Notes on Statistical Thermodynamics. Autumn 2001.

Daan Frenkel
frenkel@amolf.nl
Teaching Assistants:
Daniele Moroni: moroni@science.uva.nl
Tim Tambach: tambach@science.uva.nl

1 Introduction

Much of chemistry is devoted to the study of existing molecules or the synthesis of new ones. But, as the number of possible molecules is innumerably large, we have to distinguish between "interesting" and "uninteresting" molecules. In other words, we select molecular materials on basis of their properties. What properties are interesting, depends on the application that we have in mind. Some molecules may have interesting *chemical* properties (e.g. proteins that may catalyze a particular biochemical process). Other molecular materials may have interesting *optical* properties (e.g. cyano-biphenyls that are commonly used in liquid-crystal displays). And still other materials may have unique *mechanical* properties (e.g. carbon, in the form of diamond, is extremely hard - unlike carbon in the form of graphite or C_{60}). And clearly there are many more properties that are potentially interesting.

Obviously, we should like to be able to predict the properties of a substance, given its atomic composition. Can we? For sufficiently simple, isolated molecules such as H_2, N_2, H_2O or even a chain molecule such as polyethylene, the answer is *yes*. Starting from a quantum-mechanical description, we can predict certain properties of molecules, such as their three-dimensional structure (for instance, we can explain why benzene is flat, while NH_3 is not). We can also compute the charge distribution in a molecule. This allows us to predict the direction and size of the dipole moment of a molecule. In addition to the ground-state properties of a molecule, we can also predict the location of the various excited states (translation, rotation, vibration and electronic excitation). But, even if we know all these properties of isolated molecules, we still cannot predict the properties of a system that contains a macroscopic number of molecules ($N = \mathcal{O}(10^{23})$). For instance, how could we predict, on basis of our knowledge of the isolated molecules, that at room temperature and atmospheric pressure, H_2 forms gas, H_2O a liquid, and $NaCl$ an ionic crystal? Below, we give a very incomplete list of macroscopic properties that we should like to be able to predict:

mechanical
hardness
strength
dielectric constant
refractive index
transport
diffusion
heat conductivity
viscosity
electrical conductivity

thermodynamic
equation of state $P(\rho, T)$
heat capacity (c_v, c_p)
chemical potential (μ)
Phase diagram
chemical
equilibrium constant
reaction rate
catalytic properties

What we need is an approach that allows us to predict the "collective" macroscopic properties of molecular substances, on basis of our knowledge of the constituent molecules. Naively, one might think that the only thing that we have to do, in order to compute the properties of any system, is to solve the Schrödinger equation for that system. However, for all but the simplest systems, this is utterly impossible - certainly for a macroscopic system. How about thermodynamics? Well, thermodynamics only provides us with *relations* between measurable quantities. Relations such as

$$dU = q + w \tag{1}$$

or

$$dU = TdS - PdV + \mu dN \tag{2}$$

or

$$\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial P}{\partial T}\right)_V \tag{3}$$

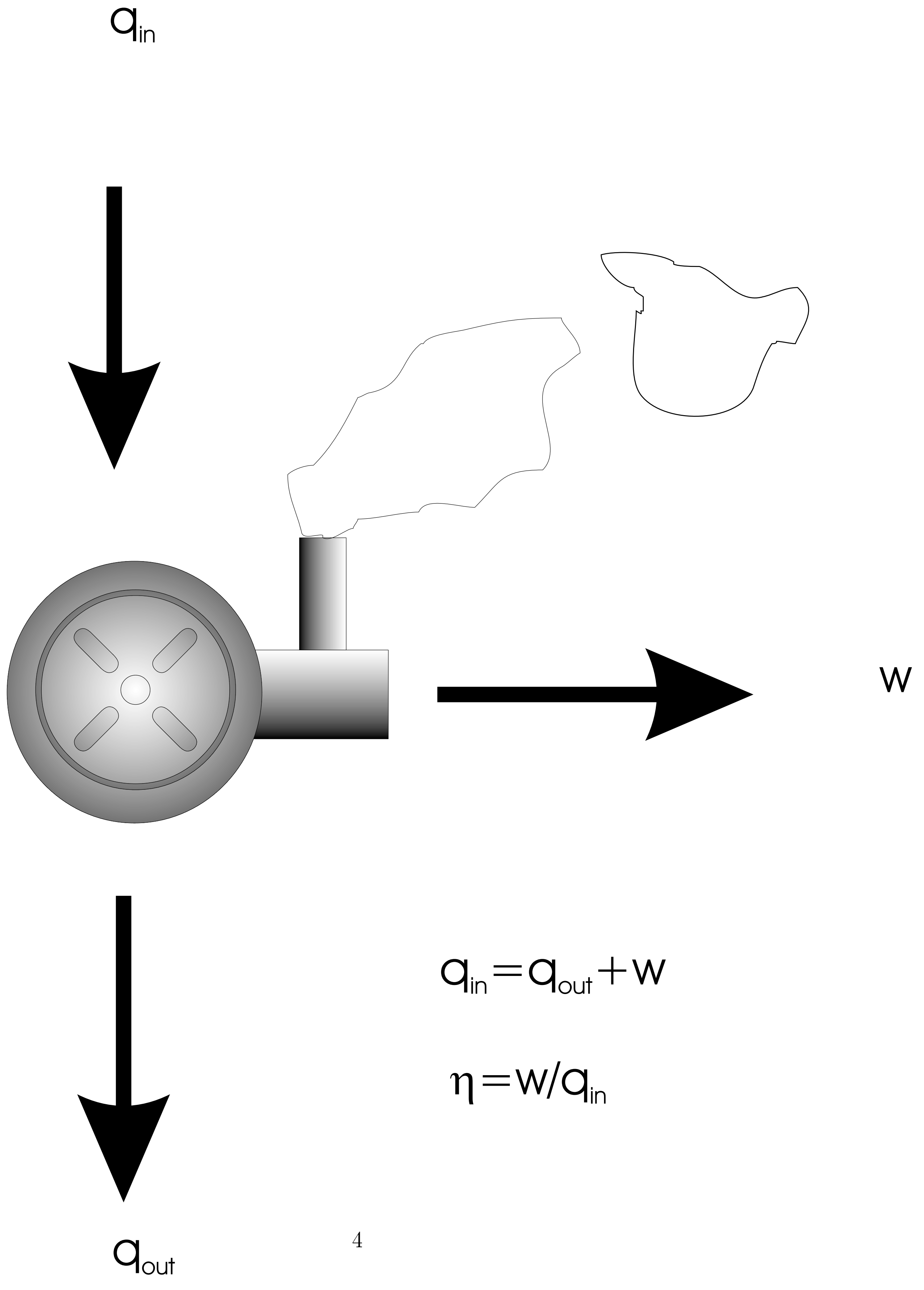
These relations are valid for *any substance.* But - precisely for this reason - it contains no information whatsoever about the molecular composition of a substance. Hence, neither quantum mechanics nor thermodynamics can provide us with the relations that we seek. *Statistical thermodynamics - the subject of this course - is the approach that provides us with the theoretical framework to relate the properties of individual molecules to the macroscopic behavior of molecular substances.* In most of the simple cases that will be discussed during these lectures, we can use pencil and paper to compute the properties of interest. However, in most practical cases the calculations become so involved that a computer has to take over the role of the pencil and paper. Yet the basic approach - statistical thermodynamics - is always the same. Before we embark on statistical thermodynamics, we need to review the core of thermodynamics.

2 Thermodynamics: a brief refresher course.

Thermodynamics is difficult because it seems so abstract. However, we should always bear in mind that thermodynamics is based on experimental observations. For instance, the First Law of thermodynamics expresses the empirical observation that energy is conserved, even though it can be converted in various forms. The internal energy of a system can be changed by performing work w on the system or by transferring an amount of heat q. It is meaningless to speak about the total amount of heat in a system, or the total amount of work. This is not mysterious: it is just as meaningless to speak about the number of train-travelers and the number of pedestrians in a train-station: people enter the station as pedestrians, and exit as train-travelers or vice versa. However if we add the sum of the changes in the number of train-travelers and pedestrians, then we obtain the change in the number of *people* in the station. And this quantity is well defined. Similarly, the sum of q and w is equal to the change in the internal energy U of the system

$$dU = q + w \tag{4}$$

This is the First Law of thermodynamics. The Second Law seems more abstract, but it is not. The Second Law is based on the *experimental* observation that it is impossible to make an engine that works by converting heat from a single heat bath (i.e. a large reservoir in equilibrium) into work. This observation is equivalent to another - equally empirical - observation, namely that heat can *never* flow spontaneously (i.e. without performing work) form a cold reservoir to a warmer reservoir. This statement is actually a bit more subtle than it seems because, before we have defined temperature, we can only distinguish *hotter* and *colder* by looking at the direction of heat flow. What the Second Law says is that it is *never* possible to make heat flow spontaneously in the "wrong" direction. How do we get from such a seemingly trivial statement to something as abstract as entropy? This is most easily achieved by introducing the concept of a *reversible* heat engine. A reversible engine is, as the word suggests, an engine that can be operated in reverse. During one cycle (a sequence of steps that is completed when the engine is returned into its original state) this engine takes in an amount of heat q_1 from a hot reservoir converts part of it into work w and delivers a remaining amount of heat q_2 to a cold reservoir. The reverse process is that, by performing an amount of work w, we can take an amount of heat q_2 form the cold reservoir and deliver an amount of heat $q_1$to the



$$q_{in} = q_{out} + w$$

$$\eta = w/q_{in}$$

hot reservoir. Reversible engines are an idealization because in any real engine there will be additional heat losses. However, the ideal reversible engine can be approximated arbitrarily closely by a real engine if, at every stage, the real engine is sufficiently close to equilibrium. As the engine is returned to its original state at the end of one cycle, its internal energy U has not changed. Hence, the First Law tells us that

$$dU = q_1 - (w + q_2) = 0 \tag{5}$$

or

$$q_1 = w + q_2 \tag{6}$$

Now consider the "efficiency" of the engine $\eta \equiv w/q_1$- i.e. the amount of work delivered per amount of heat taken in. At first, one might think that η depends on the precise design of our reversible engine. However this is not true. η is the same for *all* reversible engines operating between the same two reservoirs. To demonstrate this, we show that if different engines could have different values for η then we would contradict the Second Law in its form "heat can never spontaneously flow from a cold to a hot reservoir". Suppose therefore that we have another reversible engine that takes in an amount of heat q'_1 form the hot reservoir, delivers the same amount of work w, and then delivers an amount of heat $q'_2$to the cold reservoir. Let us denote the efficiency of this engine by η'. Now we use the work generated by the engine with the highest efficiency (say η) to drive the second engine in reverse. The amount of heat delivered to the hot reservoir by the second engine is

$$q'_1 = w/\eta' = q_1(\eta/\eta') \tag{7}$$

where we have used $w = q_1\eta$. As, by assumption, $\eta' < \eta$ it follows that $q'_1 > q_1$. Hence *there is a net heat flow form the cold reservoir into the hot reservoir.* But this contradicts the Second Law of thermodynamics. Therefore we must conclude that the efficiency of all reversible heat engines operating between the same reservoirs is *identical.* The efficiency only depends on the temperatures $t_1$and t_2 of the reservoirs (the temperatures t could be measured in any scale, *e.g.* in Fahrenheit as long as heat flows in the direction of decreasing t). As $\eta(t_1, t_2)$ depends only on the temperature in the reservoirs, then so does the ratio $q_2/q_1 = 1 - \eta$. Let us call this ratio $R(t_2, t_1)$. Now suppose that we have a reversible engine that consists of two stages: one working between reservoir 1 and 2, and the other between 2 and 3. In addition, we have another reversible engine that works directly between 1 and 3. As both engines must be equally efficient, it follows that

$$R(t_3, t_1) = R(t_3, t_2)R(t_2, t_1) \tag{8}$$

This can only be true in general if $R(t_1, t_2)$ is of the form

$$R(t_2, t_1) = \frac{f(t_2)}{f(t_1)} \tag{9}$$

where $f(t)$ is an, as yet unknown function of our measured temperature. What we do now is to introduce an "absolute" or thermodynamic temperature T given by

$$T = f(t) \tag{10}$$

Then it immediately follows that

$$\frac{q_2}{q_1} = R(t_2, t_1) = \frac{T_2}{T_1} \tag{11}$$

Note that the thermodynamic temperature could just as well have been defined as $c \times f(t)$. In practice, c has been fixed such that, around room temperature, 1 degree in the absolute (Kelvin) scale is equal to 1 degree Celsius. But that choice is of course purely historical and - as it will turn out later - a bit unfortunate.

Now, why do we need all this? We need it to introduce entropy, this most mysterious of all thermodynamic quantities. To do so, note that Eqn.11 can be written as

$$\frac{q_1}{T_1} = \frac{q_2}{T_2} \tag{12}$$

where q_1 is the heat that flows *in* reversibly at the high temperature T_1, and q_2 is the heat that flows *out* reversibly at the low temperature T_2. We see therefore that, during a complete cycle, the difference between q_1/T_1 and q_2/T_2 is zero. Recall that, at the end of a cycle, the internal energy of the system has not changed. Now Eqn.12 tells us that there is also another quantity that we call "entropy" and that we denote by S that is unchanged when we restore the system to its original state. In the language of thermodynamics, we call S a *state function*. We do not know what S is, but we do know how to compute its change. In the above example, the change in S was given by $\Delta S = (q_1/T_1) - (q_2/T_2) = 0$. In general, the change in entropy of a system due to the reversible addition of an infinitesimal amount of heat δq_{rev} from a reservoir at temperature T is

$$dS = \frac{\delta q_{rev}}{T} \tag{13}$$

We also note that S is *extensive*. That means that the total entropy of two non-interacting systems, is equal to the sum of the entropy of the individual systems. Consider a system with a fixed number of particles N and a fixed volume V. If we transfer an infinitesimal amount of heat

δq to this system,then the change in the internal energy of the system, dU is equal to δq.Hence,

$$\left(\frac{\partial S}{\partial U}\right)_{V,N} = \frac{1}{T} \tag{14}$$

The most famous (though not most intuitively obvious) statement of the Second Law of Thermodynamics is that *Any spontaneous change in a closed system (i.e. a system that exchanges neither heat nor particles with its environment) can never lead to a decrease of the entropy.* Hence, in equilibrium, the entropy of a closed system is at a maximum. Can we understand this? Well, let us first consider a system with an energy E, volume V and number of particles N that *is* in equilibrium. Let us denote the entropy of this system by $S_0(E, V, N)$. In equilibrium, all spontaneous changes that *can* happen, *have* happened. Now suppose that we want to change something in this system - for instance, we increase the density in one half and decrease it in the other. As the system was in equilibrium, this change does not occur spontaneously. Hence, in order to realize this change, we must perform a certain amount of *work,* w (for instance, my placing a piston in the system and moving it). Let us perform this work reversibly in such a way that E, the total energy of the system stays constant (and also V and N). The First Law tells us that we can only keep E constant if, while we do the work, we allow an amount of heat q, flow out of the system, such that $q = w$. Eqn.13 tells us that when an amount of heat q flows out of the system, the entropy S of the system must decrease. Let us denote the entropy of this constrained state by $S_1(E, V, N) < S_0(E, V, N)$. Having completed the change in the system, we insulate the system thermally from the rest of the world, and we remove the constraint that kept the system in its special state (taking the example of the piston: we make an opening in the piston). Now the system goes back *spontaneously* (and irreversibly) to equilibrium. However, no work is done and no heat is transferred. Hence the finally energy E is equal to the original energy (and V and N) are also constant. This means that the system is now back in its *original* equilibrium state and its entropy is once more equal to $S_0(E, V, N)$. The entropy change during this spontaneous change is equal to $\Delta S = S_0 - S_1$.But, as $S_1 < S_0$, it follows that $\Delta S > 0$. As this argument is quite general, we have indeed shown that any spontaneous change in a closed system leads to an increase in the entropy. Hence, in equilibrium, the entropy of a closed system is at a maximum.

From this point on, we can derive all of thermodynamics, except one “law” - the so-called Third Law of Thermodynamics. The Third law states that, at $T = 0$, the entropy of the equilibrium state of pure substance equals zero. Actually, the Third Law is not nearly as "basic" as the First and the Second. And, anyway, we shall soon get a more direct interpretation of its meaning.

3 Basics of Statistical Thermodynamics

Statistical thermodynamics is a theoretical framework that allows us to compute the macroscopic properties of systems containing many particles. In order to *'derive'* the basic laws of statistical thermodynamics, we need to use the language of quantum mechanics - but only very little. Specifically, we need the fact that a quantum mechanical system has discrete energy levels. Most examples discussed in quantum-chemistry textbooks concern systems with only one, or a few, particles moving in a very simple external potential (e.g. the one-dimensional harmonic oscillator or a particle in a box). For such systems, the degeneracy of energy levels (i.e. the number of states that have the same energy) will be small. However, for the systems that are of interest to statistical thermodynamics (i.e. systems with $\mathcal{O}(10^{23})$ particles), the degeneracy of energy levels is astronomically large - in what follows, it shall turn out that the word "astronomical" is actually misplaced - the numbers involved are so large that, by comparison the total number of particles in the universe is utterly negligible. In what follows, we denote by $\Omega(U, V, N)$ the number of energy levels with energy U of a system of N particles in a volume V. We now express the basic assumption of statistical thermodynamics as follows: *A system with fixed* N*,* V *and* U *is equally likely to be found in any of its* $\Omega(U)$ *energy levels.* Much of statistical thermodynamics follows from this simple (but highly non-trivial) assumption.

