# 3. **Microstates: Statistical Mechanics in the Microcanonical Ensemble**

Goal: Within the Microcanonical Ensemble, derive definitions for entropy, temperature, and for Thermodynamic quantities like free energy and Enthalpy

3(a) Sub-Systems, Entropy, & Temperature

There are many ways to set up the foundations of statistical mechanics. In this chapter we will beginning with the Microcanonical ensemble. This describes a system Σ of constant total energy, so that the only available microstates are the ones having this energy.

Notice however that if we sub-divide S into a set of M sub-systems, or 'cells', then the energy of each sub-stem is not necessarily fixed. Each of these sub-systems is able to exchange energy, either bulk kinetic or heat energy (however we will assume no particles are exchanged). The total system is however utterly closed off from the rest of the universe.
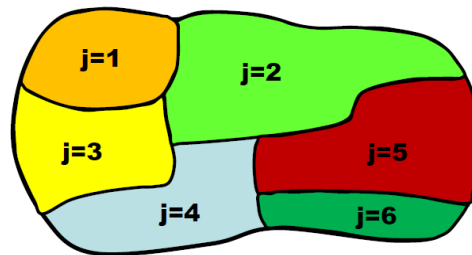
The figure shows a total system Σ which is divided into 6 different sub-systems Σ_j, with j = 1,2, …6. . We note that a given sub-system may couple to a number of other sub-systems, but unless there are long-range interactions in the system, it will not in general couple to all of the other systems. Two obvious examples of long-range interactions are gravitational interactions, and dipolar interactions (note that Coulomb interactions are usually not long-range, because they are screened).
   We assume that each sub-system is also macroscopic, so that we can make the same kind of statistical arguments for it as we do for the whole.

Two Sub-Systems: To warm up, let us begin with two sub-systems Σ_1 and Σ_2 which are separated by a *diathermal* wall, ie., one which transmits heat but not particles. The wall is for the moment considered to be fixed – we cannot move it. The situation is thus as exhibited in the figure below:
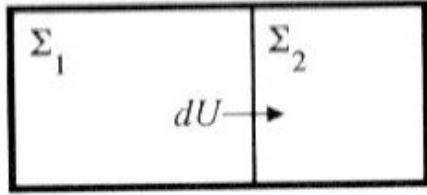
Let Ω_1 and Ω_2 be the multiplicity (ie., the number of microstates) for each of the two systems, consistent with the constraints on the two systems (here, that the total energy is constant, that the partition is fixed, and that no particles are exchanged across it). Then the total multiplicity Ω for the combined system is $\Omega = \Omega_1 \Omega_2$ so that

$$\ln\Omega = \ln\Omega_1 + \ln\Omega_2 . \tag{1}$$

Thus $\ln\Omega$ is an additive quantity whereas Ω itself is not. This of course is exactly what we would expect, since have already seen that the multiplicity increases extremely quickly with the number of particles/spins/degrees of freedom of the system.

Now in equilibrium we have already argued that Ω should be at a maximum, simply because this maximizes the number of available microstates, and they are all supposed to have equal *a priori* probability. This in turn implies that the differential $d\Omega = 0$ with respect to any changes in the system, ie., we have

$$d\ln\Omega = \frac{d\Omega}{\Omega} = d\ln\Omega_1 + d\ln\Omega_2 = 0 . \tag{2}$$

Now suppose that a small amount of thermal energy dU is transferred from Σ_1 to Σ_2 with no change in the volumes of the two systems, and where we maintain thermal equilibrium. Then we must have

$$d\ln\Omega = \left(\frac{\partial\ln\Omega_1}{\partial U_1}\right)_V dU_1 + \left(\frac{\partial\ln\Omega_2}{\partial U_2}\right)_V dU_2 \tag{3}$$

However $dU_1 = -dU_2 = dU$ by conservation of energy; it then immediately follows that

$$d\ln\Omega = \left[\left(\frac{\partial\ln\Omega_1}{\partial U_1}\right)_V - \left(\frac{\partial\ln\Omega_2}{\partial U_2}\right)_V\right] dU = 0 \tag{4}$$

Thus in thermal equilibrium we have: $\quad \left(\frac{\partial\ln\Omega_1}{\partial U_1}\right)_V = \left(\frac{\partial\ln\Omega_2}{\partial U_2}\right)_V \tag{5}$

Now we can also argue, as we did in the discussion of the 0th law of TD, that this means that these 2 systems must have something in common, which we called temperature. This argument was unsatisfactory in that it was arbitrary – why should we necessarily assume that temperature is the only thing that systems in mutual equilibrium have in common?

