# 4. The CANONICAL DISTRIBUTION

MAIN TOPIC: The canonical distribution function and partition function for a system in contact with a heat bath. The connection with thermodynamics, and the use of this distribution to analyze simple models.

4(a) Derivation of Canonical Distribution

In the discussion of the microcanonical distribution we looked at a total system that was completely isolated. We now turn to a different kind of setup, in which the system of interest Σ, which we will call the 'central system', is coupled to a 'heat bath' or thermal environment Σ_{env}, which is assumed to be very large – far larger than Σ. As a result the density of energy levels in the central system, at any given energy, is far less than it is for the environment (recall our results for the density of states for N particles moving in a box of size L, or for N oscillators). Roughly speaking we have the situation shown in the figures.
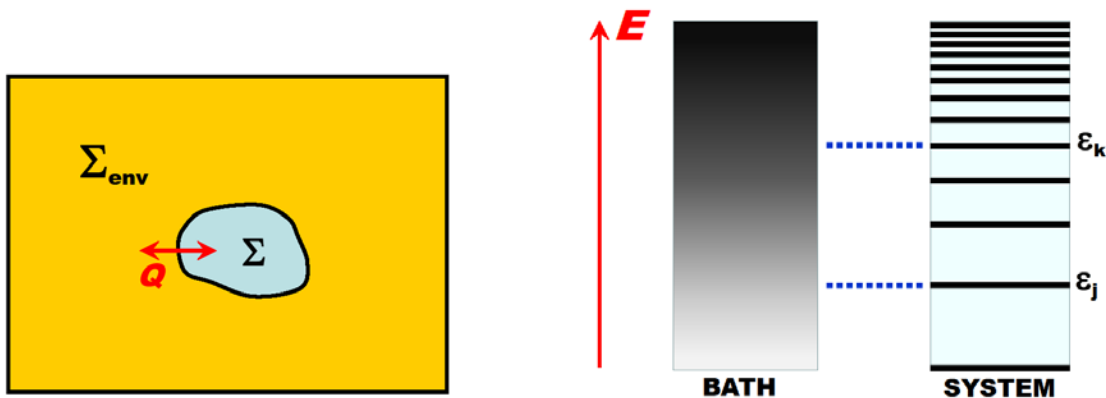
(a) System plus Bath. (b) Energy levels for each

We allow heat and other forms of energy to move between system and bath, but no matter (eg., particles) can pass between them. Now consider what the two of these look like when they are isolated from each other. In the right figure we show some of the energy levels for the system – these are very sparsely distributed compared to those for the bath, which are so densely packed in energy that we simply show them as a continuum, with the density of shading representing the density of states.

Now we want to derive the probability that we will find the system in some state of energy ε_j, under the assumption that the surrounding bath is at temperature T. I will give 2 different arguments for the result – there are many more.

Argument 1: We know that the total energy E_0 of the system is conserved. So we can define the set of energies $\{\, \mathcal{E}_k \,\}$ of the bath Σ_{env}, and ε_j for the central system Σ, so that we have a total energy

$$E_0 \;=\; \mathcal{E}_k \,+\, \varepsilon_j \;=\; \text{const} \quad \text{(for all allowed pairs k, j).} \tag{1}$$

where we include all possible states j, k of the system and bath respectively that obey eqtn. (1).

Now the bath is by hypothesis vast compared to the central system, and so for any given value of E_0 we expect that almost all of the thermal energy will be in the bath Σ_{env}, because there are so many more modes in the bath than in the central system (imagine, eg., a pollen grain in a glass of water). It is therefore a very good approximation to assume that

$$\varepsilon_j / E_0 \;<<\; 1 \tag{2}$$

for any of the relevant central system levels.

Now let us consider the combined state Σ_0 of the central system plus bath. This system is at a fixed energy, but of course there will be many microstates $\Omega_0 (E_0)$ corresponding to this energy, and the system is equally likely to be in any of them. All these different states all correspond to some pairing of states j and k from the central system and bath combined – in fact we clearly have

$$\Omega_0 (\mathcal{E}_k , \varepsilon_j) \;=\; \Omega(\varepsilon_j)\, \Omega_{env} (\mathcal{E}_k)\; \delta\, (E_0 - (\varepsilon_j + \mathcal{E}_k)) \tag{3}$$

where the delta-function which is enforces the constraint in eqtn (1).