To see this, let us first consider a system with total energy U that consists of two weakly interaction sub-systems. In this context, 'weakly interacting' means that the sub-systems can exchange energy but that we can write the total energy of the system as the sum of the energies U_1 and U_2 of the sub-systems. There are many ways in which we can distribute the total energy over the two sub-systems, such that $U_1+U_2 = U$. For a given choice of U_1, the total number of degenerate states of the system is $\Omega_1(U_1) \times \Omega_2(U_2)$. 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 sub-systems that is extensive (i.e. additive). A logical choice is to take the (natural) logarithm of the degeneracy. Hence:

$$\ln \Omega(U_1, U - U_1) = \ln \Omega_1(U_1) + \ln \Omega_2(U - U_1) \tag{15}$$

We assume that sub-systems 1 and 2 can exchange energy. In fact, in thermodynamics we often consider such a process: it is simply heat transfer (because the two systems do not exchange particles, nor does one system perform work on the other). 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 energy levels that correspond to a given distribution of the energy over the sub-systems, depends very strongly on the value of U_1. We wish to know the most likely value of U_1, i.e. the one that maximizes $\ln \Omega(U_1, U - U_1)$. The condition for this maximum is that

$$\left(\frac{\partial \ln \Omega(U_1, U - U_1)}{\partial U_1}\right)_{N,V,U} = 0 \qquad (16)$$

or, in other words,

$$\left(\frac{\partial \ln \Omega_1(U_1)}{\partial U_1}\right)_{N_1,V_1} = \left(\frac{\partial \ln \Omega_2(U_2)}{\partial U_2}\right)_{N_2,V_2} . \qquad (17)$$

We introduce the shorthand notation

$$\beta(U, V, N) \equiv \left(\frac{\partial \ln \Omega(U, V, N)}{\partial U}\right)_{N,V} . \qquad (18)$$

With this definition, we can write Eqn. 17 as

$$\beta(U_1, V_1, N_1) = \beta(U_2, V_2, N_2). \qquad (19)$$

Clearly, if initially we put all energy in system 1 (say), there will be energy transfer from system 1 to system 2 until Eqn. 19 is satisfied. From that moment on, there is no net energy flow from one sub-system to the other, and we say that the two sub-systems are in thermal equilibrium. This implies that the condition $\beta(U_1, V_1, N_1) = \beta(U_2, V_2, N_2)$ must be equivalent to the statement that the two sub-systems have the same temperature. $\ln \Omega$ is a state function (of U, V and N), just like S. Moreover, when thermal equilibrium is reached, $\ln \Omega$ of the total system is at a maximum, again just like S. This suggests that $\ln \Omega$ is closely related to S. We note that both S and $\ln \Omega$ are extensive. This suggests that S is simply proportional to $\ln \Omega$:

$$S(N, V, U) \equiv k_B \ln \Omega(N, V, U) \qquad (20)$$

where k_B is Boltzmann's constant which, in S.I. units, has the value 1.3806503 10^{-23}J/K. This constant of proportionality cannot be derived - as we shall see later, it follows from the comparison with experiment. With this identification, we see that our assumption that all degenerate energy levels of a quantum system are equally likely immediately implies that, in thermal equilibrium, the entropy of a composite system is at a maximum. Hence, in the statistical picture, the Second Law of thermodynamics is not at all mysterious, it simply states that, in thermal equilibrium, the system is most likely to be found in the state that has the largest number of degenerate energy levels. The next thing to note is that thermal equilibrium between sub-systems 1 and 2 implies that $\beta_1 = \beta_2$. In thermodynamics, we have another way to express the same thing: we say that two bodies that are brought in thermal contact are in equilibrium if their temperatures are the same. This suggests that β must be related to the absolute temperature. The thermodynamic definition of temperature is

$$1/T = \left(\frac{\partial S}{\partial U}\right)_{V,N} \tag{21}$$

If we use the same definition here, we find that

$$\beta = 1/(k_B T) \ . \tag{22}$$

It is of course a bit unfortunate that we cannot simply say that $S = \ln \Omega$. The reason is that, historically, thermodynamics preceded statistical thermodynamics. In particular, the absolute thermodynamic temperature scale (see below Eqn.11) contained an arbitrary constant that, in the previous century was chosen such that one degree Kelvin matched one degree Celsius. If we could have introduced an absolute temperature scale now, we could have chosen it such that $S = \ln \Omega$ and the new absolute temperature T' would be equal to $k_B T$. However, this would create many practical problems because, in those units, room temperature would be of the order of $5 \ 10^{-21}$Joule (that is: entropy would be dimensionless, but temperature would have the dimension of energy). Few people would be happy with such a temperature scale. So we leave things as they are.

The Third Law of thermodynamics is also easy to understand from Eqn20: it simply states that, at $T = 0$, the number of accessible states of a pure substance, (Ω) is equal to one. In other words: at absolute zero, the system is in its ground state - and this ground state is non-degenerate.

One final comment: we mentioned that Ω is usually “super-astronomically” large. Let me quantify this: at room temperature the entropy of one mol of argon is 307.2 Joule/Kelvin and hence $\Omega = 10^{10^{25}}$. That is a *large* number.... Now we can also understand the Second Law of thermodynamics - in its form “the entropy of a closed system cannot decrease spontaneously”. Strictly speaking, this is not true. The probability that a system spontaneously transform from a state with an entropy S_A into a state with a lower entropy S_B is simply equal to

$$\frac{\Omega_B}{\Omega_A} = \exp([S_B - S_A]/k_B) \tag{23}$$

For instance, the probability that the entropy of one mol of argon at room temperature spontaneously decreases by as little as 0.00000001% would still be $10^{-10^{16}}$ which is -for all practical purposes - equal to zero. This illustrates the fact that Second Law of thermodynamics is an empirical law (i.e. it states what is or is not observed) and not a *mathematical* law - mathematicians would only say that the entropy of a closed system is *unlikely* to decrease.

3.1 System at constant temperature

Now that we have defined temperature, we can consider what happens if we have a system (denoted by S) that is in thermal equilibrium with a large "heat-bath"(B). The total system is closed, i.e. the total energy $U{=}U_B{+}U_S$ 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 S is prepared in a specific state i with energy ϵ_i. The bath then has an energy $U_B = U - \epsilon_i$ and the degeneracy of the bath is given by $\Omega_B(U - \epsilon_i)$. Clearly, the degeneracy of the bath determines the probability $P(\epsilon_i)$ to find system S in state i.

$$P(\epsilon_i) = \frac{\Omega_B(U - \epsilon_i)}{\sum_i \Omega_B(U - \epsilon_i)} \; . \tag{24}$$

To compute $\Omega_B(U - \epsilon_i)$, we expand $\ln \Omega_B(U - \epsilon_i)$ around $\epsilon_i{=}0$.

$$\ln \Omega_B(U - \epsilon_i) = \ln \Omega_B(U) - \epsilon_i \frac{\partial \ln \Omega_B(U)}{\partial U} + \mathcal{O}(1/U) \tag{25}$$

or, using Eqns. 21 and 22,

$$\ln \Omega_B(U - \epsilon_i) = \ln \Omega_B(U) - \epsilon_i/k_B T + \mathcal{O}(1/U) \tag{26}$$

If we insert this result in Eqn. 24, we get

$$P(\epsilon_i) = \frac{\exp(-\epsilon_i/k_B T)}{\sum_i \exp(-\epsilon_i/k_B T)} \tag{27}$$

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 $U_S =< \epsilon >$ of the system at the given temperature T

$$
\begin{aligned}
< \epsilon > = \sum_i \epsilon_i P(\epsilon_i) &= \frac{\sum_i \epsilon_i \exp(-\epsilon_i/k_BT)}{\sum_i \exp(-\epsilon_i/k_BT)} \\
&= -\frac{\partial \ln\left(\sum_i \exp(-\epsilon_i/k_BT)\right)}{\partial 1/k_BT} \\
&= -\frac{\partial \ln Q}{\partial 1/k_BT} \,, \qquad (28)
\end{aligned}
$$

where, in the last line, we have defined the partition function Q.

$$
Q \equiv \sum_i \exp(-\epsilon_i/k_BT) \qquad (29)
$$

If we compare Eqn. 28 with the thermodynamic relation

$$
U = \frac{\partial A/T}{\partial 1/T} \,,
$$

we see that the Helmholtz free energy A is related to the partition function Q:

$$
A = -k_BT \ln Q \,. \qquad (30)
$$

Strictly speaking, A is only fixed 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 Eqn. 30 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, Eqn. 30 is the "workhorse" of equilibrium statistical thermodynamics.

3.1.1 Further links with thermodynamics

Let us go back to the formulation of the second law of thermodynamics that states that, in a closed system at equilibrium, the entropy of the system is at a maximum. As argued above, we have a simple understanding of this law: the system is most likely to be found in the state that has the largest degeneracy. Now consider again a thermal reservoir in contact with a (macroscopic) system, under the condition that system *plus* reservoir are isolated. Then we know that the entropy of this combined system should be a maximum at equilibrium. As before, we can write the total degeneracy of system plus bath as

$$
\Omega_{total} = \Omega_{system}(U_S)\Omega_{bath}(U_{total} - U_S) \qquad (31)
$$

where U_S is the internal energy of the system, and U_{total} the total internal energy of system plus bath. The condition for equilibrium is that the derivative with respect to U_S of $\ln \Omega_{total}$ vanishes. As in Eqn.26, we expand $\ln \Omega_{bath}(U_{total} - U_S)$ up to linear order in U_S and we find

$$\ln \Omega_{total} = \ln \Omega_{system}(U_S) \ + \ln \Omega_{bath}(U_{total}) - \beta U_S \tag{32}$$

Note that $\ln \Omega_{bath}(U_{total})$ does not depend on U_S. Hence, to find the maximum of $\ln \Omega_{total}$, we have to locate the maximum of $\ln \Omega_{system}(U_S) - \beta U_S$. Hence, we arrive at the statement that, for a system in contact with a heat bath, the condition for equilibrium is that $\ln \Omega_{system}(U_S) - \beta U_S$ is at a maximum or, what amounts to the same thing $\beta U_S - \ln \Omega_{system}(U_S)$ is at a minimum. Now we make use of the fact that we have identified $k_B \ln \Omega_{system}$ with the entropy S of the system. Hence, we conclude that, at constant temperature, the condition for equilibrium is

$$\beta(U_S - TS) \text{ is at a minimum} \tag{33}$$

But from thermodynamics we know that $U - TS$ is nothing else than the Helmholtz free energy A. Hence, we immediately recover the well-known statement that - at constant temperature and volume - A is at a minimum in equilibrium.

Pressure Up to this point we have been considering a system that exchanges *energy* with a bath. Now let us consider a system that can exchange *volume* with a reservoir. As before, the total energy and the total volume of system plus reservoir are fixed. Let us denote this total volume by V_{tot} and the volume of the system of interest by V. As before, the condition for equilibrium is that the total entropy is at a maximum. Hence we have to determine the maximum with respect to V of

$$\ln \Omega(V, V_{tot} - V) = \ln \Omega_{sys}(V) + \ln \Omega_{bath}(V_{tot} - V) \tag{34}$$

or, using the identification between entropy and $k_B \ln \Omega$

$$\left(\frac{\partial S_{sys}}{\partial V}\right)_{U,N} + \left(\frac{\partial S_{bath}}{\partial V}\right)_{U,N} = 0 \tag{35}$$

However, from thermodynamics we know that

$$dS = \frac{dU}{T} + \frac{P}{T}dV - \frac{\mu}{T}dN \tag{36}$$

and hence

$$\left(\frac{\partial S}{\partial V}\right)_{U,N} = \frac{P}{T} \tag{37}$$

This expresses the fact that, if a system and a bath can exchange *both* energy *and* volume, then the conditions for equilibrium are

$$T_{sys} = T_{bath} \tag{38}$$

and

$$P_{sys} = P_{bath} \tag{39}$$

In practice, it is often more convenient to use the relation between the change in volume and the change in Helmholtz free energy A

$$dA = -PdV - SdT + \mu dN \tag{40}$$

and the corresponding expression for the pressure

$$P = -\left(\frac{\partial A}{\partial V}\right)_{T,N} \tag{41}$$

to obtain

$$P = k_B T \left(\frac{\partial \ln Q}{\partial V}\right)_{T,N} \tag{42}$$

Later on, we shall use this expression to compute the pressure of gases from our knowledge of the partition function.

3.1.2 Fluctuations

Careful readers may have noticed that, in the above sections, we used two seemingly different expressions for the Helmholtz free energy, namely

$$A = -k_B T \ln Q \tag{43}$$

and

$$A = U - TS = U - k_B T \ln \Omega \tag{44}$$

It turns out that, in the thermodynamic limit that we usually consider, these two expressions yield the same result. To see this, consider the expression for the partition function Q

$$Q = \sum_{\text{all quantum states } i} \exp(-\beta \epsilon_i) \tag{45}$$

Now let us group all levels with the same energy ϵ together. As before, we denote the number of energy levels with a particular energy ϵ by $\Omega(\epsilon)$. Then

$$Q = \sum_{\text{all energy levels } \epsilon} \Omega(\epsilon) \exp(-\beta \epsilon) \tag{46}$$

In a macroscopic system, the degeneracy of energy levels is very large. But, more importantly, it increases very rapidly with ϵ. At the same time, $\exp(-\beta\epsilon)$ is a very steeply decreasing function of ϵ. As a consequence, the product $\Omega(\epsilon)\exp(-\beta\epsilon)$ is a very sharply peaked function of ϵ. Let us expand $\Omega(\epsilon)\exp(-\beta\epsilon)$ around its maximum. Let us denote the value of ϵ at this maximum by U. First of all, we note that

$$\Omega(U+\Delta U) = \Omega(U)\exp(\left(\frac{\partial \ln \Omega}{\partial U}\right)\Delta U + \frac{1}{2}\left(\frac{\partial^2 \ln \Omega}{\partial U^2}\right)\Delta U^2 + \mathcal{O}(\Delta U^3)) \tag{47}$$

But we have already determined that, at the maximum, we have

$$\left(\frac{\partial \ln \Omega}{\partial U}\right) = \beta \tag{48}$$

Hence

$$\begin{aligned}\Omega(U+\Delta U)\exp(-\beta(U+\Delta U)) = \Omega(U)\exp(-\beta U)\exp(\frac{1}{2}\left(\frac{\partial^2 \ln \Omega}{\partial U^2}\right)\Delta U^2 \\ + \mathcal{O}(\Delta U^3))\end{aligned} \tag{49}$$

Now, it is easy to verify that

$$\begin{aligned}\left(\frac{\partial^2 \ln \Omega}{\partial U^2}\right) = \left(\frac{\partial \beta}{\partial U}\right) &= \left(\frac{\partial 1/k_B T}{\partial U}\right) \\ &= -\frac{1}{k_B T^2 c_V}\end{aligned} \tag{50}$$

The important thing to notice is that c_V, the isochoric heat capacity of the system, is an *extensive* quantity. That is, $c_V \sim N$. As a consequence,

$$\left(\frac{\partial^2 \ln \Omega}{\partial U^2}\right) \sim \frac{1}{N} \tag{51}$$

and, in the thermodynamic limit it becomes vanishingly small. The net result is that, in the thermodynamic limit, we can write

$$\begin{aligned}-k_B T \ln Q &\approx -k_B T \ln\left(\Omega(U)\exp(-\beta U)\right) \\ &= U - k_B T \ln \Omega(U)\end{aligned} \tag{52}$$

and therefore the two expression for A that we used are, indeed, equivalent.

4 Towards a molecular picture

The expressions for the Boltzmann distribution (Eqn.27) and the partition function (Eqn29) are quite general. No assumption was made about the size of the system - apart from the fact that the thermal reservoir should be much larger than the system. This implies that we can also use Eqn.27 to describe the Boltzmann distribution over energy levels of a single molecule and we can use Eqn.29to describe the partition function of an isolated molecule. This is extremely important, because it will allow us to use our experimental or theoretical knowledge of the energy levels of isolated molecules to predict the thermal properties of dilute gases!

