=1}^N p_i^2/(2m_i)\right],$$

where  $p_i = \hbar k_i$ , and

$$\langle r|k\rangle \langle k|r\rangle = 1/V^N,$$

where  $V$  is the volume of the system and  $N$  the number of particles. Finally, we can replace the sum over states by an integration over all coordinates and momenta. The final result is

$$\begin{aligned} \text{Tr} \exp(-\beta\mathcal{H}) &\approx \frac{1}{h^{dN}N!} \int d\mathbf{p}^N d\mathbf{r}^N \exp\left\{-\beta \left[\sum_i p_i^2/(2m_i) + \mathcal{U}(\mathbf{r}^N)\right]\right\} \\ &\equiv Q_{\text{classical}}, \end{aligned} \quad (2.2.5)$$

where  $d$  is the dimensionality of the system and the last line defines the classical partition function. The factor  $1/N!$  has been inserted afterward to take the indistinguishability of identical particles into account. Every  $N$ -particle quantum state corresponds to a volume  $h^{dN}$  in classical phase space, but not all such volumes correspond to distinct quantum states. In particular, all points in phase space that only differ in the labeling of the particles correspond to the same quantum state (for more details, see, e.g., [43]).

Similarly, we can derive the classical limit for  $\text{Tr} \exp(-\beta\mathcal{H})A$ , and finally, we can write the classical expression for the thermal average of the observable  $A$  as

$$\langle A \rangle = \frac{\int d\mathbf{p}^N d\mathbf{r}^N \exp\left\{-\beta \left[\sum_i p_i^2/(2m_i) + \mathcal{U}(\mathbf{r}^N)\right]\right\} A(\mathbf{p}^N, \mathbf{q}^N)}{\int d\mathbf{p}^N d\mathbf{r}^N \exp\left\{-\beta \left[\sum_j p_j^2/(2m_j) + \mathcal{U}(\mathbf{r}^N)\right]\right\}}. \quad (2.2.6)$$

Equations (2.2.5) and (2.2.6) are the starting point for virtually all classical simulations of many-body systems.

### 2.2.1 Ergodicity

Thus far, we have discussed the average behavior of many-body systems in a purely static sense: we introduced only the assumption that every quantum state of a many-body system with energy  $E$  is equally likely to be occupied. Such an average over all possible quantum states of a system is called an *ensemble* average. However, this is not the way we usually think about the average behavior of a system. In most experiments we perform a series of measurements during a certain time interval and then determine the average of these measurements. In fact, the idea behind Molecular Dynamics simulations is precisely that we can study the average behavior of a many-particle system simply by computing the natural time evolution of that system numerically and averaging the quantity of interest over a sufficiently long time. To take a specific example, let us consider a fluid consisting of atoms. Suppose that we wish to compute the average density of the fluid at a distance  $r$  from a given atom  $i$ ,  $\rho_i(r)$ . Clearly, the instantaneous density depends on the coordinates  $\mathbf{r}_j$  of all particles  $j$  in the system. As time progresses, the atomic coordinates will change (according to Newton's equations of motion), and hence the density around atom  $i$  will change.

Provided that we have specified the initial coordinates and momenta of all atoms ( $\mathbf{r}^N(0), \mathbf{p}^N(0)$ ) we know, at least in principle, the time evolution of  $\rho_i(\mathbf{r}; \mathbf{r}^N(0), \mathbf{p}^N(0), t)$ . In a Molecular Dynamics simulation, we measure the time-averaged density  $\overline{\rho_i(\mathbf{r})}$  of a system of  $N$  atoms, in a volume  $V$ , at a constant total energy  $E$ :

$$\overline{\rho_i(\mathbf{r})} = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t dt' \rho_i(\mathbf{r}; t'). \quad (2.2.7)$$

Note that, in writing down this equation, we have implicitly assumed that, for  $t$  sufficiently long, the time average does not depend on the initial conditions. This is, in fact, a subtle assumption that is not true in general (see, e.g., [44]). However, we shall disregard subtleties and simply assume that, once we have specified  $N$ ,  $V$ , and  $E$ , time averages do not depend on the initial coordinates and momenta. If that is so, then we would not change our result for  $\overline{\rho_i(\mathbf{r})}$  if we average over many different initial conditions; that is, we consider the hypothetical situation where we run a large number of Molecular Dynamics simulations at the same values for  $N$ ,  $V$ , and  $E$ , but with different initial coordinates and momenta,

$$\overline{\rho_i(\mathbf{r})} = \frac{\sum_{\text{initial conditions}} \left( \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t dt' \rho_i(\mathbf{r}; \mathbf{r}^N(0), \mathbf{p}^N(0), t') \right)}{\text{number of initial conditions}}. \quad (2.2.8)$$