Let us return to our discussion of thermodynamics, and recall the definition we gave of temperature in terms of the energy dependence of the entropy. We saw that we could define the absolute temperature according to

$$T \equiv \left(\frac{\partial U}{\partial S}\right)_V = \left(\frac{\partial S}{\partial U}\right)_V^{-1} \qquad (6)$$

and we also saw that this quantity must be the same for any two bodies in mutual equilibrium. The comparison between (1) and (2) immediately tells us that the entropy must be proportional to $\ln \Omega$, and so we write the fundamental relation first written by Boltzmann in 1872, viz.,

$$S \equiv k_B \ln \Omega$$

(in fact the equation was first written in this form by Planck). The constant k_B is known as Boltzmann's constant. Since both S and $\ln \Omega$ are macroscopic (of order N), ie., they are both extensive variables, we have no indication of how big k_B is; its actual value is fixed by the units of temperature we employ. By modern convention

$$k_B \equiv 1.38 \times 10^{-23} J/K \quad \text{(SI units)} \qquad (7)$$

which defines the temperature T in terms of the Kelvin scale (degrees K). The reason why k_B is so small here is that it is measured in macroscopic unit (Joules), whereas the definition of entropy is in terms of microstates.

Let us also note that the equation above as written down before we knew about Quantum Mechanics. At first glance it seems to support the 3rd law of thermodynamics in the form originally written down, viz., that the entropy goes to zero at T = 0. We now know, of course, that the ground state of a quantum system may be degenerate, involving a finite number of states. However, this does not make the above equation invalid – it simply means that at T = 0, the entropy goes to $S_0 = k_B \ln D_0$, where D_0 is the degeneracy of the ground state.

N Sub-Systems: Clearly we can generalize the arguments above to deal with a system split into N sub-systems – the argument can be done iteratively starting from the 2 sub-system argument as formulated above. We simply employ the same argument in reverse,

to argue that any other system j with the same value of $\left(\dfrac{\partial \ln \Omega_j}{\partial U_j} \right)_V$ must in thermal equilibrium with systems 1 and 2. Alternatively we can imagine adding a third system, and go through the same arguments as above – the energy flows are more complicated, but by going through the exercise you will see that we can generalize iteratively to M sub-systems.

These results are established using statistical arguments regarding the number of microstates in each system, and defining equilibrium via the stationarity of $\ln \Omega$. Notice however that we have actually established the 0th law of thermodynamics using these arguments.

Let us however go into this a little more deeply. We have seen that for a given system, the number of available microstates in the microcanonical ensemble depends only on the energy E. We might ask whether it can depend on anything else – eg., the total angular momentum $\mathbf{L}$, or the linear momentum $\mathbf{P}$ of the system. In general it can – in fact we expect it to depend on all the conserved quantities for the system (which we recall is isolated).

Let us ignore the dependence on $\mathbf{L}$ and $\mathbf{P}$ (which can in any case be removed by choosing an appropriate reference frame). Then we can say that $S(E) = k_B \ln \Omega(E)$, ie., the entropy depends only on the energy. However we can also say a little more – since both the energy E and the entropy are additive quantities (extensive variables) it follows that we can write $S(E) = a + bE$, ie. the entropy must be a linear function of the energy. Notice that this also means that the average entropy (if it is changing in time) must also be a linear function of the *average* energy, ie. we have

$$< S(E) > \;=\; k_B < \ln \Omega(E) > \;=\; a + b <E> \tag{8}$$

Now this is very important, because we can generalize this result immediately to all the different sub-systems in the total system – even though their energies can all individually change, nevertheless their *average energies* will obey (8). In other words we can add the results (8) over all the different sub-systems, to get

$$S(E) \;=\; k_B \, \Sigma_j < \ln \Omega_j (E_j) > \;=\; S_0 + bE \tag{9}$$

ie.,

$$=\; \Sigma_j \, S_j \tag{10}$$

where $E = \Sigma_j <E_j>$ is the total energy of the system (which is of course a constant), and the $\{ S_j \}$ are the entropies of the individual sub-systems.