Before computing specific molecular partition functions, it is useful to make a few general comments that will simplify the analysis. We are interested in the calculation of partition functions of the form

$$q = \sum_{i}^{\text{all levels}} \exp(-\beta u_i) \tag{53}$$

Such sums can easily be evaluated in a number of simple cases where we know the energy levels from quantum mechanics. Examples (that will be discussed below) are the case of a one-dimensional harmonic oscillator, a particle in a one-dimensional box or a rigid, linear rotor. However, in practice we shall always be interested in systems containing very many particles that are subject to translation, vibration and rotation (and possibly also to electronic or spin excitations). It would seem that the calculation of the partition function of such a complex system is totally intractable. In fact, for macroscopic systems of strongly interacting particles, this statement is very nearly true. This is what is commonly called the "many-body problem". Computing the partition function of - say - liquid water is *not* possible analytically and sophisticated numerical simulation techniques have been developed that allow us to make progress in this direction. However, there is an important class of problems where progress can be made without recourse to numerical simulation, namely the case where the total energy of the system can be written as a *sum* of simple contributions. This situation typically arises when there is virtually no interaction between the different particles in the system (i.e. in the case of an ideal gas). Let us consider a simple example. Suppose that we have a system with a total energy E, such that

$$E = \sum_{i=1}^{N} \epsilon_i(n_i) \tag{54}$$

where $\epsilon_i(n_i)$ is the energy of the i-th molecule in a quantum state labeled by the quantum number n_i. If all molecules have the same energy levels[1],

---

[1] We assume that, although the levels of different molecules have the same energy, they are not physically the same levels. Later, we shall consider what happens if the levels really are identical.

the partition function of this system is given by

$$\begin{aligned}
Q &= \sum_{n_1=1}^{\infty}\sum_{n_2=1}^{\infty}...\sum_{n_N=1}^{\infty}\exp(-\beta\sum_{i=1}^{N}\epsilon_i(n_i)) \\
&= \sum_{n_1=1}^{\infty}\sum_{n_2=1}^{\infty}...\sum_{n_N=1}^{\infty}\prod_{i=1}^{N}\exp(-\beta\epsilon_i(n_i)) \\
&= \left(\sum_{n_1=1}^{\infty}\exp(-\beta\epsilon_1(n_1))\right)\left(\sum_{n_2=1}^{\infty}\exp(-\beta\epsilon_2(n_2))\right)...\left(\sum_{n_N=1}^{\infty}\exp(-\beta\epsilon_N(n_N))\right) \\
&= q^N \qquad (55)
\end{aligned}$$

where

$$q \equiv \sum_{n=1}^{\infty}\exp(-\beta\epsilon(n)) \qquad (56)$$

In what follows, we shall often make use of this trick. To give a simple example, assume that a molecule has only two levels: a ground state with energy $\epsilon_0 = 0$ and an excited state with energy $\epsilon_1 = \epsilon$. In that case,

$$\begin{aligned}
q &= \exp(-\beta\times 0) + \exp(-\beta\epsilon) \\
&= 1 + \exp(-\beta\epsilon) \qquad (57)
\end{aligned}$$

and hence

$$Q = (1 + \exp(-\beta\epsilon))^N \qquad (58)$$

4.1 Single-molecule partition function

Quantum mechanics allows us to compute the energy levels of (simple) isolated molecules. Once we know these energy levels, we can use Eqn.29 to compute the molecular partition function. From this quantity, we can the compute a host of important thermodynamical properties of the individual molecules and, as we shall see later, also of macroscopic systems containing many ($\mathcal{O}(10^{23})$) molecules. For instance, we shall be able to compute the internal energy, the heat-capacity, the pressure and the chemical potential. From the latter quantity, we shall be able to *predict* the equilibrium constants of chemical reactions! But, before we do this, let us first consider the most important classes of molecular partition functions. Recall that the internal energy of a molecule can be due to various kinds of motion: translation, rotation, vibration and electronic excitation are the most common. Usually the energy of electronic excitations is very large compared to k_BT. As a consequence, the internal energy of a molecule at room temperature is usually due to translation, rotation and vibration. Translation is the motion of the molecule as a whole. As the intra-molecular motion (rotation and vibration) does not depend on its center-of-mass motion, we can write the energy of a molecule as the *sum* of the translational energy ϵ_{tr} and the intra-molecular energy ϵ_{int} due to rotation and vibration. It is often a fair approximation to assume that rotation and vibration are uncoupled. In that case, the total intra-molecular energy can be written as

$$\epsilon_{int} = \epsilon_{rot} + \sum_{i=1}^{m} \epsilon_{vib}^{i} \tag{59}$$

where ϵ_{rot} is the rotational energy and ϵ_{vib}^{i} is the energy associated with i-th vibration (diatomic molecules have only one vibration, poly-atomic molecules have many - we denote this number by m). The total energy of a molecule is then

$$\epsilon_{tot} = \epsilon_{trans} + \epsilon_{rot} + \sum_{i=1}^{m} \epsilon_{vib}^{i} \tag{60}$$

Of course, the value of ϵ_{tot} depends on the quantum state that the molecule is in. This quantum state is characterized by a set of quantum numbers k_α characterizing the translational motion in directions α ($\alpha = x, y, z$), a quantum number J describing the rotational motion and quantum numbers n_i describing the quantum-state associated with the i-th vibration. If we want to compute the molecular partition function, we have to evaluate the sum

$$q_{mol} = \sum_{k_x,k_y,k_z,,J,n_1,\ldots,n_m} \exp(-\beta\epsilon_{tot}) \tag{61}$$

This expression can be simplified appreciably if the internal energy of the molecule can be written as the sum of translational, rotational and vibrational energies. Using

$$\exp(-\beta[\epsilon_{trans}+\epsilon_{rot}+\sum_{i=1}^{m}\epsilon_{vib}^{i}]) = \exp(-\beta\epsilon_{trans})\exp(-\beta\epsilon_{rot})\prod_{i=1}^{m}\exp(-\beta\epsilon_{vib}^{i}) \tag{62}$$

we can write the molecular partition function as a product of a translational, a rotational and m vibrational partition functions

$$q_{mol} = q_{trans} q_{rot} \prod_{i=1}^{m} q_{vib}^{i} \tag{63}$$

where

$$\begin{aligned} q_{trans} &= \sum_{k_x,k_y,k_z} \exp(-\beta\epsilon_{trans}(k_x,k_y,k_z)) \\ q_{rot} &= \sum_J \exp(-\beta\epsilon_{rot}(J)) \\ q^i_{vib} &= \sum_{n_i} \exp(-\beta\epsilon_{vib}(n_i)) \end{aligned}$$

Hence, we can compute the partition functions for translation, rotation and vibration separately. The total molecular partition function is then simply the product. Actually, the situation is slightly more complex: when computing the total energy of a molecule, we need to know ϵ_{ground}, the energy of the molecule in its ground state. ϵ_{ground} is a measure for the binding energy of the molecule. It will turn out to be important later on (when we compute chemical equilibria). However, for the time being we ignore this term and simply pretend that the energy of a molecule in its ground state is zero. Now let us see if we can write down expressions for the translational, rotational and vibrational partition functions.

4.1.1 Vibration

Let us start with vibration (because it is easiest). If the vibrational motion is harmonic, we know from quantum mechanics that the vibrational levels are non-degenerate and all equally spaced. If the vibration frequency of the molecule is ν_i then the spacing between two levels is $h\nu_i$. If we take the potential energy at the bottom of the harmonic potential to be zero, then the energy of the n-th vibrational state is $\epsilon_n = (n + 1/2)h\nu_i$. The quantity $h\nu_i/2$ is the zero-point energy of the harmonic oscillator with frequency ν_i. The partition function associated with the i-th vibrational mode is

$$q^i_{vib} = \exp(-\beta h\nu_i/2) \sum_{n=0}^{\infty} \exp(-n\beta h\nu_i) \tag{64}$$

If we denote $\exp(-\beta h\nu_i)$ by x_i, then we see that the vibrational partition function is of the form

$$q^i_{vib} = x^{1/2} \sum_{n=0}^{\infty} x_i^n \tag{65}$$

But this is nothing but an infinite geometric series. As $\exp(-h\nu_i/k_BT) = x < 1$, we can sum this series and we obtain

$$q^i_{vib} = \frac{x^{1/2}}{1-x} = \frac{\exp(-\beta h\nu_i/2)}{1-\exp(-\beta h\nu_i)} \tag{66}$$

The probability to find this vibration is in its n-th excited state is

$$p(n_i) = \frac{\exp(-(n+1/2)\beta h\nu_i)}{q^i_{vib}} = [1-\exp(-\beta h\nu_i)]\exp(-n\beta h\nu_i) \tag{67}$$

We can also compute the mean vibrational energy. Using Eqn.28 we find that

$$< \epsilon^i_{vib} >= -\frac{\partial \ln q^i_{vib}}{\partial \beta} = h\nu_i/2 + \frac{h\nu_i \exp(-\beta h\nu_i)}{1-\exp(-\beta h\nu_i)} = h\nu_i/2 + \frac{h\nu_i}{\exp(\beta h\nu_i)-1} \tag{68}$$

Note that in the classical limit ($h \to 0$), we find that $< \epsilon^i_{vib} >= \beta^{-1} = k_B T$, irrespective of the frequency.

4.1.2 Translation

Now we use the quantum-mechanical result for the energy levels of a particle in a box. For a one-dimensional box of length L, these energy levels are given by

$$\epsilon(n) = \frac{n^2h^2}{8mL^2} \tag{69}$$

The lowest energy level is $\epsilon(1)$. However, for macroscopic systems the average energy is very much large than $\epsilon(1)$. Hence, in what follows we use Eqn.69 to compute the energy of level n, rather than $\epsilon(n) - \epsilon(1)$. We can now write down the expression for the translational partition function for a one-dimensional system:

$$q^x_{trans} = \sum_{n=1}^{\infty} \exp(-\frac{\beta n^2 h^2}{8mL^2}) \tag{70}$$

In macroscopic systems, the translational level spacing is typically much smaller than k_BT. As a consequence, we can replace the sum in the expression for q^x_{trans} by an integral

$$q^x_{trans} \approx \int_1^{\infty} dn \exp(-\frac{\beta n^2 h^2}{8mL^2}) \tag{71}$$

We now change variables. We define $x \equiv n\sqrt{\beta h^2/8mL^2}$. Then

$$q^x_{trans} \approx \sqrt{8mL^2/\beta h^2} \int_{\sqrt{\beta h^2/8mL^2}}^{\infty} dx \exp(-x^2) \tag{72}$$

But, because $\sqrt{\beta h^2/8mL^2} \ll 1$, we can replace the lower limit of this integral by 0. The integral

$$\int_0^{\infty} dx \exp(-x^2) = \frac{1}{2}\sqrt{\pi} \tag{73}$$

and hence

$$q^x_{trans} = \sqrt{2\pi m L^2/\beta h^2} \tag{74}$$

The result for a three-dimensional box follows because the total translational partition function is simply

$$q_{trans} = q^x_{trans} q^y_{trans} q^z_{trans} \tag{75}$$

For convenience, we assume that the particle is contained in a cubic box with volume $V = L^3$. Then

$$q_{trans} = \left(2\pi m L^2/\beta h^2\right)^{3/2} = V\left(2\pi m k_B T/h^2\right)^{3/2} \tag{76}$$

We can now easily compute the average translational energy of a particle. First we write q_{trans} as

$$q_{trans} = V\left(2\pi m/h^2\right)^{3/2} \beta^{-3/2} \tag{77}$$

Next, we use the relation

$$\begin{aligned} < \epsilon_{trans} > &= -\frac{\partial \ln q_{trans}}{\partial \beta} \\ &= -\frac{\partial \ln \beta^{-3/2}}{\partial \beta} = \frac{3}{2}\frac{\partial \ln \beta}{\partial \beta} \\ &= \frac{3}{2\beta} = \frac{3}{2} k_B T \end{aligned} \tag{78}$$

Hence the average translational energy per particle is $\frac{3}{2}k_B T$. If we have a system consisting of N non-interacting particles, we can simply add the translational energies of all particles and hence the total translational energy becomes

$$E_{trans} = \frac{3}{2} N k_B T \tag{79}$$

4.1.3 Rotation

The rotational energy levels of a molecule depend on its symmetry. Here, we only consider the simplest case: a linear rotor (examples are HCl and HCN, but *not* H_2O). The energy levels of a linear rotor depend on its moment of inertia I (see Appendix C.1.2):

$$E_J = J(J+1)\frac{\hbar^2}{2I} \tag{80}$$

and the degeneracy of an energy level with quantum number J is equal to $g_J = (2J+1)$. The rotational partition function of a linear rotor is then

$$q_{rot} = \sum_{J=0}^{\infty} (2J+1) \exp[-\beta J(J+1)\hbar^2/2I] \tag{81}$$

This sum can be easily evaluated numerically. But in the "classical" limit $\beta\hbar^2/2I \ll 1$ it is simple to get an analytical result. In that case we can write

$$q_{rot} \approx \int_0^\infty dJ\ (2J+1)\exp[-\beta J(J+1)\hbar^2/2I] \tag{82}$$

As the largest contribution to this integral comes from terms with $J \gg 1$, we can simplify the integral by noting that $2J+1 \approx 2J$ and $J(J+1) \approx J^2$. Then

$$\begin{aligned} q_{rot} &\approx \int_0^\infty dJ\ \ 2J\exp[-\beta J^2\hbar^2/2I] \\ &= \int_0^\infty dJ^2\ \ \exp[-\frac{\beta\hbar^2}{2I}J^2] \end{aligned} \tag{83}$$

The latter integral is easy to evaluate if we make the substitution $J^2 \equiv x$

$$\begin{aligned} q_{rot} &= \int_0^\infty dx\ \ \exp[-\frac{\beta\hbar^2}{2I}x] \\ &= \frac{2I}{\beta\hbar^2} = \frac{2Ik_BT}{\hbar^2} \end{aligned} \tag{84}$$

As before, we can now compute the average rotational energy per particle. It is

$$\begin{aligned} < \epsilon_{rot}^{lin} > &= -\frac{\partial \ln q_{rot}^{lin}}{\partial\beta} \\ &= \beta^{-1} = k_BT \end{aligned} \tag{85}$$

In the case of *non-linear* molecules, $q_{rot} \sim T^{3/2}$ in the classical limit. In that case

$$< \epsilon_{rot}^{non-lin} > = \frac{3}{2}k_BT \tag{86}$$

Note that, in the first paragraph of this section, we mentioned a heteronuclear diatomic such as HCl as an example of a linear molecule, but not the simple homo-nuclear molecules such as H_2 or N_2.The reason is that, due to the symmetry of these molecules, certain rotational states are forbidden. Which states are forbidden depends on the nature of the nuclei. For instance, ^{14}N is a boson (its nuclear spin is zero). For a diatomic molecule consisting of bosons, the total wave-function should remain identical if we permute the two nuclei. However, the wave-functions of a linear rotor have the property that they change sign if we rotate the molecule over 180 ° if J is odd. But rotating the molecule over 180 ° is equivalent to permuting the two nuclei. And for bosons the wave function should *not* change sign when we do that. Hence in that case, *all states with odd J are excluded.* The result is that the rotational partition function is reduced. In the classical limit, this reduction amounts to a factor 2. Similarly, if the molecule contains two identical fermionic nuclei (e.g. H_2 - nuclear spin 1/2) then the total wavefunction must change sign if we permute the two nuclei. This would suggest that *only* odd J-states are allowed. However, matters are a bit more complex because it makes a difference whether the two nuclear spins are parallel or anti-parallel. The total wave-function is in this case a product of the rotational wave function (that is changes sign when J is odd) and a *spin* wavefunction that changes sign when the spins are *anti-parallel.* As the *total* wavefunction should change sign under permutation of the two nuclei, we find that if the spins are parallel, then *J must be odd.* And if the spins are anti-parallel then *J must be even.* In either event, we have to exclude every second J value (either odd or even). In the classical limit, the net result is that the total rotational partition function is reduced by a factor 2,*just as for bosons.* We call this factor the *symmetry number* (σ) for the homo-nuclear diatomics. For more complex molecules we can compute the symmetry number systematically using group-theoretical arguments (but this would carry too far).

4.2 Many-particle partition function

In the above section, we have given expressions for q_{mol}, the partition function of isolated molecules in an ideal gas, and for the total energy of the gas. However, we did not yet write down the partition function for an ideal gas of N molecules. It would seem that this is trivial. By analogy to Eqn.55 we would expect that the total partition function simply equals q_{mol}^N. However, this is *not* the case. In fact, we can demonstrate experimentally that

$$Q_N \neq q_{mol}^N \tag{87}$$

To see this, consider the following experiment. We start with an ideal gas containing N particles that are all of one and the same kind - say, argon atoms. In that case, q_{mol} is equal to the translational partition function. Of these N atoms, we put $N/2$ in one half of a container with volume V and the other $N/2$ particles in the other half. The two halves of the container are kept at a temperature T. They are separated by a removable wall. If Eqn.87 were valid, than the total Helmholtz free energy of these two systems together would be

$$A_{\text{total}}^{\text{initial}} = -2k_BT\ln\left(\frac{V}{2}\left(2\pi mk_BT/h^2\right)^{3/2}\right)^{N/2} \tag{88}$$

$$= -Nk_BT\ln\left(\frac{V}{2}\left(2\pi mk_BT/h^2\right)^{3/2}\right) \tag{89}$$

Now assume that we take out the wall separating the two sub-systems. If the two halves of the volume had contained *different* gases, then we would see irreversible mixing, and the free energy would decrease. But, as the two halves contain identical particles, removing the wall between them has no effect whatsoever on the total free energy. Yet, if the relation $Q_N = q_{mol}^N$ were true, we would predict that the new free energy would be

$$\begin{aligned}A_{\text{total}}^{\text{final}} &= -k_BT\ln\left(V\left(2\pi mk_BT/h^2\right)^{3/2}\right)^N \\ &= -Nk_BT\ln\left(V\left(2\pi mk_BT/h^2\right)^{3/2}\right) \end{aligned} \tag{90}$$

This would imply that the Helmholtz free energy would change by an amount

$$A_{\text{total}}^{\text{final}} - A_{\text{total}}^{\text{initial}} = -Nk_BT\ln 2 \tag{91}$$

Clearly, something is wrong with the assumption that, for identical molecules, $Q_N = q_{mol}^N$. The resolution of this problem lies in the fact that we have been overcounting the number of distinct quantum states in the system. If N identical particles can occupy the same set of quantum levels, then any permutation of the particles over these levels still yields the same overall quantum state. That is, we can label the overall quantum state by indicating what levels $\{i, j, k, l, ...\}$ are occupied. But it is meaningless to say that particle 1 occupies level i and particle 2 occupies levels j and so on. There is no way in which we could distinguish this situation from the one in which particle 2 occupies level i and particle 1 occupies level j. In short: permutations of identical particles over a shared set of energy levels, do not correspond to distinct quantum states. This means that there is only *one* way to place N identical particles in N distinct quantum states, rather than $N!$ ways. Now look at Eqn.55, in particular at the expression

$$Q = \sum_{n_1=1}^{\infty}\sum_{n_2=1}^{\infty}\cdots\sum_{n_N=1}^{\infty}\exp(-\beta\sum_{i=1}^{N}\epsilon_i(n_i)) \tag{92}$$