We now consider the limiting case where we average over all initial conditions compatible with the imposed values of  $N$ ,  $V$ , and  $E$ . In that case, we can replace the sum over initial conditions by an integral:

$$\frac{\sum_{\text{initial conditions}} f(\mathbf{r}^N(0), \mathbf{p}^N(0))}{\text{number of initial conditions}} \rightarrow \frac{\int_E d\mathbf{r}^N d\mathbf{p}^N f(\mathbf{r}^N(0), \mathbf{p}^N(0))}{\Omega(N, V, E)}, \quad (2.2.9)$$

where  $f$  denotes an arbitrary function of the initial coordinates  $\mathbf{r}^N(0), \mathbf{p}^N(0)$ , while  $\Omega(N, V, E) = \int_E d\mathbf{r}^N d\mathbf{p}^N$  (we have ignored a constant factor<sup>1</sup>). The subscript  $E$  on the integral indicates that the integration is restricted to a shell of constant energy  $E$ . Such a “phase space” average, corresponds to the classical limit of the *ensemble* average discussed in the previous sections.<sup>2</sup> We

<sup>1</sup>If we consider a quantum mechanical system, then  $\Omega(N, V, E)$  is simply the number of quantum states of that system, for given  $N$ ,  $V$ , and  $E$ . In the classical limit, the number of quantum states of a  $d$ -dimensional system of  $N$  distinguishable, structureless particles is given by  $\Omega(N, V, E) = (\int d\mathbf{p}^N d\mathbf{r}^N) / h^{dN}$ . For  $N$  indistinguishable particles, we should divide the latter expression by a factor  $N!$ .

<sup>2</sup>Here we consider the classical equivalent of the so-called microcanonical ensemble, i.e., the ensemble of systems with fixed  $N$ ,  $V$ , and  $E$ . The classical expression for phase space integrals in the microcanonical ensemble can be derived from the quantum mechanical expression involving a sum over quantum states in much the same way that we used to derive the classical constant  $N, V, T$  (“canonical”) ensemble from the corresponding quantum mechanical expression.

denote an ensemble average by  $\langle \dots \rangle$  to distinguish it from a time average, denoted by a bar. If we switch the order of the time averaging and the averaging over initial conditions, we find

$$\overline{\rho_i(\mathbf{r})} = \lim_{t \rightarrow \infty} \frac{1}{t} \int dt' \langle \rho_i(\mathbf{r}; \mathbf{r}^N(0), \mathbf{p}^N(0), t') \rangle_{NVE}. \quad (2.2.10)$$

However, the ensemble average in this equation does not depend on the time  $t^*$ . This is so, because there is a one-to-one correspondence between the initial phase space coordinates of a system and those that specify the state of the system at a later time  $t'$  (see e.g., [44, 45]). Hence, averaging over all initial phase space coordinates is equivalent to averaging over the time-evolved phase space coordinates. For this reason, we can leave out the time averaging in equation (2.2.10), and we find

$$\overline{\rho_i(\mathbf{r})} = \langle \rho_i(\mathbf{r}) \rangle_{NVE}. \quad (2.2.11)$$

This equation states that, if we wish to compute the average of a function of the coordinates and momenta of a many-particle system, we can *either* compute that quantity by time averaging (the “MD” approach) *or* by ensemble averaging (the “MC” approach). It should be stressed that the preceding paragraphs are meant only to make equation (2.2.11) *plausible*, not as a proof. In fact, that would have been quite impossible because equation (2.2.11) is not true in general. However, in what follows, we shall simply assume that the “ergodic hypothesis”, as equation (2.2.11) is usually referred to, applies to the systems that we study in computer simulations. The reader, however, should be aware that many examples of systems are not ergodic *in practice*, such as glasses and metastable phases, or even *in principle*, such as nearly harmonic solids.

## 2.3 Questions and Exercises

### Question 1 (Number of Configurations)

1. Consider a system  $A$  consisting of subsystems  $A_1$  and  $A_2$ , for which  $\Omega_1 = 10^{20}$  and  $\Omega_2 = 10^{22}$ . What is the number of configurations available to the combined system? Also, compute the entropies  $S$ ,  $S_1$ , and  $S_2$ .
2. By what factor does the number of available configurations increase when  $10 \text{ m}^3$  of air at  $1.0 \text{ atm}$  and  $300 \text{ K}$  is allowed to expand by  $0.001\%$  at constant temperature?
3. By what factor does the number of available configurations increase when  $150 \text{ kJ}$  is added to a system containing  $2.0 \text{ mol}$  of particles at constant volume and  $T = 300 \text{ K}$ ?