Using the result in eqn. (9) we can also rewrite things in a very well known form, as follows. Note first that, for a sub-system with total multiplicity $\Omega^{(j)}$, the probability of getting a microstate is given by

$$P^{(j)} \;=\; 1/\Omega^{(j)} \tag{11}$$

so that the logarithm of this probability is equal to

$$\ln P^{(j)} \;=\; -\ln \Omega^{(j)} \tag{12}$$

Suppose now we consider the entire set of subsystems, and suppose that at energy E there is but one microstate. Then from comparing (9) and (12) we find that the entropy of the entire system is equal to

$$S(E) \;=\; k_B \, \Sigma_j <\ln \Omega j \, (Ej) > \;=\; - k_B \, \Sigma j \; (P^{(j)} \ln \; P^{(j)}) \tag{13}$$

where the constant S_0 has disappeared in the normalization of the probabilities (so that we have $\Sigma j \, P^{(j)} = 1$).

Now of course in general the degeneracy (ie., the multiplicity) of the macrostate with energy E may not be 1, and certainly the multiplicity of the states of each sub-system will not be. But this makes no difference to the argument – we can simply rewrite (13) assuming that the sum is over all collections of states, labeled by $\{\,j,n\,\}$, for a given sub-system j, such that the set of states labeled by n for subsystem j has multiplicity $W_n^{(j)}$. The argument goes through exactly as before, and we find that

$$S(E) \;=\; - k_B \, \Sigma_{j,n} \; (P_n^{(j)} \ln \; P_n^{(j)}) \tag{14}$$

Now this result is very famous – it was first written down by Gibbs in 1878. Much later, in 1948, it was re-interpreted by Shannon in information-theoretical terms (similar ideas having been given earlier by, eg., Szilard); I shall talk about this later. What is basically says is that the entropy is the average value, over different sub-systems in the total system, or the logarithm of the probabilities for the different sub-systems to be in the allowed states.

3(b) **Some Examples**

A good way to get the hang of all of this is to look at some examples – we will look at three of these.

Example 1 (spin ½ Magnet): Let us consider the entropy of a finite spin ½ magnetic system in a field B. We consider a model system of $N = 100$ spin ½ particles, and write again the difference $n \equiv N_\uparrow - N_\downarrow$. As before, the available energies are

$$E_n = -n\mu B \qquad (100 > n > -100) \tag{15}$$

A key feature of this system is thus that its quantum spectrum is bounded both above and below; the states only exist between energies E such that $-N\mu B < E < N\mu B$, with $N{=}100$.

As we have already seen, the density of states as a function of energy is very strongly peaked around $n = 0$; we can show this here by calculating a few of the multiplicities for different values of n, to get

n	E_n	Ω_n
N	$-N\mu B$	1
$N-2$	$-(N-2)\mu B$	$C^N_1 = N$
$N-4$	$-(N-4)\mu B$	$C^N_2 = N(N-1)/2$
$N-100 = 0$	0	$C^N_{50} \sim 10^{29}$

We see that the entropy in the lower half of the spectrum, as a function of energy, obeys $\left(\dfrac{\partial S}{\partial U}\right)_V > 0$ as we would expect; indeed the increase is quite rapid.

However, because the density of states or multiplicity then decreases once we have passed the zero energy point – it is of course symmetric in energy – we arrive at what seems a quite peculiar result, viz, that if the system happens to be in some macrostate with net positive energy, its temperature will be *negative*!
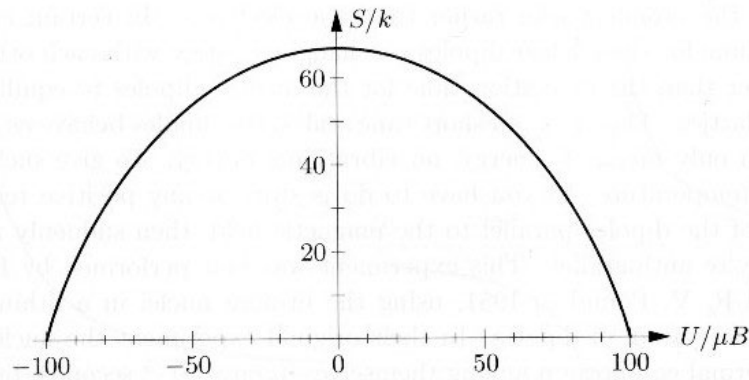

Figure 3.8. Entropy as a function of energy for a two-state paramagnet consisting of 100 elementary dipoles.

It is important to realize that this is not a contradiction or an incorrect result. It arises here because the spectrum is bounded above, with a maximum density of states in an intermediate range of energies, above which the density of states begins to decrease. We will also see, when we come to the canonical ensemble, that for any system in contact with a heat bath – without which it is hard to define a consistent notion of temperature – the probability of a state with energy E_n being occupied, when one is in equilibrium, will be proportional to $\exp[-E_n/k_B T]$. Thus in this case the problem will never arise.