This expression is valid if all sums over the quantum numbers n_i refer to *different* levels. This is, for instance, the case if we look at N harmonic oscillators: the situation where oscillator 1 is in the ground state and oscillator 2 is in the first excited state is *not* the same as the situation where oscillator 2 is in the ground state and oscillator 1 is in the first excited state. However, we have two identical particles in the same box, then the situation where particle 1 is in the ground state and particle 2 is in the first excited state is the same as the one where the two particles are permuted. For that reason, we should count all terms in the sum in Eqn.92 only once. This means that we should divide the sum in Eqn.55 by $N!$. Hence, the partition function of a system of N ideal gas molecules is

$$\begin{aligned} Q &= \frac{1}{N!}\sum_{n_1=1}^{\infty}\sum_{n_2=1}^{\infty}...\sum_{n_N=1}^{\infty}\exp(-\beta\sum_{i=1}^{N}\epsilon_i(n_i)) \\ &= \frac{q_{mol}^N}{N!} \end{aligned} \tag{93}$$

If we now look again at the two volumes with identical gases that were being combined, we find that the initial Helmholtz free energy was

$$A_{\text{total}}^{\text{initial}} = -2k_BT\ln\left\{\frac{\left(\frac{V}{2}\left(2\pi mk_BT/h^2\right)^{3/2}\right)^{N/2}}{(N/2)!}\right\} \tag{94}$$

and

$$A_{\text{total}}^{\text{final}} = -k_BT\ln\left\{\frac{\left(V\left(2\pi mk_BT/h^2\right)^{3/2}\right)^{N}}{N!}\right\} \tag{95}$$

The difference between these two free energies is

$$A_{\text{total}}^{\text{final}} - A_{\text{total}}^{\text{initial}} = -Nk_BT\ln 2 + k_BT\ln\left(\frac{N!}{\{(N/2)!\}^2}\right) \tag{96}$$

Using the Stirling approximation $\ln N! \approx N\ln N - N$, we find that

$$A_{\text{total}}^{\text{final}} - A_{\text{total}}^{\text{initial}} \approx -Nk_BT\ln 2 + Nk_BT\ln 2 = 0 \tag{97}$$

Hence the assumption that identical particles are truly indistinguishable is compatible with the experimental observation that the Helmholtz free energy does not change when we bring two reservoirs containing identical particles into contact.

We should add that the above discussion is not completely honest: we made the implicit assumption that no two particles would be in exactly the same quantum state. For an ideal gas at room temperature, this *assumption is quite reasonable. This can be seen as follows: recall that the molecular partition function is a measure for the number of quantum state accessible to a molecule. For the translational quantum states, this number is equal to*

$$n_{states} \approx V \left(2\pi m k_B T/h^2\right)^{3/2} \equiv \frac{V}{\Lambda^3} \tag{98}$$

where we have defined the "thermal wavelength" Λ of a particle. The number of accessible states is equal to the number of volumes Λ^3 that fit into the macroscopic volume V. To give a specific example: for argon at room temperature, $\Lambda \approx 1.7\ 10^{-11}$ m. Hence, the number of accessible quantum states in one cubic meter is 2 10^{32}. This is much larger than Avogadro's number. Hence, the probability that two atoms in a dilute gas will occupy the same level is very small indeed. However, Λ increases as the temperature is decreased and at very low temperatures we should take the possibility into account that two particles may occupy the same quantum state. Then it becomes important to distinguish between particles that cannot occupy the same quantum state (fermions) and particles that can (bosons). We shall not discuss this interesting topic here.

In what follows, we shall therefore assume that the partition function of an ideal gas of identical molecules is given by

$$Q = \frac{q_{mol}^N}{N!} \tag{99}$$

Now we start to make true on our promise that Statistical Thermodynamics allows us to predict the thermodynamical properties of a macroscopic system solely on basis of our (quantum-mechanical) knowledge of the constituent molecules. In particular, we are now in a position to *compute* all thermodynamical properties of a dilute molecular gas.

Let us begin with the pressure. This is a very important relation because it allows us to justify the identification

$$\beta = 1/k_B T.$$

Using the relation

$$P = \beta^{-1} \left(\frac{\partial \ln Q}{\partial V}\right)_{N,\beta} \tag{100}$$

we find

$$P = N\beta^{-1} \left(\frac{\partial \ln q_{mol}}{\partial V}\right)_{N,\beta} \tag{101}$$

But note that of the different terms in $q_{mol} = q_{trans} q_{rot} \prod_{i=1}^{m} q_{vib}^i$, only q_{trans} depends on the volume V. Hence

$$P = N\beta^{-1} \left(\frac{\partial \ln q_{trans}}{\partial V}\right)_{N,T} = \frac{N}{\beta V} \tag{102}$$

If we compare this to the experimental (Boyle/Gay-Lussac) relation

$$P = \frac{nRT}{V} \tag{103}$$

where n is the number of moles of the gas and R is the gas constant, then it follows that

$$\beta = \frac{N}{nRT} \tag{104}$$

The ratio N/n is simply the number of molecules per mole, i.e. Avogadro's number $L = 6.022\ 10^{23}$. We denote the ratio R/L by k_B (Boltzmann's constant). Then

$$\beta = \frac{1}{k_B T} \tag{105}$$

where k_B has the value $1.3806503\ 10^{-23}$J/K.

We can also compute the internal energy of an ideal gas. However, now we should take translation, rotation and vibration into account. But, as the total molecular partition function is a product of these individual terms, the internal energy

$$U = -N\left(\frac{\partial \ln q_{mol}}{\partial \beta}\right) \tag{106}$$

is the *sum* of a translational, a rotational and a vibrational part.

4.2.1 Monatomic gas.

Let us first consider a monatomic gas. In that case $q_{mol} = q_{trans}$ and we get (from Eqn.78) that

$$U = \frac{3}{2} N k_B T \tag{107}$$

It then follows immediately that the isochoric heat capacity of an ideal monatomic gas is

$$c_V = \left(\frac{\partial U}{\partial T}\right)_{N,V} = \frac{3}{2} N k_B \tag{108}$$

Let us next compute the Helmholtz free energy of an ideal monatomic gas

$$A = -k_B T \ln\left\{\frac{V^N}{N!\Lambda^{3N}}\right\} \tag{109}$$

Using the Stirling approximation for $\ln N!$,we get

$$\begin{aligned} A &= -Nk_BT\ln(V/\Lambda^3) + Nk_BT\ln N - Nk_BT \\ &= Nk_BT\{\ln(\rho\Lambda^3) - 1\} \end{aligned} \tag{110}$$

where $\rho \equiv (N/V)$ is the number density of the gas. It is then easy to obtain an expression for the Gibbs free energy $G = A + PV$.

$$\begin{aligned} G &= Nk_BT\{\ln(\rho\Lambda^3) - 1\} + Nk_BT \\ &= Nk_BT\ln(\rho\Lambda^3) \end{aligned} \tag{111}$$

When we recall that, for a one-component system, $G = N\mu$, where μ is the chemical potential, then we find that

$$\mu = k_BT\ln(\rho\Lambda^3) \tag{112}$$

or, if we use the ideal gas law, in the form $P = \rho k_BT$,

$$\mu = k_BT\ln(P\Lambda^3/k_BT) \tag{113}$$

Using $A = U - TS$, we can also get a expression for the entropy S from the expressions for the Helmholtz free energy A and for the internal energy U. We then find the famous Sackur-Tetrode equation for the entropy of a monatomic gas

$$\begin{aligned} S &= (U - A)/T \\ &= Nk_B\left(\frac{5}{2} - \ln(\rho\Lambda^3)\right) \end{aligned} \tag{114}$$

4.2.2 Polyatomic gases

When discussing the statistical thermodynamics of ideal polyatomic gases, it is important to bear in mind that the partition function of such a gas can be written as

$$Q_{total} = \frac{q_{trans}^N}{N!}q_{internal}^N \tag{115}$$

where $q_{internal}$ is the part of the molecular partition function that contains the contributions due to rotation, vibration and electronic excitation. Due to the factorization of the partition function we can write

$$A_{total} = A_{trans} - Nk_BT\ln q_{internal} \tag{116}$$

Therefore, we can simply add to the translational part of the free energy the contribution due to the internal degrees of freedom the molecules. This also holds for the Gibbs free energy G:

$$\begin{aligned} G_{total} &= G_{trans} - Nk_BT\ln q_{internal} \\ &= Nk_BT\ln(\rho\Lambda^3) - Nk_BT\ln q_{internal} \end{aligned} \tag{117}$$

Of particular interest is the expression for the chemical potential of an ideal gas of polyatomic molecules of species α

$$\mu_\alpha = k_B T \ln(\rho_\alpha \Lambda_\alpha^3) - k_B T \ln q_{internal}^\alpha \tag{118}$$

It is useful to cast this expression for the chemical potential in the form that is often used in thermodynamics

$$\mu(P) = \mu^\ominus + k_B T \ln(P/p^\ominus) \tag{119}$$

where $p^\ominus$ is the pressure of the reference state (usually: $p^\ominus$= 1 bar). For an ideal gas, $\rho_\alpha = P_\alpha / k_B T$, where P_α is the partial pressure of species α. Hence, we can write Eqn.113 as

$$\begin{aligned} \mu_\alpha &= k_B T \ln(\Lambda_\alpha^3 P_\alpha / k_B T) - k_B T \ln q_{internal}^\alpha \\ &= k_B T \ln(P_\alpha / p^\ominus) + k_B T \ln(\Lambda_\alpha^3 p^\ominus / k_B T) - k_B T \ln q_{internal}^\alpha \end{aligned} \tag{120}$$

and hence

$$\mu^\ominus = -k_B T \ln\left(\frac{q_{internal}^\alpha k_B T}{\Lambda_\alpha^3 p^\ominus}\right) \tag{121}$$

One of the quantities that is often measured in experiments is the (isochoric) heat capacity c_V.

$$c_V = \left(\frac{\partial U}{\partial T}\right)_{N,V} \tag{122}$$

In the late 19th century, when statistical thermodynamics was being developed on basis of *classical* mechanics, the behavior of the heat capacity of poly-atomic gases was a mystery. It is easy to see why. The total internal energy of an ideal gas is equal to the sum of the translational, rotational and vibrational energies. We have computed all these quantities in the classical limit:

$$\begin{aligned} U_{trans} &= \frac{3}{2} N k_B T \\ U_{rot} &= \frac{3}{2} N k_B T \quad \text{(for non-linear molecules)} \\ U_{vib} &= N k_B T \qquad \text{(for every vibrational mode)} \end{aligned}$$

The number of vibrational modes of a non-linear molecule containing n atoms is equal to $3n - 6$. Hence, classically, the total energy should be

$$U_{class} = (3n - 3) N k_B T \tag{123}$$

It then follows immediately that, in the classical limit,

$$c_V^{classical} = (3n - 3) N k_B \tag{124}$$

However, this was not at all what was found in experiment. To be more precise: at high temperatures, this expression appeared to become valid, but at low temperatures it failed completely. It appeared as if the vibrations - and sometimes even the rotations - were not contributing to the heat capacity. This problem was resolved with the advent of quantum mechanics. Above we have discussed the correct expression for the average rotational and vibrational energy. For a vibration with frequency ν, the contribution to the internal energy is

$$\begin{aligned} U_{vib} = N < \epsilon_{vib} >= -N \left(\frac{\partial \ln q_{vib}}{\partial \beta} \right) = \\ = Nh\nu/2 + \frac{Nh\nu}{\exp(\beta h\nu) - 1} \end{aligned} \tag{125}$$

At low temperatures, this energy approaches a constant value ($h\nu/2$ per oscillator). Let us write $\beta h\nu \equiv x$. Then the expression for the heat capacity is

$$c_V = Nk_B \frac{x^2 \exp(-x)}{(1 - \exp(-x))^2} \tag{126}$$

At high temperatures ($x \to 0$) c_V approaches Nk_B. But at low temperatures, the heat capacity goes to zero quite steeply. The explanation of this "freezing out" of the vibrational heat capacity was one of the early successes of the merger of quantum mechanics and statistical thermodynamics.

4.2.3 Interacting systems

It is of course nice to be able to compute the thermodynamical properties of ideal gases but most substances that we know are not in a phase that resembles the ideal gas. This means that we should also be able to compute properties of systems consisting of *interacting* molecules. It turns out that this problem is most easily tackled in the classical limit. For convenience, we shall limit the discussion to simple, spherical particles and we shall ignore the internal degrees of freedom of this particles. This is an excellent approximation for atoms, and it is quite reasonable for simple molecules. We then have to consider only the translation partition function. Above, we have already computed this partition function for an ideal gas molecule

$$q_{trans} = V \left(2\pi m k_B T/h^2\right)^{3/2} = \frac{V}{\Lambda^3} \tag{127}$$

Now recall that the molecular partition function was nothing but a measure for the number of quantum states that was accessible to a molecule. It is simply equal to the number of volumes of size Λ^3 that are contained in a volume V. In this example, we have assumed that the ground-state energy of the molecule was zero. Now suppose that the molecule moves in a constant external potential U. In that case, all energy levels are shifted by an amount U and the translational partition function of the molecule becomes

$$q'_{trans} = \frac{V \exp(-\beta U)}{\Lambda^3} \tag{128}$$

or, in terms of the number of accessible quantum states: $n_{acc} = V \exp(-\beta U)/\Lambda^3$. We now assume that the total volume of the system can be divided into small volumes ΔV, such that the external potential may vary from one box to the next, but is constant within a given box. The number of accessible quantum states in the i-th box is $n_{acc}(i) = \Delta V \exp(-\beta U_i)/\Lambda^3$ and the total number of accessible quantum states (i.e. q_{trans}) is equal to

$$\begin{aligned} q_{trans} &= \sum_i \Delta V \exp(-\beta U_i)/\Lambda^3 \\ &\approx \frac{1}{\Lambda^3} \int d\mathbf{r} \;\exp(-\beta U(\mathbf{r})) \end{aligned} \tag{129}$$

Next consider the case that we have N interacting molecules. Every molecule (denoted by i) is located in a volume ΔV centered around a position $\mathbf{r}_i$. ΔV is so small that the total potential energy due to the interaction between the molecules does not depend on *where* inside this volume the molecules are located. Then the number of accessible quantum states for this specific molecular configuration is:

$$n_{states}(\mathbf{r}_1, \mathbf{r}_2, .., \mathbf{r}_N) = \frac{\Delta V^N}{N!\Lambda^{3N}} \exp(-\beta U(\mathbf{r}_1, \mathbf{r}_2, .., \mathbf{r}_N)) \tag{130}$$

the factor $1/N!$ appears again because permuting identical molecules over the different volumes ΔV does not yield a new quantum state. Finally, to compute the total number of accessible quantum states for the entire system, we should sum over all possible positions of the volume elements containing the molecules

$$\begin{aligned} Q_{trans} &= \frac{1}{N!\Lambda^{3N}} \sum_{\mathbf{r}_1} \sum_{\mathbf{r}_2} ... \sum_{\mathbf{r}_N} \Delta V^N \exp(-\beta U(\mathbf{r}_1, \mathbf{r}_2, .., \mathbf{r}_N)) \\ &\approx \frac{1}{N!\Lambda^{3N}} \int d\mathbf{r}_1 \int d\mathbf{r}_2 ... \int d\mathbf{r}_N \;\exp(-\beta U(\mathbf{r}_1, \mathbf{r}_2, .., \mathbf{r}_N)) \end{aligned} \tag{131}$$

This is the (classical) partition function for a system of N interacting molecules. If we could compute it for any substance, we would be able to predict all equilibrium properties of matter. In practice, we can only compute the partition function of a system of interacting particles in a few very special cases (e.g. for a perfect crystal at low temperatures). In general, we need to use numerical simulations to compute the thermodynamic properties of normal liquids or solids. The technicalities of these simulations do not concern us here. The essential fact is that Statistical Thermodynamics tells us exactly *what* to compute.