4. A sample consisting of five molecules has a total energy  $5\epsilon$ . Each molecule is able to occupy states of energy  $\epsilon_j$ , with  $j = 0, 1, 2, \dots, \infty$ . Draw up a table with columns by the energy of the states and write beneath them all configurations that are consistent with the total energy. Identify the type of configuration that is most probable.

**Question 2 (Thermodynamic Variables in the Canonical Ensemble)** Starting with an expression for the Helmholtz free energy ( $F$ ) as a function of  $N, V, T$

$$F = -\frac{\ln [Q(N, V, T)]}{\beta}$$

one can derive all thermodynamic properties. Show this by deriving equations for  $U, p$ , and  $S$ .

**Question 3 (Ideal Gas (Part 1))** The canonical partition function of an ideal gas consisting of monoatomic particles is equal to

$$Q(N, V, T) = \frac{1}{h^{3N} N!} \int d\Gamma \exp[-\beta H] = \frac{V^N}{\lambda^{3N} N!}$$

in which  $\lambda = h/\sqrt{2\pi m/\beta}$  and  $d\Gamma = dq_1 \dots dq_N dp_1 \dots dp_N$ .

Derive expressions for the following thermodynamic properties:

- $F(N, V, T)$  (hint:  $\ln(N!) \approx N \ln(N) - N$ )
- $p(N, V, T)$  (which leads to the ideal gas law !!!)
- $\mu(N, V, T)$  (which leads to  $\mu = \mu_0 + RT \ln \rho$ )
- $U(N, V, T)$  and  $S(N, V, T)$
- $C_V$  (heat capacity at constant volume)
- $C_p$  (heat capacity at constant pressure)

**Question 4 (Ising Model)** Consider a system of  $N$  spins arranged on a lattice. In the presence of a magnetic field,  $H$ , the energy of the system is

$$U = -\sum_{i=1}^N H \mu s_i - J \sum_{i>j} s_i s_j$$

in which  $J$  is called the coupling constant ( $J > 0$ ) and  $s_i = \pm 1$ . The second summation is a summation over all pairs ( $D \times N$  for a periodic system,  $D$  is the dimensionality of the system). This system is called the Ising model.

1. Show that for positive  $J$ , and  $H = 0$ , the lowest energy of the Ising model is equal to

$$U_0 = -DNJ$$

in which  $D$  is the dimensionality of the system.

2. Show that the free energy per spin of a 1D Ising model with zero field is equal to

$$\frac{F(\beta, N)}{N} = -\frac{\ln(2 \cosh(\beta J))}{\beta}$$

when  $N \rightarrow \infty$ . The function  $\cosh(x)$  is defined as

$$\cosh(x) = \frac{\exp[-x] + \exp[x]}{2}. \quad (2.3.1)$$

3. Derive equations for the energy and heat capacity of this system.

**Question 5 (The Photon Gas)** An electromagnetic field in thermal equilibrium can be described as a phonon gas. From the quantum theory of the electromagnetic field, it is found that the total energy of the system ( $U$ ) can be written as the sum of photon energies:

$$U = \sum_{j=1}^N n_j \omega_j \hbar = \sum_{j=1}^N n_j \epsilon_j$$

in which  $\epsilon_j$  is the characteristic energy of a photon with frequency  $\omega$ ,  $j$ ,  $n_j = 0, 1, 2, \dots, \infty$  is the so-called occupancy number of mode  $j$ , and  $N$  is the number of field modes (here we take  $N$  to be finite).

1. Show that the canonical partition function of the system can be written as

$$Q = \prod_{j=1}^N \frac{1}{1 - \exp[-\beta \epsilon_j]}. \quad (2.3.2)$$

Hint: you will have to use the following identity for  $|x| < 1$ :

$$\sum_{i=0}^{\infty} x^i = \frac{1}{1-x}. \quad (2.3.3)$$

For the product of partition functions of two independent systems  $A$  and  $B$  we can write

$$Q_A \times Q_B = Q_{AB} \quad (2.3.4)$$

when  $A \cap B = \emptyset$  and  $A \cup B = AB$ .