However it is also important to realize that what appear, for all practical purposes, to be negative temperatures, can be achieved in the real world over limited time periods. What is required is some sub-system of a physical system that can be thermally isolated from the bulk for long periods, and then prepared in a non-equilibrium state in which higher energy levels are preferentially occupied. Then, for a limited time period, it will behave as though it is at negative temperature.
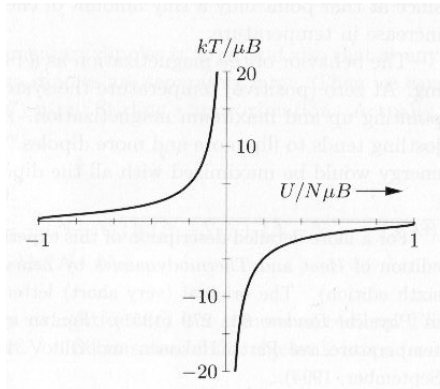
As an example, experiments have prepared the nuclear spin sub-system in a negative T state, in NMR experiments. One polarizes the proton spins in a static magnetic field (so that one is in a set of microstates with $n \sim N$), and then applies a 180 degree RF pulse to suddenly invert the polarization (alternatively one could suddenly invert the external field, but this is more difficult!). The spins are then out of equilibrium, with the higher energy states more occupied than the lower energy states. The system is in internal thermal equilibrium at negative T, for a timescale T_2 (the nuclear spin-spin relaxation rate).

How does this nuclear spin sub-system see the external thermal environment? This is done by interaction of the nuclear spins with phonons, which are themselves everywhere in the system, and which equilibrate internally very fast (mainly via phonon-phonon interactions). The spin-phonon coupling transfers energy from the spins to the phonons over a timescale T_1.

However nuclear spins at low temperatures are coupled very poorly to phonons (at mK temperatures in magnetic insulators it can take months!). Thus the nuclear spin-spin relaxation rate ($1/T_2$) is much higher than the spin-phonon relaxation rate ($1/T_1$), and so the nuclear spins are in internal quasi-equilibrium with themselves as they slowly relax back to the crystal temperature.

These remarks re-emphasize that one needs to be careful in trying to understand the notion of thermal equilibrium, and how it depends in an essential way on timescales.

Example 2: Entropy of mixing: Suppose we have two separate crystals A and B with N_A atoms and N_B atoms respectively. We will neglect all other sources of entropy except the "*configurational entropy*" associated with the atomic positions (so, eg., we neglect the entropy associated with vibrational motion, ie., the phonon entropy).

At $T=0$, the system will be completely ordered, so that

$$\Omega = 1, \quad S = 0 \qquad (T=0) \qquad\qquad (16)$$

Now assume that we have randomly mixed the atoms in the 2 crystals, such that the probability of finding atom A at any given site is $x = \dfrac{N_A}{N}$ where $N = N_A + N_B$. Then

$$\Omega = \frac{N!}{(xN)![(1-x)N]!} \tag{17}$$

Keeping the first two terms in Stirling's approximation $\ln N! \cong N \ln N - N$

$$
\begin{aligned}
\ln \Omega &= N \ln N - N - xN \ln xN + xN - (1-x)N \ln(1-x)N + (1-x)N \\
&= N \ln N - N - xN \ln x - xN \ln N + xN - (1-x)N \ln(1-x) - (1-x)N \ln N + (1-x)N \\
&= -xN \ln x - (1-x)N \ln(1-x) \\
&= N \ln 2 \quad for \; x = 1/2
\end{aligned} \tag{18}
$$

Thus the maximum increase in entropy $S = Nk_B \ln 2$ is comparable to a change in temperature of factor of 2. Note this term is T independent. Thus if we could arrange for the system to be in this mixed state at T=0, it would have a rather large entropy! The physical question then becomes whether or not this is actually an equilibrium states in the T → 0 limit. Typically it will not be – the system will prefer to order – but how long this takes will depend on energy barriers for the motion of the atoms, and it could be a very long time.