As an example of a calculation that involves intermolecular interactions, let us consider deviations from the ideal gas law. Experimentally, we know that the pressure of a real gas does not satisfy the ideal-gas relation $PV/Nk_BT = 1$. Rather, we find that, as the number density ρ ($=N/V$) is increased, the deviations from this relation occur

$$\frac{PV}{Nk_BT} = 1 + B_2\rho + B_3\rho^2 + ... \tag{132}$$

where B_2, B_3 etc. are called the second, third etc. *virial coefficients.* The virial coefficients depend on the intermolecular interactions. Here we shall derive an expression for B_2. First, we multiply and divide the translational partition function by V^N. This yields

$$\begin{aligned} Q_{trans} &= \frac{V^N}{N!\Lambda^{3N}} \int \frac{d\mathbf{r}_1}{V} \int \frac{d\mathbf{r}_2}{V} ... \int \frac{d\mathbf{r}_N}{V} \ \exp(-\beta U(\mathbf{r}_1, \mathbf{r}_2, .., \mathbf{r}_N)) \\ &= \frac{V^N}{N!\Lambda^{3N}} \langle \exp(-\beta U) \rangle \end{aligned} \tag{133}$$

where the angular brackets denote the average of the Boltzmann factor $\exp(-\beta U(\mathbf{r}_1, \mathbf{r}_2, .., \mathbf{r}_N))$ over all possible positions $\{\mathbf{r}_1, \mathbf{r}_2, .., \mathbf{r}_N\}$ within the volume V. At extremely low densities, the molecules are almost always too far apart to interact and hence the average Boltzmann factor is simply equal to one. At higher densities, we shall notice the effect of interactions between molecules. Let us assume that the interactions between molecules are such that a molecule must be within a certain distance r_c in order to experience the potential of the other molecules. Or, phrased in another way, if there are no molecules within a volume $v_c = (4\pi/3)r_c^3$ of a given molecule, then that molecule does not contribute to the interaction energy. Let us denote by P_0 the probability that no two molecules are within a distance r_c. At very low densities, we can write the average of the Boltzmann factor as

$$\langle \exp(-\beta U) \rangle = P_0 \times 1 + P_1 < \exp(-\beta U) >_{pair} \tag{134}$$

where P_1 denotes the probability that there is exactly one pair of molecules at a distance less than r_c. Because the density is very low, we can ignore the probability that there will be more than two molecules at a distance less than r_c. In other words, we have either *no* molecules that are interacting (probability P_0) or just *one pair* (probability P_1). Clearly, $P_0 + P_1 = 1$ and hence $P_0 = 1 - P_1$. Now we should still compute the average Boltzmann factor for a pair of molecules at a distance less than r_c. It is

$$< \exp(-\beta U) >_{pair} = \frac{1}{v_c} \int_{v_c} d\mathbf{r} \exp(-\beta u(\mathbf{r})) \tag{135}$$

where $u(\mathbf{r})$ is the potential energy of interaction of a *pair* of molecules at a distance $\mathbf{r}$. We can now write

$$\begin{aligned} \langle \exp(-\beta U) \rangle &= P_0 + P_1 \frac{1}{v_c} \int_{v_c} d\mathbf{r} \exp(-\beta u(\mathbf{r})) \\ &= 1 - P_1 + P_1 \frac{1}{v_c} \int_{v_c} d\mathbf{r} \exp(-\beta u(\mathbf{r})) \\ &= 1 - P_1 \frac{1}{v_c} \int_{v_c} d\mathbf{r} + P_1 \frac{1}{v_c} \int_{v_c} d\mathbf{r} \exp(-\beta u(\mathbf{r})) \\ &= 1 + P_1 \frac{1}{v_c} \int_{v_c} d\mathbf{r} [\exp(-\beta u(\mathbf{r})) - 1] \end{aligned} \tag{136}$$

As we have assumed that the intermolecular interaction vanishes outside v_c (and hence $\exp(-\beta u(r)) - 1 = 0$), we need not limit the integration to the volume v_c but can extend it over all space. Now we should still compute P_1, the probability that there is a single pair of (randomly distributed) molecules within the same volume v_c. At low densities, the probability that there is another molecule in a volume around a *given* molecule is simply equal to ρv_c (where ρ is the number density of the molecules). As there are N molecules in the system, and we could have taken any of these molecules as our "central" molecules, the probability to find a pair is $N/2$ times larger (the factor $1/2$ comes in to avoid double counting). Hence, at low densities, $P_1 = N\rho v_c/2$ and hence

$$\langle \exp(-\beta U) \rangle = 1 + \frac{N\rho}{2} \int d\mathbf{r} [\exp(-\beta u(\mathbf{r})) - 1] \tag{137}$$

With this result, we can write

$$Q_{trans} \approx \frac{V^N}{N!\Lambda^{3N}} \left(1 + \frac{N\rho}{2} \int d\mathbf{r} [\exp(-\beta u(\mathbf{r})) - 1] \right) \tag{138}$$

The pressure P is given by

$$\begin{aligned} \frac{P}{k_B T} &= \frac{\partial \ln Q_{trans}}{\partial V} \\ &\approx \frac{N}{V} - \frac{\frac{\rho^2}{2} \int d\mathbf{r} [\exp(-\beta u(\mathbf{r})) - 1]}{1 + \frac{N\rho}{2} \int d\mathbf{r} [\exp(-\beta u(\mathbf{r})) - 1]} \\ &\approx \rho + \frac{\rho^2}{2} \int d\mathbf{r} [1 - \exp(-\beta u(\mathbf{r}))] \end{aligned} \tag{139}$$

where, in the third line, we have used the fact that, at sufficiently low densities, $\frac{N\rho}{2}\int d\mathbf{r}[\exp(-\beta u(\mathbf{r}))-1]\ll 1$. If we compare this expression with the virial series (Eqn.132), we find that the second virial coefficient is equal to

$$B_2=\frac{1}{2}\int d\mathbf{r}[1-\exp(-\beta u(\mathbf{r}))] \tag{140}$$

Again, this is a very important result because it shows that a measurement of the second virial coefficient provides information about the intermolecular interactions. To give a specific example: assume that molecules are hard spheres with a diameter σ. For $r>\sigma$, $u(r)=0$ and hence $\exp(-\beta u(r))=1$. For $r<\sigma$, $u(r)=\infty$ and hence $\exp(-\beta u(r))=0$. Therefore,

$$B_2^{HS}=\frac{1}{2}\int_0^\sigma 4\pi r^2 dr=\frac{2\pi\sigma^3}{3} \tag{141}$$

Of course, real molecules do not only repel each other at short distances, they also attract at larger distances. A very simple model potential that exhibits both features is the so-called *square-well potential.* The square-well potential is equal to the hard-sphere potential for $r<\sigma$. But for $\sigma<r<\lambda\sigma$ (with $\lambda>1$). the square well potential is attractive:

$$u_{sw}(r)=-\epsilon \quad (\text{for } \sigma<r<\lambda\sigma) \tag{142}$$

and

$$u^{SW}(r)=0 \ (\text{for } r>\lambda\sigma) \tag{143}$$

We can easily compute the second virial coefficient for this model potential. It is

$$\begin{aligned} B_2^{SW} &= \frac{1}{2}\left(\int_0^\sigma 4\pi r^2 dr-(\exp(\beta\epsilon)-1)\int_\sigma^{\lambda\sigma}4\pi r^2 dr\right)\\ &= \frac{2\pi\sigma^3}{3}\left(1-(\exp(\beta\epsilon)-1)(\lambda^3-1)\right) \end{aligned} \tag{144}$$

At very high temperatures ($\beta\to 0$), B_2^{SW} is equal to the hard-sphere second virial coefficient. However, at low temperatures, the term with $\exp(\beta\epsilon)$ dominates, and B_2 becomes large and negative. The point where B_2 changes sign is called the Boyle temperature. It follows from the equation:

$$\left(1-(\exp(\beta\epsilon)-1)(\lambda^3-1)\right)=0 \tag{145}$$

which yields

$$\exp(\beta\epsilon)=\frac{\lambda^3}{\lambda^3-1} \tag{146}$$

or

$$\frac{k_B T}{\epsilon} = \frac{1}{\ln[\lambda^3/(\lambda^3 - 1)]} \tag{147}$$

If we keep ϵ fixed and vary λ, we see that for large λ (long-ranged attraction), $k_B T/\epsilon$ is large (high Boyle temperature), while for $\lambda \to 1$, $k_B T/\epsilon \to 0$ (low Boyle temperature).

It should be stressed that only the lowest few virial coefficients can easily be computed. Hence, when we are interested in the behavior of dense systems (liquids, solids), the properties of such systems cannot be computed analytically from Eqn.131. The standard approach to resolve this problem is to use numerical simulations instead. With the help of computer simulations, it is possible to compute the thermodynamical properties of arbitrarily complex (macro)molecular systems provided that: 1.) we can use a "classical" description (i.e. Eqn.131 is valid) and 2.) we have a good knowledge of the potential energy function $U(\mathbf{r}_1, \mathbf{r}_2, .., \mathbf{r}_N)$.

5 Chemical Equilibria

One of the important results of classical thermodynamics was that it allowed us to write down the conditions for equilibrium. It starts from the condition that, at constant temperature and pressure, the Gibbs free energy G is at a minimum in equilibrium. A variation of the Gibbs free energy can be written as

$$dG = -SdT + VdP + \sum_\alpha \mu_\alpha dn_\alpha \tag{148}$$

where n_α denotes the number of moles of species α and μ_α its chemical potential. At constant pressure and temperature, dP and dT are zero, and hence the condition for equilibrium is

$$\sum_\alpha \mu_\alpha dn_\alpha = 0 \tag{149}$$

In a chemical reaction, the variations dn_α are related through the stoichiometry of the reaction. This is usually denoted by

$$dn_\alpha = \nu_\alpha d\xi \tag{150}$$

where ξ measures the progress of the reaction and ν_α is the stoichiometric coefficient of species α. Hence, the condition for equilibrium is

$$\sum_\alpha \mu_\alpha \nu_\alpha d\xi = 0 \tag{151}$$

or, as $d\xi$ is arbitrary

$$\sum_\alpha \mu_\alpha \nu_\alpha = 0 \tag{152}$$

Using Eqn.118 for the chemical potential of species α

$$\begin{aligned}\mu_\alpha &= k_BT\ln(\rho_\alpha \Lambda_\alpha^3) - k_BT\ln q_{internal}^\alpha \\ &= k_BT\ln(\frac{p_\alpha}{k_BT}\Lambda_\alpha^3) - k_BT\ln q_{internal}^\alpha \end{aligned}\tag{153}$$

we find that

$$\sum_\alpha \nu_\alpha \left(\ln(\frac{p_\alpha/p^\ominus}{k_BT/p^\ominus}\Lambda_\alpha^3) - \ln q_{internal}^\alpha\right) = 0 \tag{154}$$

where $p^\ominus$ is the pressure of the reference state (e.g. one bar). It then follows that

$$\sum_\alpha \nu_\alpha \ln(p_\alpha/p^\ominus) = \sum_\alpha \nu_\alpha \left(\ln(\frac{k_BT}{p^\ominus \Lambda_\alpha^3}) + \ln q_{internal}^\alpha\right) \tag{155}$$

But

$$\sum_\alpha \nu_\alpha \ln(p_\alpha/p^\ominus) = \ln(\prod_\alpha (p_\alpha/p^\ominus)^{\nu_\alpha}) = \ln K \tag{156}$$

where K is the equilibrium constant of the reaction. Hence we have now derived a "molecular" expression for the equilibrium constant of a chemical reaction of (ideal) gases

$$\ln K = \sum_\alpha \nu_\alpha \ln(\frac{k_BT}{p^\ominus \Lambda_\alpha^3} q_{internal}^\alpha) \tag{157}$$

Before we can actually use this expression, we should note one important point. Until now, we have, somewhat arbitrarily chosen the ground state of the molecule as our zero of energy. However, when we compare the relative stability of different molecules, the difference in energy between the molecular ground states become very important. We can account for this part of the internal energy by writing

$$q_{internal}^\alpha = \exp(-\beta\epsilon_0^\alpha) q_{internal}^{\alpha,0} \tag{158}$$

where $q_{internal}^{\alpha,0}$ is the internal molecular partition function that is obtained when the zero of energy coincides with the molecular ground state. When we introduce the notation

$$\Delta\epsilon_0 = \sum_\alpha \nu_\alpha \epsilon_0^\alpha \tag{159}$$

we obtain

$$\ln K = -\frac{\Delta\epsilon_0}{k_B T} + \ln \prod_\alpha \left(\frac{k_B T}{p^\ominus \Lambda_\alpha^3} q_{internal}^{\alpha,0} \right)^{\nu_\alpha} \tag{160}$$

Remember that we can determine all terms on the right-hand side of this equation either from quantum-mechanical calculations or from spectroscopic measurements. Hence, we have now achieved one of the prime objectives of Statistical Thermodynamics: we have obtained an expression that allows us to *predict* the equilibrium constant of a gas reaction. Take for instance the reaction $A_2 + B_2 \rightleftharpoons 2AB$. For this reaction, the equilibrium constant is given by

$$\ln K = -\frac{\Delta\epsilon_0}{k_B T} + \ln \frac{\left(\frac{q_{internal}^{AB,0}}{\Lambda_{AB}^3} \right)^2}{\left(\frac{q_{internal}^{AA,0}}{\Lambda_{AA}^3} \right) \left(\frac{q_{internal}^{BB,0}}{\Lambda_{BB}^3} \right)} \tag{161}$$

Using

$$\Lambda^3 = \left(h^2 / 2\pi m k_B T \right)^{3/2} \tag{162}$$

$$q_{rot} = \frac{2 I k_B T}{\sigma \hbar^2} \tag{163}$$

$$q_{vib}^i = \frac{1}{1 - \exp(-\beta h \nu_i)} \tag{164}$$

we can work out the individual (translational, rotational and vibrational) contributions to the equilibrium constant.

5.0.4 Example

To give an example, consider the reaction

$$H_2 + D_2 \rightleftharpoons 2HD \tag{165}$$

As H and D are isotopes, it is a reasonable approximation to assume that all three molecules have the same bondlength and that they have the same vibrational force constant, k_{vib}. The molecular ground-states are, however, not the same, because the zero-point energy ($h\nu/2$) of the molecular vibration is different. Hence,

$$\Delta\epsilon_0 = \frac{1}{2}(2h\nu_{HD} - h\nu_{H_2} - h\nu_{D_2}) \tag{166}$$

The individual vibration frequencies can be computed using

$$v_{HD} = \frac{1}{2\pi} \sqrt{\frac{k_{vib}[m_H + m_D]}{m_H m_D}} \tag{167}$$

and similar expressions for ν_{H_2} and ν_{D_2}. Let us compute the equilibrium constant in the high-temperature limit.

$$\begin{aligned}\ln K = & -\frac{\Delta\epsilon_0}{k_BT} + \ln\frac{\Lambda_{H_2}^3\Lambda_{D_2}^3}{\Lambda_{HD}^6} \\ & + \ln\frac{(q_{HD}^{rot})^2}{q_{H_2}^{rot}q_{D_2}^{rot}} \\ & + \ln\frac{(q_{HD}^{vib})^2}{q_{H_2}^{vib}q_{D_2}^{vib}}\end{aligned} \tag{168}$$

Let us compute the equilibrium constant in the high-temperature limit where $q_{rot} = 2Ik_BT/(\sigma\hbar^2)$ and $q_{vib}^i = kT/h\nu_i$. The translational contribution is

$$\begin{aligned}\ln\frac{\Lambda_{H_2}^3\Lambda_{D_2}^3}{\Lambda_{HD}^6} & = \ln\frac{m_{HD}^3}{m_{H_2}^{3/2}m_{D_2}^{3/2}} \\ & = \ln\frac{3^3}{2^{3/2}4^{3/2}} \\ & = \ln\frac{9^{3/2}}{2^{3/2}4^{3/2}} \\ & = \frac{3}{2}\ln\frac{9}{8}\end{aligned} \tag{169}$$

The rotational contribution is

$$\begin{aligned}\ln\frac{(q_{HD}^{rot})^2}{q_{H_2}^{rot}q_{D_2}^{rot}} & \approx \ln\frac{I_{HD}^2/\sigma_{HD}^2}{(I_{H_2}/\sigma_{H_2})(I_{D_2}/\sigma_{D_2})} \\ & = \ln\frac{(2/3)^2}{(1/4)(1/2)} = \ln\frac{32}{9}\end{aligned} \tag{170}$$

where we have used $\sigma_{H_2} = \sigma_{D_2} = 2$ and $\sigma_{HD} = 1$. And, in addition,

$$I = \mu r_e^2 \tag{171}$$

with $r_e(H_2) \approx r_e(D_2) \approx r_e(HD)$, while $\mu_{H_2} = \frac{1}{2}m_H$, $\mu_{D_2} = m_H$ and $\mu_{HD} = \frac{2}{3}\ m_H$. The vibrational contribution is

$$\begin{aligned}\ln\frac{(q_{HD}^{vib})^2}{q_{H_2}^{vib}q_{D_2}^{vib}} & \approx \ln\frac{\nu_{H_2}\nu_{D_2}}{\nu_{HD}^2} \\ & - \frac{1}{2}\ln\frac{9}{8}\end{aligned} \tag{172}$$

and the total equilibrium constant is

$$\ln K = -\frac{\Delta\epsilon_0}{k_BT} + \ln 4 \tag{173}$$

At very high temperatures, the first term becomes small, and the equilibrium constant approaches 4. This seems surprising: why should the equilibrium be shifted towards the HD? The reason is that, due to the symmetry of H_2 and D_2, the number of accessible (rotational states) of these molecules is divided by two, and hence the product is divided by four. And, after all, the equilibrium constant simply measures the ratio of the number of accessible states of the products and reactants.

6 Macromolecules

Small molecules such as H_2 and N_2 are very interesting, but most of (bio)chemistry deals with molecules that are much more complex and very much larger. Examples are polymers, proteins, DNA etc... For such complex molecules it is usually very difficult, if not outright impossible, to compute the molecular partition function directly from quantum mechanics. It would then seem that we cannot apply the framework of Statistical Thermodynamics directly to such macromolecular systems. Although there is some truth to this statement - it is for instance very difficult to predict the properties of proteins in solution - there are many examples where the fact that the molecule is very large, actually simplifies matters. An important example of “simple” behavior of complex molecules is to be found in polymers. Polymers are long chain molecules. Such molecules tend to be very flexible - so much so that it is impossible to speak of *the* conformation of the molecule. A polymer in solution is very disordered: it resembles a random coil. It would seem that this great disorder would make the description of polymers difficult, but the converse is true: the random coil provides the simplest model of a polymers and, as we shall see below, it has several non-trivial properties that can be observed experimentally.