2. Show that the average occupancy number of state  $j$ ,  $\langle n_j \rangle$ , is equal to

$$\langle n_j \rangle = \frac{\partial \ln Q}{\partial (-\beta \epsilon_j)} = \frac{1}{\exp[\beta \epsilon_j] - 1}. \quad (2.3.5)$$

3. Describe the behavior of  $\langle n_j \rangle$  when  $T \rightarrow \infty$  and when  $T \rightarrow 0$ .

**Question 6 (Ideal Gas (Part 2))** An ideal gas is placed in a constant gravitational field. The potential energy of  $N$  gas molecules at height  $z$  is  $Mgz$ , where  $M = mN$  is the total mass of  $N$  molecules. The temperature in the system is uniform and the system infinitely large. We assume that the system is locally in equilibrium, so we are allowed to use a local partition function.

1. Show that the grand-canonical partition function of a system in volume  $V$  at height  $z$  is equal to

$$Q(\mu, V, T, z) = \sum_{N=0}^{\infty} \frac{\exp[\beta\mu N]}{h^{3N} N!} \int d\Gamma \exp[-\beta(H_0 + Mgz)] \quad (2.3.6)$$

in which  $H_0$  is the Hamiltonian of the system at  $z = 0$ .

2. Explain that a change in  $z$  is equivalent to a change in chemical potential,  $\mu$ . Use this to show that the pressure of the gas at height  $z$  is equal to

$$p(z) = p(z=0) \times \exp[-\beta mgz]. \quad (2.3.7)$$

(Hint: you will need the formula for the chemical potential of an ideal gas.)

### Exercise 1 (Distribution of Particles)

Consider an ideal gas of  $N$  particles in a volume  $V$  at constant energy  $E$ . Let us divide the volume in  $p$  identical compartments. Every compartment contains  $n_i$  molecules such that

$$N = \sum_{i=1}^{i=p} n_i. \quad (2.3.8)$$

An interesting quantity is the distribution of molecules over the  $p$  compartments. Because the energy is constant, every possible eigenstate of the system will be equally likely. This means that in principle it is possible that one of the compartments is empty.

1. On the book's website you can find a program that calculates the distribution of molecules among the  $p$  compartments. Run the program for different numbers of compartments ( $p$ ) and total number of gas molecules ( $N$ ). Note that the code has to be completed first (see the file *distribution.f*). The output of the program is the probability of finding  $x$  particles in a particular compartment as a function of  $x$ . This is printed in the file *output.dat*.
2. What is the probability that one of the compartments is empty?
3. Consider the case  $p = 2$  and  $N$  even. The probability of finding  $N/2 + n_1$  molecules in compartment 1 and  $N/2 - n_1$  molecules in compartment 2 is given by

$$P(n_1) = \frac{N!}{(N/2 - n_1)! (N/2 + n_1)! 2^N}. \quad (2.3.9)$$

Compare your numerical results with the analytical expression for different values of  $N$ . Show that this distribution is a Gaussian for small  $n_1/N$ . Hint: For  $x > 10$ , it might be useful to use Stirling's approximation:

$$x! \approx (2\pi)^{\frac{1}{2}} x^{x+\frac{1}{2}} \exp[-x]. \quad (2.3.10)$$

### Exercise 2 (Boltzmann Distribution)

Consider a system of  $N$  energy levels with energies  $0, \epsilon, 2\epsilon, \dots, (N-1)\epsilon$  and  $\epsilon \gg 0$ .

1. Calculate, using the given program, the occupancy of each level for different values of the temperature. What happens at high temperatures?
2. Change the program in such a way that the degeneracy of energy level  $i$  equals  $i + 1$ . What do you see?
3. Modify the program in such a way that the occupation of the energy levels as well as the partition function ( $q$ ) is calculated for a hetero nuclear linear rotor with moment of inertia  $I$ . Compare your result with the approximate result

$$q = \frac{2I}{\beta \hbar^2} \quad (2.3.11)$$

for different temperatures. Note that the energy levels of a linear rotor are

$$U = J(J+1) \frac{\hbar^2}{2I} \quad (2.3.12)$$

with  $J = 0, 1, 2, \dots, \infty$ . The degeneracy of level  $J$  equals  $2J + 1$ .

### Exercise 3 (Coupled Harmonic Oscillators)

Consider a system of  $N$  harmonic oscillators with a total energy  $U$ . A single harmonic oscillator has energy levels  $0, \epsilon, 2\epsilon, \dots, \infty$  ( $\epsilon > 0$ ). All harmonic oscillators in the system can exchange energy.