Example 3: 1D polymers and the entropic force: Consider a 1D polymer with N links of length l. We will use a very crude model, in which the bonds between the links are restricted to either 0 or 180 degrees, as shown.
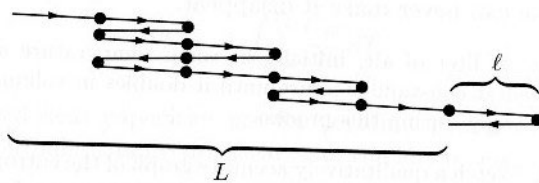
The internal energy is assumed to be independent of the configuration (ie., no interaction between the links, or energy associated with the bond angles), and we will ignore vibrational entropy. Thus, the potential energy of all the different possible configurations is the same!

However, as before, there is configurational entropy in the system. Let L be the position of one end relative to the other. Let us assume that L/l is an even number, so that one possibility is that L is zero. Clearly there is only one way to be in a fully stretched state whereas there are in general a large number of configurations which yield a short distance.

We see immediately that the problem maps onto the spin ½ magnet. Thus in the limit $L \ll Nl$ we have

$$\Omega(L) = \Omega_0 \exp\left[-\frac{L^2}{2Nl^2} \right] \qquad (19)$$

so that the configurational entropy is:

$$S_{con}(L) = S(0) - \frac{k_B L^2}{2Nl^2} \qquad (20)$$

Now assume we fix one end and apply a force to the other end. There will be a restoring force f, so that the polymer tries to shorten its length – this is simply because doing so will reduce its entropy, and free energy (but notice that the internal energy U is unaffected). Thus the sign of this force is opposite to the $-pdV$ term for a gas. We can then write the energy of the system as

$$dU = TdS + fdL \quad = \quad 0 \qquad (21)$$

and the free energy as

$$dF = -SdT + fdL \qquad (22)$$

so that we find a result for the force, viz.,

$$f = \left(\frac{\partial F}{\partial L} \right)_T \qquad (23)$$

Suppose now we ask - what is the effect of phonons on this result? The internal energy of the system will be modified by the energy $U = U_{vib}$ associated with atomic vibrations.

Thus the total entropy must have contributions from both atomic vibrations and from the polymer configuration; we have

$$S = S_{vib} + S_{con} \qquad (24)$$

However the vibrational entropy has no explicit dependence on L, so that

$$\left(\frac{\partial S_{vib}}{\partial L} \right)_T = 0 \qquad (25)$$

and thus we have

$$\left(\frac{\partial S}{\partial L}\right)_T = \left(\frac{\partial S_{con}}{\partial L}\right)_T \qquad (26)$$

Since we are at finite temperature, and the problem is clearly one governed by entropy, it makes to consider the free energy $F = U - TS$, and thus we can write:

$$f = \left(\frac{\partial F}{\partial L}\right)_T = -T\left(\frac{\partial S_{con}}{\partial L}\right)_T = \frac{k_B T}{N l^2} L \qquad (27)$$

This is a very interesting result – notice that it has the form of Hooke's Law, viz., $L = Yf$, where Y is Young's modulus. We see that the restoring force is driven by the tendency to increase entropy and thereby minimize the free energy – it has nothing to do with any internal forces between the elementary constituents of the system. The entropy driven force, or "entropic force", increases with temperature, and it resists the external force f. Since we have no internal entropy here, the polymer will tend to collapse at any finite T. If we had repulsion between the bonds, or an energy associated with bending of the bonds (which will certainly exist in a real polymer), then this will of course change.

What happens if one now stretches the polymer by a small amount L? Let us do this isentropically in a vacuum. Then the internal energy increases by an amount:

$$\Delta U_{vib} = f\Delta L \qquad (28)$$

The total entropy is unchanged so

$$\Delta S_{vib} = -\Delta S_{con} = \frac{k_B L \Delta L}{N l^2} \qquad (29)$$

Note ΔS_{vib} is positive since ΔS_{con} is negative. Thus the work is converted into vibrational energy and the entropy of configuration is converted to vibrational entropy.

If you think about this for some concrete system like an elastic band, then we see that in the absence of the vibrational contribution (ie., ignoring local vibrations in the molecule), we must introduce an external force to stretch the band, solely in order to overcome the entropic force which tries to collapse the band No work is done in this process against any internal forces, nor do we raise the energy associated with different configurations – they all have the same energy.

On the other hand the different vibrational modes DO have different energies (physically we excite the modes to higher energies), and the excited states have higher entropy (there are more of them, as we will see in detail in the next section). So when entropy is passed from configurational to vibrational modes, we find that in pulling on the string/elastic band, we are actually doing work to raise the energy of the vibtational modes.

You might care to think about the following question:

What happens if the polymer is stretched isothermally? What is the change in the internal energy and what is the change in the configurational and vibrational entropies?