6.1 Ideal polymers

Let us consider a polymer that consists of N segments of length l. When this polymer is fully stretched, its total length is equal to Nl. Now consider what happens if the polymer is fully flexible - that is, every segment can rotate freely around the point at which it is connected to the previous segment. Clearly, this is an oversimplification. However, in practice, many polymers behave like such freely jointed chain (although the length l is larger than the size of a monomer). In addition, we assume that different segments do not interact - this assumption is similar to the one we make when we say that, to a first approximation, a dilute gas can be described as an ideal gas where the molecules do not interact. The difference in the case of ideal polymers is that, although different segments do not interact, adjacent segments *are* connected. Clearly, as an ideal polymer can coil up, its size will be less than its contour length Nl. In fact, it turns out that we can easily compute the average size of an ideal polymer. In particular, we shall consider the end-to-end distance of an ideal polymer. First of all, we note that every segment of length l has a direction. It is therefore convenient to denote the distance between the beginning and end-point of the i-th segment of the polymer by a vector $\mathbf{l}_i$. The total (vector) distance between one end of the polymer and the other end is

$$\mathbf{R}_{ee} = \sum_{i=1}^{N} \mathbf{l}_i \tag{174}$$

What is the *average* end-to-end distance ? To compute this, we should average over all possible orientations of the segments

$$< \mathbf{R}_{ee} >=< \sum_{i=1}^{N} \mathbf{l}_i >= \sum_{i=1}^{N} < \mathbf{l}_i > \tag{175}$$

However, all segments can rotate freely. That means that all orientations are equally likely. In particular, a segment is just as likely to point in the direction $+\mathbf{l}_i$ as in the direction $-\mathbf{l}_i$. Hence the average $< \mathbf{l}_i >= 0$. And therefore $< \mathbf{R}_{ee} >= 0$. This does *not* mean that the size of a polymer is vanishingly small, but simply that the polymer has no preferred orientation. A better measure for the size of a coiled ideal polymer is the *mean-squared* end-to-end distance $< \mathbf{R}_{ee}^2 >$.

$$\begin{aligned} < \mathbf{R}_{ee}^2 > &= \left\langle \left(\sum_{i=1}^{N} \mathbf{l}_i \right)^2 \right\rangle \\ &= \left\langle \left(\sum_{i=1}^{N} \mathbf{l}_i \right) \cdot \left(\sum_{j=1}^{N} \mathbf{l}_j \right) \right\rangle \\ &= \left\langle \sum_{i=1,j=1}^{N} \mathbf{l}_i \cdot \mathbf{l}_j \right\rangle = \sum_{i=1,j=1}^{N} < \mathbf{l}_i \cdot \mathbf{l}_j > \\ &= \sum_{i=1,j=1}^{N} l^2 < \cos\theta_{ij} > \end{aligned} \tag{176}$$

where θ_{ij} is the angle between segments i and j. Now recall that every segment is free to rotate. That means that the orientation of different segments i and j are uncorrelated and hence $< \cos\theta_{ij} >= 0$, for all $i \neq j$. However, for $i = j$, $\theta_{ii} = 0$ and hence $\cos\theta_{ii} = 1$. As a consequence, all the cross-terms with $i \neq j$ drop out of Eqn.176 and we find

$$< \mathbf{R}_{ee}^2 >= \sum_{i=1}^{N} l^2 = Nl^2 \tag{177}$$

Hence, the root mean-squared end-to-end distance is

$$\sqrt{<\mathbf{R}_{ee}^2>} = l\sqrt{N} \tag{178}$$

When we compare this with the contour length R_c of the polymer (i.e. the length it would have if it were fully stretched), we find that

$$\frac{\sqrt{<\mathbf{R}_{ee}^2>}}{R_c} = \frac{l\sqrt{N}}{lN} = \frac{1}{\sqrt{N}} \tag{179}$$

As N is usually quite large (ranging from hundreds to millions), the typical size of a coiled polymer is much less than that of a stretched polymer.

We can do better than simply calculating the mean-squared end-to-end distance of a polymer: we can actually compute the probability to find any given value for $\mathbf{R}_{ee}$. However, to simplify matters, we will only consider the one-dimensional case. Again, we consider a polymer consisting of N segments, but now the segments can only point either to the right or to the left. Let us denote the number of segments pointing to the right by N_R and the number pointing to the left by N_L. Let us denote the difference $N_R - N_L$ by n. Clearly, the end-to-end distance is equal to $\Delta x = nl$. We can compute the probability to find the chain in a conformation with a given value of $N_R - N_L$. First, consider the number of ways in which we can have N_R segments point to the right, and N_L segments point to the left. This number is

$$\begin{aligned}\Omega(N_R, N_L) &= \frac{N!}{N_R! N_L!} \\ &= \frac{N!}{\{\frac{1}{2}(N+n)\}!\{\frac{1}{2}(N-n)\}!}\end{aligned} \tag{180}$$

where we have used the fact that $N_R = (N+n)/2$ and $N_L = (N-n)/2$. The total number of possible conformations is simply 2^N (because every segment has two possible orientations, and there are N segments). The probability to have a conformation with $N_R - N_L = n$ is then

$$P(\frac{1}{2}(N+n), \frac{1}{2}(N-n)) = \frac{N!}{\{\frac{1}{2}(N+n)\}!\{\frac{1}{2}(N-n)\}!}\left(\frac{1}{2}\right)^N \tag{181}$$

or

$$\ln P(\frac{1}{2}(N+n), \frac{1}{2}(N-n)) = \ln N! - \ln\{\frac{1}{2}(N+n)\}! - \ln\{\frac{1}{2}(N-n)\}! - N\ln 2 \tag{182}$$

We will consider the case that $N \gg 1$, and $n \ll N$. In that case, we can use the Stirling approximation ($\ln x! \approx x \ln x - x + \ln \sqrt{2\pi x}$) for all factorials in the above equation. Then (see Appendix):

$$\begin{aligned}\ln P(\frac{1}{2}(N+n), \frac{1}{2}(N-n)) = {} & N \ln N + \ln \sqrt{2\pi N} \\ & - \ln \sqrt{\pi(N+n)} - \ln \sqrt{\pi(N-n)} \\ & - \frac{1}{2}(N+n)\ln(N+n) - \frac{1}{2}(N-n)\ln(N-n)\end{aligned} \tag{183}$$

This can be rewritten as

$$\begin{aligned}\ln P(\frac{1}{2}(N+n), \frac{1}{2}(N-n)) = {} & -\frac{1}{2} N \ln(1 - \frac{n^2}{N^2}) + \ln \sqrt{2\pi N} \\ & - \ln \sqrt{\pi(N+n)} - \ln \sqrt{\pi(N-n)} \\ & - \frac{1}{2} n \ln(1 + \frac{n}{N}) + \frac{1}{2} n \ln(1 - \frac{n}{N})\end{aligned} \tag{184}$$

We now make use of the fact that $n/N \ll 1$, and that $\ln(1 \pm x) \approx \pm x$ for $x \ll 1$. We then find

$$\ln P(\frac{1}{2}(N+n), \frac{1}{2}(N-n)) \approx -\frac{n^2}{2N} - \ln \sqrt{\pi N/2} \tag{185}$$

From which it follows that

$$P(\frac{1}{2}(N+n), \frac{1}{2}(N-n)) = \frac{1}{\sqrt{\pi N/2}} \exp(-n^2/2N) \tag{186}$$

However, in this form, the above expression is rarely used. We should recall that we started by considering a polymer of length N of which N_R segments were pointing to the right. In this expression, N_R can take the values $\{0, 1, 2, \cdots, N\}$. We then converted to the variable $n \equiv N_R - N_L$. n can take the values $\{-N, -N+2, \cdots, -4, -2, 0, 2, 4, \cdots, N-2, N\}$ (where we have assumed that N is even. For N is odd, the values would be $\{-N, -N+2, \cdots, -3, -1, 1, 3, \cdots, N-2, N\}$). The point to note is that when N_R changes by one, n changes by two. In the limit of large N we are usually not interested in the question whether n has a *specific value* (say 3), but in the probability to find n in a given interval Δn. When we compute this probability per unit interval, we have to take into account that $P(\frac{1}{2}(N+n), \frac{1}{2}(N-n))$ is given by Eqn.186 for all *even* (c.q. *odd*) values of n and *zero* otherwise. The probability per *unit* interval is then equal to the average

$$\begin{aligned}
P(n) &= (P(\frac{1}{2}(N+n),\frac{1}{2}(N-n))_{n=odd} + P(\frac{1}{2}(N+n),\frac{1}{2}(N-n))_{n=even})/2 \\
&= \frac{\left(0+\frac{1}{\sqrt{\pi N/2}}\exp(-n^2/2N)\right)}{2} \\
&= \frac{1}{\sqrt{2\pi N}}\exp(-n^2/2N) \qquad (187)
\end{aligned}$$

This distribution is then properly normalized, in the sense that

$$\int_{-\infty}^{\infty} P(n)dn = 1 \qquad (188)$$

It is the distribution $P(n)$ that is most commonly used when discussing the distribution of end-to-end distances in polymers. In summary: the distribution of end-to-end distances of an ideal polymer is a Gaussian. From the width of the Gaussian, we can deduce that the mean-squared end-to-end distance of the one dimensional chain is

$$< X_{ee}^2 >= l^2 < n^2 >= Nl^2 \qquad (189)$$

Finally, we can make a link with thermodynamics. We can compute the change in entropy involved in stretching a polymer. We start with Eqn.180 for the number of conformations of a polymer with end-to-end distance $(N_R - N_L)l$. Using the relation between entropy and the number of states of a system, we can write

$$S(n) = k_B \ln \Omega(N_R, N_L) \qquad (190)$$

and, proceeding as above, we find that

$$S(n) = -\frac{n^2 k_B}{2N} + \text{constant} \qquad (191)$$

where the constant accounts for all terms that do not depend on n. Let us denote the end-to-end elongation of the polymer by $X_{ee} = nl$. Then

$$S(x) = -\frac{X_{ee}^2 k_B}{2Nl^2} + \text{constant} \qquad (192)$$

Now, just as for an ideal gas, the internal energy U of an *ideal* polymer,depends only on the temperature. If we use

$$dU = w + q \qquad (193)$$

with $q = TdS$ and $w = FdX_{ee}$, where F is the restoring force exerted by the polymer when stretched, then we have

$$dU = TdS + FdX_{ee} \tag{194}$$

or, at constant temperature,

$$\left(\frac{\partial U}{\partial X_{ee}}\right)_T = 0 = T\left(\frac{\partial S}{\partial X_{ee}}\right)_T + F \tag{195}$$

and hence

$$F = -T\left(\frac{\partial S}{\partial X_{ee}}\right)_T \tag{196}$$

If we use Eqn.192, we find that

$$F = \frac{X_{ee}k_BT}{Nl^2} \tag{197}$$

Hence, the restoring force is linear in the deformation (just as for a normal spring) but, unlike a metal spring, the spring-constant is proportional to the temperature T. The reason is that the restoring force of an ideal polymer is purely entropic. Hence, if you pull a rubber band, the restoring force that you feel is entropic in origin!

One final interesting observation: we can also predict what will happen if we stretch a polymer (or, in practice, a rubber band) adiabatically - i.e. without exchanging heat with the environment. Using the thermodynamic relation

$$\left(\frac{\partial T}{\partial X_{ee}}\right)_S \left(\frac{\partial X_{ee}}{\partial S}\right)_T \left(\frac{\partial S}{\partial T}\right)_{X_{ee}} = -1 \tag{198}$$

we can get an expression for the change in temperature due to the adiabatic stretching of the polymer

$$\left(\frac{\partial T}{\partial X_{ee}}\right)_S = -\frac{\left(\frac{\partial S}{\partial X_{ee}}\right)_T}{\left(\frac{\partial S}{\partial T}\right)_{X_{ee}}} \tag{199}$$

or, using the relation $\left(\frac{\partial S}{\partial T}\right)_{X_{ee}} = c_P/T$, we get

$$\left(\frac{\partial T}{\partial X_{ee}}\right)_S = \frac{X_{ee}k_BT}{c_PNl^2} \tag{200}$$

As the heat capacity c_P is always positive, it follows that the temperature *increases* when the polymer material is stretched adiabatically (or *decreases* when the force on a polymer material is released). Try it!

Note that we can combine Eqn.197 with Eqn.189 to obtain

$$F = \frac{X_{ee} k_B T}{< X_{ee}^2 >} \tag{201}$$

In other words, the "spring-constant" κ of an ideal polymer is equal to

$$\kappa = \frac{k_B T}{< X_{ee}^2 >} \tag{202}$$

In the Appendix, we show that this result is also valid for real (i.e. non-ideal) polymers.

7 Diffusion

7.1 Fick's first law

Thus far, we have only considered systems in equilibrium. Let us now look at a simple example of molecular transport. Suppose that we have a solution of molecules with concentration c. If an external force F is acting on this molecules then the molecules will start to move in the direction of the force. When the molecules move with a drift velocity $< v >$ with respect to the solvent, then they will experience a friction force

$$F_{frict} = -f < v > \tag{203}$$

where f is the friction constant. In steady state, the friction force F_{frict} on the molecule exactly balances the external force F_{ext} and hence

$$F_{ext} = f < v > \tag{204}$$

The net flux of molecules J_{drift} (i.e. the number of molecules that pass per second through a unit area) is equal to the concentration of the molecules c times their average velocity

$$J_{drift} = c < v >= c\frac{F_{ext}}{f} \tag{205}$$

Now consider a seemingly unrelated situation: a system in which there is a concentration gradient. This gradient will give rise to a diffusion flux that counteracts the concentration gradient. The expression for the diffusion flux is given by Fick's first law

$$J_{diff} = -D\frac{\partial c}{\partial x} \tag{206}$$

where D is the diffusion coefficient of the molecules. Now consider what will happen if we apply an external force to molecules in solution. The molecules will start to move in the direction of the force. But, as more and more molecules are transported, a concentration gradient will build up. This concentration gradient will give rise to a diffusion current *in the opposite direction*. Eventually, the diffusion current will exactly balance the flux due to the external force. When this happens, the system has reached equilibrium (no more fluxes). But we know that, in equilibrium, the concentration of the molecules is given by the Boltzmann distribution

$$c(x) = c(0)\exp(-\beta(U(x) - U(0)) \tag{207}$$

where $U(x)$ is the (external) potential energy of the molecules at position x (we assume that the solution is sufficiently dilute that we can ignore the interactions between the molecules). Hence

$$\begin{aligned} J_{diff} &= -D\frac{\partial}{\partial x}[c(0)\exp(-\beta(U(x) - U(0))] \\ &= D\beta\frac{\partial U(x)}{\partial x}c(x) \end{aligned} \tag{208}$$

This flux is equal and opposite to the flux due to the external force F

$$J_{drift} = c(x)\frac{F_{ext}}{f} = -\frac{c(x)}{f}\frac{\partial U(x)}{\partial x} \tag{209}$$

where we have used the fact that

$$F_{ext} = -\frac{\partial U(x)}{\partial x} \tag{210}$$

In equilibrium

$$J_{drift} + J_{diff} = 0 \tag{211}$$

or

$$-\frac{c(x)}{f}\left(\frac{\partial U(x)}{\partial x}\right) = -D\beta c(x)\left(\frac{\partial U(x)}{\partial x}\right) \tag{212}$$

This shows that the friction coefficient f is related to the diffusion coefficient D by

$$f = \frac{kT}{D} \tag{213}$$

This relation was first derived by Einstein.

7.2 Fick's second law

Let us consider a system in which there is a concentration gradient. We assume that no external forces act on the system. Due to diffusion, the local concentration will change until it is constant throughout the system. How does the local concentration $c(x,t)$ change in time? Let us first consider how the number of particles in a thin slab of fluid between x and $x + \Delta x$ changes with time. This slab is bounded by two surfaces S at x and at $x + \Delta x$. The number of molecules that pass per second at point x through the surface into the slab, is equal to $J(x,t)S$. The number of molecules that leave the slab per second at point $x + \Delta x$ is equal to $J(x + \Delta x, t)S$. The rate of change of the number of molecules in the slab is then equal to

$$\frac{\partial N(x,t)}{\partial t} = S(J(x,t) - J(x + \Delta x, t)) \tag{214}$$

If we now consider the limit $\Delta x \Rightarrow 0$, we can write

$$\frac{\partial N(x,t)}{\partial t} = -S\Delta x \frac{\partial J(x,t)}{\partial x} \tag{215}$$

But $S\Delta x$ is imply the volume of the slab and the concentration $c(x,t)$ is equal to $N(x,t)/(S\Delta x)$. Hence,

$$\frac{\partial c(x,t)}{\partial t} = -\frac{\partial J(x,t)}{\partial x} \tag{216}$$

This equation simply expresses the conservation of the total number of dissolved particles. If we now combine the conservation law (Eqn.216) with Fick's first law (Eqn.206), we get

$$\frac{\partial c(x,t)}{\partial t} = \frac{\partial}{\partial x}\left(D\frac{\partial c(x,t)}{\partial x}\right) \tag{217}$$

and if the diffusion constant does not depend on x, we find

$$\frac{\partial c(x,t)}{\partial t} = D\frac{\partial^2 c(x,t)}{\partial x^2} \tag{218}$$

This is Fick's second law: it describes how the local concentration changes in time due to diffusion. The solution of this equation depends on the boundary conditions. Let us consider the special case that initially (i.e. at $t = 0$) the concentration is sharply peaked at $x = 0$. In that case, the solution of Eqn.218 is

$$c(x,t) = \frac{n_0}{\sqrt{4\pi D t}} \exp(-x^2/4Dt) \tag{219}$$

where n_0 is the total number of dissolved molecules. Note that the concentration profile is a Gaussian. The width of the Gaussian is equal to $2Dt$. That is: for short time, the Gaussian is very narrow (the concentration is highly localized) but for longer times it broadens. Using

$$\int_{-\infty}^{\infty} dx \exp(-ax^2) = \sqrt{\frac{\pi}{a}} \tag{220}$$

it is easy to show that the integral

$$\int_{-\infty}^{\infty} dx c(x,t) = n_0 \tag{221}$$

In other words, the total number of dissolved molecules is conserved - as it should.