1. Invent a computational scheme to update the system at constant total energy ( $U$ ). Compare your scheme with the scheme that is incorporated in the computer code that you can find on the book's website (see the file *harmonic.f*).
2. Make a plot of the energy distribution of the first oscillator as a function of the number of oscillators for a constant value of  $U/N$  (*output.dat*). Which distribution is recovered when  $N$  becomes large? What is the function of the other  $N - 1$  harmonic oscillators? Explain.
3. Compare this distribution with the canonical distribution of a single oscillator at the same average energy (use the option *NVT*).
4. How does this exercise relate to the derivation of the Boltzmann distribution for a system at temperature  $T$ ?

**Exercise 4 (Random Walk on a 1D Lattice)**

Consider the random walk of a single particle on a line. The particle performs jumps of fixed length 1. Assuming that the probability for forward or backward jumps is equal, the mean-squared displacement of a particle after  $N$  jumps is equal to  $N$ . The probability that, after  $N$  jumps, the net distance covered by the particle equals  $n$  is given by

$$\ln [P(n, N)] \approx \frac{1}{2} \ln \left( \frac{2}{\pi N} \right) - \frac{n^2}{2N}.$$

1. Derive this equation using Stirling's approximation for  $\ln x!$ .
2. Compare your numerical result for the root mean-squared displacement with the theoretical prediction (the computed function  $P(n, N)$ , see the file *output.dat*). What is the diffusivity of this system?
3. Modify the program in such a way that the probability to jump in the forward direction equals 0.8. What happens?

**Exercise 5 (Random Walk on a 2D Lattice)**

Consider the random walk of  $N$  particles on a  $M \times M$  lattice. Two particles cannot occupy the same lattice site. On this lattice, periodic boundaries are used. This means that when a particle leaves the lattices it returns on the opposite side of the lattice; i.e., the coordinates are given modulo  $M$ .

1. What is the fraction of occupied sites ( $\theta$ ) of the lattice as a function of  $M$  and  $N$ ?
2. Make a plot of the diffusivity  $D$  as a function of  $\theta$  for  $M = 32$ . For low values of  $\theta$ , the diffusivity can be approximated by

$$D \approx D_0 (1 - \theta).$$

Why is this equation reasonable at low densities? Why does it break down at higher densities?

3. Modify the program in such a way that the probability to jump in one direction is larger than the probability to jump in the other direction. Explain the results.
4. Modify the program in such a way that periodic boundary conditions are used in one direction and reflecting boundary conditions in the other. What happens?

## Chapter 3

# Monte Carlo Simulations

In the present chapter, we describe the basic principles of the Monte Carlo method. In particular, we focus on simulations of systems of a fixed number of particles ( $N$ ) in a given volume ( $V$ ) at a temperature ( $T$ ).

### 3.1 The Monte Carlo Method

In the previous chapter, we introduced some of the basic concepts of (classical) statistical mechanics. Our next aim is to indicate where the Monte Carlo method comes in. We start from the classical expression for the partition function  $Q$ , equation (2.2.5):

$$Q = c \int d\mathbf{p}^N d\mathbf{r}^N \exp[-\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N)/k_B T], \quad (3.1.1)$$

where  $\mathbf{r}^N$  stands for the coordinates of all  $N$  particles, and  $\mathbf{p}^N$  for the corresponding momenta. The function  $\mathcal{H}(\mathbf{q}^N, \mathbf{p}^N)$  is the Hamiltonian of the system. It expresses the total energy of an isolated system as a function of the coordinates and momenta of the constituent particles:  $\mathcal{H} = \mathcal{K} + \mathcal{U}$ , where  $\mathcal{K}$  is the kinetic energy of the system and  $\mathcal{U}$  is the potential energy. Finally,  $c$  is a constant of proportionality, chosen such that the sum over quantum states in equation (2.1.15) approaches the classical partition function in the limit  $\hbar \rightarrow 0$ . For instance, for a system of  $N$  identical atoms,  $c = 1/(h^{3N} N!)$ . The classical equation corresponding to equation (2.2.1) is

$$\langle A \rangle = \frac{\int d\mathbf{p}^N d\mathbf{r}^N A(\mathbf{p}^N, \mathbf{r}^N) \exp[-\beta \mathcal{H}(\mathbf{p}^N, \mathbf{r}^N)]}{\int d\mathbf{p}^N d\mathbf{r}^N \exp[-\beta \mathcal{H}(\mathbf{p}^N, \mathbf{r}^N)]}, \quad (3.1.2)$$

where  $\beta = 1/k_B T$ . In this equation, the observable  $A$  has been expressed as a function of coordinates and momenta. As  $\mathcal{K}$  is a quadratic function of