How far does a molecule travel from the origin in time t? As the Gaussian distribution is an even function of x, the average displacement $< x(t) >= 0$. However, the mean-squared displacement is not equal to zero. In fact, it is[2]

$$\begin{aligned} < x^2(t) > &= \frac{\int dx c(x,t) x^2}{\int dx c(x,t)} \\ &= \int dx \frac{x^2}{\sqrt{4\pi Dt}} \exp(-x^2/4Dt) \\ &= 2Dt \end{aligned} \tag{222}$$

and the root mean-squared displacement $\sqrt{< x^2 >} = \sqrt{2Dt}$.

7.3 Relation to ideal polymers

The distribution of end-to-end distances of a polymer of N_s segments with length l_s is given by a Gaussian distribution, and so is the distribution of distances traveled in a time interval t by a diffusing molecule. This analogy is more than superficial. In the case of the polymer, the Gaussian distribution followed when we considered the number of ways in which an ideal (one-dimensional) polymer could have N_R segments pointing to the right and N_L segments pointing to the left, such that $(N_R - N_L)l_s = x$. Diffusion, on the other hand, is due to the random (Brownian) motion of molecules. Let us describe the motion of such a molecule as a sequence of jumps with a characteristic length l_j (in a gas, this length l_j would be of the order of the mean free path of a molecule between collisions). Let us denote by Γ the average number of jumps that a molecule makes per second. The total number of jumps that a molecule makes in time t is then $N_{jump} = \Gamma t$. Any jump is equally likely to be to the right or to the left. The probability that a molecule has made N_R jumps to the right and N_L jumps to the left is

$$P(N_R, N_L) = \frac{N_{jump}!}{N_R! N_L!} \left(\frac{1}{2}\right)^{N_{jump}} \tag{223}$$

---

[2] This is easy to show, using

$$\int_{-\infty}^{\infty} dx \exp(-ax^2) x^2 = -\frac{d}{da} \int_{-\infty}^{\infty} dx \exp(-ax^2)$$

The distance that such a molecule will have traveled is $x = (N_R - N_L)l_j$. For $x \ll N_{jump} l_j$ we can show (using the same techniques as for the ideal polymers - Eqn.188) that

$$P(x,t) = \frac{1}{\sqrt{2\pi\Gamma l_j^2 t}} \exp(-x^2/(2\Gamma l_j^2 t)) \tag{224}$$

Clearly, the description of the random walk of a diffusing molecule is completely equivalent to the description of the random conformations of an ideal polymer if we make the following identifications:

$$N_{jump} = \Gamma t = N_s \tag{225}$$

$$l_j = l_s \tag{226}$$

and

$$2D = \Gamma l_j^2 \tag{227}$$

As Richard Feynman has said: *The same equations have the same solutions...*

A Appendix: Essential mathematics.

Elementary Statistical Thermodynamics does not make use of any sophisticated mathematics. Here we briefly review the mathematics that is most frequently used. Below, we give neither a proper derivation nor a proof of any of the results that we quote. However, in some cases we do provide a non-rigorous "justification". We assume that the reader is familiar with the most common functions and algebraic manipulations.

- **Properties of** $\ln x$ and $\exp x$. The essential properties of logarithms and exponentials are - of course - well known $\cdots$ but still often forgotten.

$$\ln(a \times b) = \ln a + \ln b \tag{228}$$

$$\exp(a + b) = (\exp a) \times (\exp b) \tag{229}$$

- **Chain rule**

  When differentiating a function $F(u)$, where $u(x)$ is a function of the independent variable x, we can use the so-called chain rule

$$\frac{\partial F(u(x))}{\partial x} = \frac{\partial F(u)}{\partial u} \times \frac{\partial u(x)}{\partial x} \tag{230}$$

  More generally, if F is a function of u and u is a function of v $\cdots$ and y is a function of z, then

$$\frac{\partial F}{\partial z} = \frac{\partial F(u)}{\partial u} \times \frac{\partial u(v)}{\partial v} \times \frac{\partial v(w)}{\partial w} \times \cdots \times \frac{\partial y(z)}{\partial z}$$

- **Derivative of** $\exp(ax)$ and $\ln x$

  The derivative of $\exp(ax)$:

$$\frac{\partial \exp(ax)}{\partial x} = a \exp(ax) \tag{231}$$

  This result can easily be derived from the definition of $\exp(ax)$:

$$\exp(ax) = \lim_{n\to\infty} (1 + \frac{ax}{n})^n \tag{232}$$

  Coversely, the primitive function of $\exp(ax)$ is $a^{-1}\exp(ax)$.

  The derivative of $\ln x$ with respect to x is

$$\frac{\partial \ln x}{\partial x} = \frac{1}{x} \tag{233}$$

This is easily derived from Eqn.231. If $y = \ln x$, then $x = \exp(y)$, hence

$$\frac{\partial \ln x}{\partial x} = \frac{\partial y}{\partial \exp y} = \frac{1}{\exp y} = \frac{1}{x} \tag{234}$$

Conversely, the primitive function of $1/x$ is $\ln x$.

- **Taylor Expansion**

  If $f(x)$ and all its derivatives are smooth functions of x, then we can write:

  $$f(x+a) = f(x)+\left(\frac{\partial f}{\partial x}\right)_x a+\frac{1}{2!}\left(\frac{\partial^2 f}{\partial x^2}\right)_x a^2+\cdots+\frac{1}{n!}\left(\frac{\partial^n f}{\partial x^n}\right)_x a^n+\cdots \tag{235}$$

  The first two terms in the Taylor expansion are often used to approximate $f(x+a)$ if a is suuficiently small

  $$f(x+a) \approx f(x) + \left(\frac{\partial f}{\partial x}\right)_x a$$

  Specific examples are:

  $$\begin{aligned} \exp(x) &\approx 1+x \\ \ln(1+x) &\approx x \\ \sqrt{1+x} &\approx \frac{1}{2}x \\ (1+x)^n &\approx 1+nx \\ \sin(x) &\approx x \end{aligned}$$

  where, in all cases, it has been assumed that $x \ll 1$.

- **Permuting summation and multiplication**

  $$\sum_i \sum_j \sum_k \cdots \sum_n a_i b_j c_k \cdots z_n = \left(\sum_i a_i\right)\left(\sum_j b_j\right)\left(\sum_k c_k\right)\cdots\left(\sum_n z_n\right) \tag{236}$$

  (assuming that all sums converge uniformly)

- Permuting integration and multiplication

  $$\begin{aligned} &\int\int\int\cdots\int dx_1 dx_2 \cdots dx_n \; f(x_1)g(x_2)\cdots q(x_n) \\ &= \left(\int dx_1 \;\; f(x_1)\right)\left(\int dx_2 \; g(x_2)\right)\cdots\left(\int dx_n \; q(x_n)\right) \end{aligned} \tag{237}$$

In the special case that the functions $f, g, \cdots$,are all the same

$$\begin{aligned}
&\int\int\int\cdots\int dx_1 dx_2\cdots dx_n\ f(x_1)f(x_2)\cdots f(x_n) \\
&= \left(\int dx_1\ f(x_1)\right)\left(\int dx_2\ f(x_2)\right)\cdots\left(\int dx_n\ f(x_n)\right) \\
&= \left(\int dx\ f(x)\right)^n \qquad (238)
\end{aligned}$$

- **Geometric series**

  Consider the sum

  $$S = \sum_{i=0}^{n} a x^i \qquad (239)$$

  Clearly,

  $$xS = \sum_{i=0}^{n} a x^{i+1} = S - a + a x^{n+1} \qquad (240)$$

  Hence

  $$S(1-x) = a(1-x^{n+1}) \qquad (241)$$

  or

  $$S = \frac{a(1-x^{n+1})}{1-x} \qquad (242)$$

  If $|x| < 1$,we can take the limit $n \to \infty$:

  $$S_{n\to\infty} = \frac{a}{1-x} \qquad (243)$$

- **Factorials and permutations**

  The symbol $N!$ denotes the "factorial" of N. For positive, integer N, it is defined as

  $$N! = N \times (N-1) \times (N-2) \times \cdot \times 2 \times 1$$

  In addition, $0! \equiv 1$. The number of permutations of a set of N labeled objects is equal to $N!$. This can be demonstrated by induction. The number of ways in which a single object can be ordered is clearly equal to 1, which is equal to 1!. Hence, the relation holds for $N = 1$. The next step is to show that if the relation holds for N objects, it also holds for $N+1$ objects. This is easily demonstrated as follows. Assuming that there are $N!$ permutations for N objects, then for every permutation there are $N+1$ positions in the sequence where we could insert object $N+1$.

Hence the total number of permutations for $(N+1)$ objects is $(N+1) \times N! = (N+1)!$. This completes the proof.

Now consider the following question: we have N labeled objects and we wish to count the number of distinct ways that these objects can be divided into two sets, such that one set contains M elements and the other $N-M$ elements. For instance, 3 objects can be distributed in 3 ways over a subset of size one and a subset of size 2:

$$(1,23) \qquad (2,31) \qquad \text{and} \qquad (3,12)$$

Note that we do not count different permutations within one subset as distinct. To compute this number in general, we consider all possible permutations of N objects. There are $N!$ such permutations. For every permutation, we attribute the first M elements to one set, and the remaining $N-M$ elements to the other. In this way, we get that the total number of permutations with M elements in one set and $N-M$ in the other is equal to $N!$. However, in this counting procedure, we have considered different permutations of the objects in either set as distinct. To get the total number of ways to distribute N objects over the two subsets, we should divide by the number of permutations in the set of M objects and in the set of $N-M$ objects. The result is that the number of ways to divide N objects over two subsets of size M and $N-M$ respectively, is given by

$$\frac{N!}{M!(N-M)!}$$

- **Binomial and multinomial distributions.**

The number of ways to distribute N objects over two classes, in such a way that M objects end up in class I and $N-M$ objects in class II is given by

$$\frac{N!}{M!(N-M)!} \equiv \binom{N}{M} \tag{244}$$

For example: the number of ways in which I can throw N coins, such that M are head and $N-M$ are tail, is $\binom{N}{M}$. If we assume that the probability of head and tail are both equal to 1/2, then the probability that I throw M heads and $N-M$ tails is

$$P(M, N-M) = \binom{N}{M} 2^{-N} \tag{245}$$

In the more general case that the probabilities for the two events are not equal - say the probability to throw head (tail) is p $(1-p)$, then the probability to throw head M times and tail $N-M$ times is

$$P(M, N-M) = \binom{N}{M} p^M (1-p)^{N-M} \tag{246}$$

Of course, the sum of the probabilities of all different outcomes should add up to one

$$\sum \binom{N}{M} p^M (1-p)^{N-M} = (p + (1-p))^N = (1)^M = 1 \tag{247}$$

To give a specifi example, consider two containers, one with volume V_1 and the other with volume V_2. We assume that the probability that a molecule will be in volume 1 is equal to $V_1/(V_1+V_2)$. The probability to find a molecule in volume 2 is then $1 - V_1/(V_1+V_2) = V_2/(V_1+V_2)$. The probability to find M molecules in $V_1$and $N-M$ molecules in V_2 is then

$$P(M, N-M) = \binom{N}{M} \frac{V_1^M V_2^{N-M}}{(V_1+V_2)^N} \ . \tag{248}$$

The probability to find *all* molecules in volume 1 is

$$P(N, 0) = \frac{V_1^N}{(V_1+V_2)^N} \tag{249}$$

In case we distribute N objects over a larger number of classes - say m - the number of realizations is given by

$$\frac{N!}{\prod_{i=1}^{m} M_i!} \tag{250}$$

where M_i is the number of objects in class i and $\sum M_i = N$.

- **Some integrals.**

Certain integrals occur time and again in statistical physics. First of all, there are the integrals of the type:

$$\int_0^\infty dx\ x^n \exp(-ax) \tag{251}$$

All these integrals can be derived through integration by parts from the integral

$$\int_0^\infty dx\ \ \exp(-ax) = 1/a \tag{252}$$

For instance

$$\int_0^\infty dx\ x\exp(-ax) = -\frac{x}{a}\exp(-ax)|_0^\infty + \int_0^\infty dx\ \frac{\exp(-ax)}{a}$$
$$= \frac{1}{a^2} \tag{253}$$

The general result is

$$\int_0^\infty dx\ x^n\exp(-ax) = \frac{n!}{a^{n+1}} \tag{254}$$

This result can also be obtained by noting that

$$x^n\exp(-ax) = (-1)^n\left(\frac{\partial^n\exp(-ax)}{\partial a^n}\right) \tag{255}$$

and that therefore

$$\int_0^\infty dx\ x^n\exp(-ax) = (-1)^n\left(\frac{\partial^n(1/a)}{\partial a^n}\right) = \frac{n!}{a^{n+1}} \tag{256}$$

A second type of integral of particular importance is the Gaussian integral

$$I = \int_{-\infty}^\infty dx\ \ \exp(-cx^2) \tag{257}$$

A trick to compute this integral, is to consider its square

$$I^2 = \left(\int_{-\infty}^\infty dx\ \ \exp(-cx^2)\right)^2 = \int_{-\infty}^\infty dx\ \ \exp(-cx^2)\int_{-\infty}^\infty dy\ \ \exp(-cy^2) \tag{258}$$

We can write the latter product of integrals as

$$\int_{-\infty}^\infty dx\ \ \exp(-cx^2)\int_{-\infty}^\infty dy\ \ \exp(-cy^2) = \int_{-\infty}^\infty\int_{-\infty}^\infty dy\ dx\ \ \exp(-cx^2)\exp(-cy^2) \tag{259}$$

The latter integral is a two-dimensional integral. It can be simplified by using the polar coordinates r and ϕ, such that $x = r\cos\phi$ and $y = r\sin\phi$. Clearly, $x^2+y^2 = r^2$. The integration range for ϕ is $\{0, 2\pi\}$ and r ranges from 0 to ∞. Finally, we replace the area element $dx\ dy$ by $rd\phi\ dr$. We can then write

$$\begin{aligned} I^2 &= \int_0^{2\pi} d\phi\int_0^\infty dr\ r\exp(-cr^2) \\ &= 2\pi\int_0^\infty \frac{1}{2}dr^2\ \ \exp(-cr^2) \\ &= \pi\int_0^\infty dr^2\ \ \exp(-cr^2) \\ &= \frac{\pi}{c} \end{aligned} \tag{260}$$

where, in the third line, we have used $dr^2 = 2r\ dr$. To arrive at the last equality, we used Eqn.252. Hence

$$\int_{-\infty}^{\infty} dx\ \ \exp(-cx^2) = \sqrt{\frac{\pi}{c}} \tag{261}$$

- **Stirling's approximation**

From equations254 and 261 above, we can derive Stirling's approximation for $N!$.

$$N! = \int_0^\infty dx\ x^N \exp(-x) = \int_0^\infty dx\ \ \exp(-x + N \ln x) \tag{262}$$

where we have used Eqn.254 with $a = 1$. The integrand is sharply peaked at $x = N$. The value of the exponent at $x = N$ is $-N + N \ln N$. The first derivative is zero (we are at a maximum). The second derivative is $-1/N$. Hence, we can approximate the integral by

$$\begin{aligned} N! &\approx \int_0^\infty dx\ \ \exp(-N + N \ln N - \frac{(x-N)^2}{2N}) \\ &= \int_{-N}^\infty du\ \ \exp(-N + N \ln N - \frac{u^2}{2N}) \end{aligned} \tag{263}$$

where we have defined $u \equiv x - N$. As the function is sharply peaked, we can replace the lower limit of the integration by $-\infty$. We then have

$$\begin{aligned} N! &\approx \exp(-N + N \ln N) \int_{-\infty}^\infty du\ \ \exp(-\frac{u^2}{2N}) \\ &= \exp(-N + N \ln N)\sqrt{2\pi N} \\ &= N^N \exp(-N)\sqrt{2\pi N} \end{aligned} \tag{264}$$

where we have used Eqn.261. This is Stirling's approximation for $N!$. In fact, Stirling's approximation is the first term of a series

$$N! = N^N \exp(-N)\sqrt{2\pi N}\left(1 + \frac{1}{12N} + \frac{1}{288N^2} - \frac{139}{51840N^3} + \cdots\right) \tag{265}$$

B Appendix: Essential Thermodynamics

The First Law of Thermodynamics expresses the conservation of energy. It can be written as

$$dU = \ q + \ w \tag{266}$$

The change of the internal energy of the system is equal to the amount of heat transferred to the system plus the amount of work performed on the system. The energy of a system is a state function: i.e. when, at the end of a process, a system is returned to its initial state, then the internal energy of the system remains unchanged in the process. Note that work (w) and heat (q) are *not* state functions.

The Second Law of Thermodynamics was discussed in Section2. The law is a consequence of the experimental observation that it is not possible to devise a cyclic process that has as its only consequence that heat, taken from a reservoir, is completely converted into work. Note that such a process is not forbidden by the First Law. As explained in Section2, this allows us to define an absolute temperature scale (T) and it leads to the introduction of a state function, S, called "Entropy". In terms of these quantities, the Second Law of Thermodynamics can be expressed in a different way: *A spontaneous process in an isolated system always results in an increase of the entropy.* The change in the entropy of a system can be written as

$$dS_{tot} = \frac{q_{rev}}{T} \tag{267}$$

where q_{rev} denotes the amount of heat that is reversibly transferred to the system. The work performed on the system can have many different forms. It can be mechanical, electrical or chemical.

Having defined the internal energy (U), the temperature (T) and the entropy (S), we can now use the first and second law of thermodynamics to introduce a number of additional useful relations. Let us first consider a very large, isolated system. Inside this system, we have a small system with volume V, internal energy U and entropy S. Let us first assume that this system can exchange only heat with the big reservoir around it. Initially, the reservoir and the system are not in equilibrium. Hence, spontaneous changes take place that must increase the total entropy. Let us compute the change in the total entropy of the system due to an infinitesimal transfer of heat (q) between the reservoir and the small system. The total change in entropy is

$$dS_{tot} = dS_{res} + dS_{sys} \geq 0 \tag{268}$$

As the system can only exchange heat with the reservoir, the First Law implies that $dU_{sys} = q$. The change in entropy of the reservoir is equal to $dS_{res} = -\ q/T$. Hence, the total change in entropy can now be expressed in terms of state functions of the system (and the temperature of the reservoir):

$$dS_{tot} = \frac{-dU_{sys}}{T} + dS_{sys} = -(dU_{sys} - TdS_{sys})/T \geq 0 \tag{269}$$

As the temperature of the (effectively infinite) reservoir is constant, we can write

$$-(dU_{sys} - dTS_{sys})/T \geq 0 \tag{270}$$

or

$$d(U_{sys} - TS_{sys}) \leq 0 \tag{271}$$

As both U and S are state functions, the combination $U - TS$ is also a state function. The combination $U - TS$ is called the *Helmholtz free energy* of the system. It is denoted by the symbol A. The second law of thermodynamics therefore implies: *for a system with a fixed volume and a fixed number of particles at constant temperature, the Helmholtz free energy is at a minimum.* Similarly, we can derive that, for a system with a fixed number of particles *at constant temperature and pressure*, the state function $G \equiv U + PV - TS$ is at a minimum. G is called the *Gibbs free energy.* The only difference is that now we must write dU_{sys} $=q+$ w, $w = -PdV_{sys}$. Hence, $q_{rev} = dU_{sys} + PdV_{sys}$. The change in the total entropy is then

$$dS_{tot} = \frac{-dU_{sys} - PdV_{sys}}{T} + dS_{sys} = -\frac{d(U_{sys} + PV_{sys} - TS_{sys})}{T} \geq 0 \tag{272}$$

or $dG_{sys} \leq 0$. Finally, we consider a situation where the small system can exchange heat *and* particles with the reservoir. In that case, the work is chemical work

$$w = \sum_i \mu_i dN_i \tag{273}$$

Then we get

$$dS_{tot} = \frac{-dU_{sys} + \sum_i \mu_i dN_i}{T} + dS_{sys} = -\frac{dU_{sys} - \sum_i \mu_i dN_i - TdS_{sys}}{T} \geq 0 \tag{274}$$

Another state function that is often useful, is the *enthalpy* $H \equiv U + PV$.

The general relations for infinitesimal variations in S, U, H, A and G are

$$dS = \frac{dU}{T} + \frac{P}{T}dV - \sum_i \frac{\mu_i}{T} dN_i \tag{275}$$

$$dU = TdS - PdV + \sum_i \mu_i dN_i \tag{276}$$

$$dH = TdS + VdP + \sum_i \mu_i dN_i \tag{277}$$

$$dA = -SdT - PdV + \sum_i \mu_i dN_i \tag{278}$$

$$dG = -SdT + VdP + \sum_i \mu_i dN_i \tag{279}$$

From the above expressions, we can derive a number of Maxwell relations. Consider an arbitrary state function $Q(x, y, z)$ where x, y and z are the independent variables that characterize the state point (e.g. the set (x, y, z) could correspond to (T, V, N)). We can then write

$$dQ = \left(\frac{\partial Q}{\partial x}\right)_{y,z} dx + \left(\frac{\partial Q}{\partial y}\right)_{x,z} dy + \left(\frac{\partial Q}{\partial z}\right)_{x,y} dz \tag{280}$$

Clearly,

$$\left(\frac{\partial^2 Q}{\partial x \partial y}\right) = \left(\frac{\partial^2 Q}{\partial y \partial x}\right) \tag{281}$$

If we apply this (for instance) to the thermodynamic state function A, we get

$$\left(\frac{\partial P}{\partial T}\right)_{V,\{N_i\}} = \left(\frac{\partial S}{\partial V}\right)_{T,\{N_i\}} \tag{282}$$

$$\left(\frac{\partial \mu_i}{\partial T}\right)_{V,\{N_i\}} = \left(\frac{\partial S}{\partial N_i}\right)_{T,V,\{N_j\}} \tag{283}$$

Other important thermodynamic quantities are the heat capacity at constant volume (C_V) and the heat capacity at constant pressure (C_P):

$$C_V = \left(\frac{\partial U}{\partial T}\right)_{V,\{N_i\}} \tag{284}$$

$$C_P = \left(\frac{\partial H}{\partial T}\right)_{P,\{N_i\}} \tag{285}$$

Finally, any function of state functions is a state function. This is sometimes useful. For instance, consider the state function A/T.

$$d\frac{A}{T} = \frac{dA}{T} + Ad\frac{1}{T} = \frac{-SdT - PdV + \sum_i \mu_i dN_i}{T} - A\frac{dT}{T^2} \tag{286}$$

at constant volume and constant number of particles, this becomes

$$\begin{aligned}\left(d\frac{A}{T}\right)_{V,\{N_i\}} =& \frac{-SdT}{T} - A\frac{dT}{T^2} = -(TS+A)\frac{dT}{T^2} \\ =& Ud\frac{1}{T} \end{aligned} \tag{287}$$

and hence

$$\left(\frac{\partial(A/T)}{\partial(1/T)}\right)_{V,\{N_i\}} = U \tag{288}$$

C Appendix: Essential quantum mechanics.

Statistical thermodynamics relates the equilibrium behavior of macroscopic systems to the physical properties of the constituent molecules. As the behavior of individual molecules is determined by the laws of quantum mechanics, statistical thermodynamics makes extensive use the results of quantum-mechanical calculations. However, only a very limited amount of quantum mechanics is needed for the present course.

The most important result of quantum mechanics that we use is the following: *A system can only be found in discrete states.* For a macroscopic isolated system, the number of these discrete states may be extremely large, but it is *finite* and, at least in principle, *countable.* The basic assumption of statistical thermodynamics is that, in an isolated system (*i.e.* a system with fixed volume, energy and number of particles) every quantum state occurs with equal probability. If the total number of states of the system is denoted by $\Omega(N,V,U)$, then the probability to find the system in any one of these states is $1/\Omega(N,V,U)$.

C.1 Molecular energy levels

In general, we cannot compute the quantum levels of a macroscopic system. However, for isolated molecules, the situation is better. Using the techniques of quantum chemistry, it is possible to compute molecular quantum levels with high accuracy. We shall not do this. For the purpose of this course, we shall assume that the motion of a molecule can be decomposed into individual modes of motion that are uncoupled. Three modes of motion can be described in the language of classical mechanics. These are: *translation* (i.e. the motion of the molecular center of mass), *rotation* (i.e. the overall rotation of a molecule around its center of mass) and *vibration* (i.e. the oscillatory motion that the atoms in the molecule carry out when we stretch or bend chemical bonds in the molecule). Loosely, one could call translation, vibration and rotation "semi-classical" forms of motion. By this I mean that the energy levels for these forms of motion can be estimated by combining our knowledge of the corresponding classical motion with simple “quantization rules”. In addition, certain molecular energy levels cannot be deduced in any simple way from classical mechanics. These are the levels that are associated with excitation of *electronic motion* in the molecule, and energy levels that differ in the state of the *spin* of the nuclei or electrons in the atom. If we assume that the various modes of motion are uncoupled, we can describe them all independently. This is a great simplification. However, it should be stressed that, in reality, there is always some coupling between the different modes of motion.

C.1.1 Translation

It is easy to compute the energy levels for a particle of mass m moving in a one-dimensional box of length L. Classically, the energy of a such a particle is simply the kinetic energy $E = \frac{1}{2}mv^2 = \frac{p^2}{2m}$, where v is the velocity of the particle and p is its momentum. The only quantum mechanics that we use is that the momentum p can be written as h/λ where h is Planck's constant and λ is the “wavelength” of the particle and that every energy level of the particle in the box corresponds to a standing wave. The condition for a standing wave is: $L = n\lambda/2$, with n an arbitrary positive integer. Hence, $\lambda = 2L/n$. If we insert this value for λ in the expression for the classical kinetic energy, we find:

$$E_n = \frac{(h/\lambda)^2}{2m} = \frac{n^2h^2}{8mL^2} \tag{289}$$

Although our “derivation” was oversimplified, the answer is correct. Of course, we live in a three-dimensional world. Hence, the more relevant question is: *what are the energy levels of a particle in a three-dimensional box ?* Let us assume that we have a rectangular box with sides L_x, L_y and L_z. The energy levels of a particle of mass m in such a box are

$$E_{n,m,l} = \frac{n^2h^2}{8mL_x^2} + \frac{m^2h^2}{8mL_y^2} + \frac{l^2h^2}{8mL_z^2} \tag{290}$$

For a cubic box with sides L, the levels are

$$E_{n,m,l} = \frac{(n^2 + m^2 + l^2)h^2}{8mL^2} \tag{291}$$

C.1.2 Rotation

The quantum-mechanical energy levels of a rotating particle depend on the moment of inertia of the particle. The simplest case is that of a linear molecule with a moment of inertia I. In general, the moment of inertia of a linear molecule consisting of n atoms, is given by

$$I = \sum_{i=1}^{n} m_i(x_i - X_{CM})^2$$

where m_i is the mass of the i-th atom, x_i is its position and X_{CM} is the position of the center of mass - where we have assumed that the molecule is oriented along the x-axis.

$$X_{CM} = \sum_{i=1}^{n} m_i x_i / M$$

where $M = \sum m_i$. For a diatomic molecule:

$$I = m_1(x_1 - X_{CM})^2 + m_2(x_2 - X_{CM})^2$$

If we insert the expression for X_{CM} we find

$$\begin{aligned} I = & \frac{m_1}{M^2}(m_2 x_1 - m_2 x_2)^2 + \\ & \frac{m_2}{M^2}(m_1 x_2 - m_1 x_1)^2 \end{aligned}$$

or

$$\begin{aligned} I &= \frac{m_1 m_2}{m_1 + m_2}(x_1 - x_2)^2 \\ &\equiv \mu r_e^2 \end{aligned}$$

where

$$\mu \equiv \frac{m_1 m_2}{m_1 + m_2}$$

defines the reduced mass of the diatomic molecule and

$$r_e = |x_1 - x_2|$$

defines its equilibrium length.

Classically, the rotational kinetic energy of such a molecule would be $\frac{1}{2} I \omega^2 = J^2/(2I)$, where ω is the angular velocity of the molecule and J its angular momentum. Quantum mechanically, only discrete values of J^2 are allowed, namely $J^2 = j(j+1)\hbar^2$, where $\hbar = h/(2\pi)$ and j is an arbitrary non-negative integer. Hence the energy levels of a linear rotor are given by:

$$E_j = \frac{j(j+1)\hbar}{2I} \qquad (j = 0, 1, 2, \dots) \tag{292}$$

Classically, the angular momentum vector can have any orientation in space. Quantum mechanically, only discrete orientations are allowed. The number of allowed orientations of the angular momentum of a linear molecule with rotational quantum number j is equal to $(2j+1)$. Hence, different rotational energy levels have different degeneracies. For instance, for $j = 0$, the degeneracy is 1, for $j = 1$, the degeneracy is 3 and so on ...

The rotational energy levels of nonlinear molecules can also be computed on basis of the knowledge of their moment of inertia (which is now a tensor). For details, see *e.g.* Atkins, section 16.4.

C.1.3 Vibration

Suppose that we have a classical spring with spring-constant k, *i.e.* the relation between the extension x of the spring and the restoring force f is

$$f = -kx \tag{293}$$

From classical mechanics, it is easy to derive that if a particle of mass m is connected to the spring, then an initial displacement of the spring will lead to an oscillatory motion with angular frequency $\omega = \sqrt{k/m}$. The classical energy of a harmonic oscillator is: $E_{cl} = \frac{1}{2}m\omega^2x^2 + \frac{1}{2}mv^2$. For a periodic motion (*e.g.* $x(t) = x(0)\cos(\omega t)$ and $v(t) = -x(0)\omega\sin(\omega t)$), the total kinetic energy is: $E_{cl} = \frac{1}{2}m\omega^2x^2(0)$. However, quantum-mechanically not all energies are allowed. It turns out that the allowed energy levels satisfy

$$E_n = (n + \frac{1}{2})h\nu, \tag{294}$$

where ν is the vibrational frequency ($v = w/(2\pi)$). The energy levels of a single, one-dimensional harmonic oscillator are non-degenerate.

C.1.4 Electronic

To find the electronically energy levels of a molecule ($\epsilon_0, \epsilon_1, \epsilon_2$ etc.), one really needs to solve the full Schrödinger equation for this system. Moreover, different molecules will have different electronic energy levels. We will therefore not attempt to write down any "general" form for the electronic energy levels of a molecule.

C.1.5 Spin

Electrons have a "spin" of $\frac{1}{2}$. This means that they carry angular momentum. Similarly, many nuclei have a spin too. Depending on the nature of the nucleus 7the nuclear spin can be either be integer valued (0, 1, 2 etc.) or half-integer valued ($\frac{1}{2}$, $\frac{3}{2}$, etc.). It turns out that particles with half-integer spins ("fermions") behave very differently from particles with integer spins ("bosons"). The angular momentum of a particle with spin can have only discrete orientations in space. The number of discrete orientations is equal to $(2s + 1)$ where s is the integer or half-integer that denotes the spin state of the particle. For example, a single electron can have 2 ($= 2\frac{1}{2} + 1$) spin states.

D Appendix: Spring constant of an arbitrary polymer

Consider the average end-to-end distance ($< X_{ee} >_F$) of an arbitrary polymer that is being stretched by a force F. A non-ideal polymer with a conformation Γ_i has an internal energy $U_{int}(\Gamma_i)$ that depends on the details of that conformation. In addition, due to the external force, the potential energy of a polymer with end-to-end distance X_{ee} is lowered by an amount $X_{ee}F$. The average end-to-end distance of the polymer is

$$< X_{ee} >_F = \frac{\sum_{i=1}^{\Omega_T} X_{ee}(\Gamma_i)\exp(-\beta(U(\Gamma_i) - X_{ee}(\Gamma_i)F))}{\sum_{i=1}^{\Omega_T} \exp(-\beta(U(\Gamma_i) - X_{ee}(\Gamma_i)F))} \tag{295}$$

where Ω_T denotes the total number of conformations of the polymer. If the applied force is sufficiently small, we can write $\exp(\beta X_{ee}(\Gamma_i)F) \approx 1 + \beta X_{ee}(\Gamma_i)F$.

$$< X_{ee} >_F \approx \frac{\sum_{i=1}^{\Omega_T} X_{ee}(\Gamma_i)(1 + \beta X_{ee}(\Gamma_i)F)\exp(-\beta(U(\Gamma_i))}{\sum_{i=1}^{\Omega_T}(1 + \beta X_{ee}(\Gamma_i)F)\exp(-\beta(U(\Gamma_i))} \tag{296}$$

If we now divide and multiply by $\sum_{i=1}^{\Omega_T} \exp(-\beta(U(\Gamma_i))$, we get

$$\begin{aligned} < X_{ee} >_F \approx & \frac{\sum_{i=1}^{\Omega_T} X_{ee}(\Gamma_i)(1 + \beta X_{ee}(\Gamma_i)F)\exp(-\beta(U(\Gamma_i))}{\sum_{i=1}^{\Omega_T} \exp(-\beta(U(\Gamma_i))} \\ \times & \frac{\sum_{i=1}^{\Omega_T} \exp(-\beta(U(\Gamma_i))}{\sum_{i=1}^{\Omega_T}(1 + \beta X_{ee}(\Gamma_i)F)\exp(-\beta(U(\Gamma_i))} \\ = & \frac{< X_{ee} >_{F=0} + \beta F < X_{ee}^2 >_{F=0}}{1 + \beta F < X_{ee} >_{F=0}} \end{aligned} \tag{297}$$

Note that all the averages appearing in this equation are averages for the *unperturbed* polymer (i.e. in the absence of an external force). If we now make use of the fact that $< X_{ee} >_{F=0}$vanishes (due to symmetry) we get

$$< X_{ee} >_F = \beta F < X_{ee}^2 >_{F=0} \tag{298}$$

This is the relation between force and elongation for a harmonic spring with force constant

$$\kappa = \frac{k_B T}{< X_{ee}^2 >} \tag{299}$$