The next thing we observe is that the multiplicity $\Omega_{env} (\mathcal{E}_k)$ for the bath (ie., its many-body density of states $\boldsymbol{N}(\varepsilon)$) is far larger than that of the central system (see figure on previous page); and moreover, it increases with energy incredibly fast (recall that the many-body density of states at energy E goes roughly as E^{N-1}, where N is the number of degrees of freedom of the system). Thus we can immediately see that the most probable energy for the bath energy $\mathcal{E}_k$ is just its maximum possible value E_0, ie., the energy of the total "system + bath" Σ_0.

This argument seems to give nonsense – it says that the most probable value for the energy ε_j of the central sub-system is zero! However this is actually correct. What counts now is – how does the probability of occupation for the central sub-system *vary as a function of energy away from zero*? To find this out, we must expand away from the maximum probability.

To do this we could just use Stirling's approximation – but it would have to be done with a modification using Lagrange multipliers, to implement the constraints in (1) and (3). We we will come to this method later (in an Appendix), and employ here the simpler expedient of asking how the log of the multiplicity varies with energy away from E_0. This is easy, because we already know how ln Ω varies with energy, from the usual result

that T = dU/dS, from eqtn (1.11) or (3.6) (using in addition the Boltzmann relation between S and Ω, viz., that S = k_B ln Ω, to relate the two). We can therefore write

$$\ln \Omega_{env}(E) \quad = \quad \ln \Omega_{env}(E_0) + (d \ln \Omega_{env} / d\, E)\big|_{E=E0} \;\; (E_0 - E) + \; .... \tag{4}$$

$$= \quad \ln \Omega_{env}(E_0) - \beta\varepsilon + \; .... \tag{5}$$

where we use the fact that the energy of the central sub-system here is just $\varepsilon = E_0 - E$, if the bath energy is E. As before, $\beta = 1/k_B T$, the inverse temperature.

Now because over the tiny energy range involved here, the multiplicity of the central system is hardly varying, this result (5) is also giving us the probability that the central system is in a state with energy ε. In fact, what we can say is the joint probability that the bath will have energy E and the central system the corresponding energy $\varepsilon = E_0 - E$ is just given by the multiplicity $\Omega_{env}(E) = \Omega_{env}(E_0 - \varepsilon)$. So we immediately get, by exponentiating back the result in eqtn. (5), that for the central system the probability of having energy ε is just proportional to exp[-$\beta\varepsilon$]. This is a really important result. Let us rewrite it in terms of the discrete energies ε_j that we had before – we then get

$$P(\varepsilon_j) \quad = \quad \exp[-\beta\varepsilon_j] \, / \, Z \tag{6}$$

$$Z(\beta) \quad = \quad \Sigma_j \; \exp[-\beta\varepsilon_j] \tag{7}$$

where Z(β) is usually called the canonical partition function, and is just acting as a normalizing factor here, so that the probabilities sum to unity.

Argument 2: Assume to start off, to make things simple, that the central system Σ has just two possible energies E_1 and E_2 separated by ΔE (here I change the notation somewhat, dropping the symbols ε_j for the central system energies). Let p_1 and p_2 be the probabilities that Σ has energy E_1 and E_2 respectively. Suppose that if Σ has energy E_1 then the entropy of Σ_{env} is S_1; the entropy in Σ will be 0 (the state has a multiplicity of 1), so that the total entropy is also S . Then the *total* multiplicity of the combined system plus bath is :

$$\Omega_1 \; = \; \exp[S_1 / k_B T] \tag{8}$$

If on the other hand Σ has energy E_2 in state 2 then an amount of energy ΔE is given up to the bath Σ_{env}, and the entropy increases to $S_1 + \Delta E/T$. The total multiplicity for system plus bath is now:

$$\Omega_2 \; = \; \exp[S_1/k_B + \Delta E/k_B T] \tag{9}$$

and consequently the ratio of the 2 probabilities for occupation must be

$$\frac{p_1}{p_2} = \frac{\Omega_1}{\Omega_2} = \exp[-\Delta E / k_B T] = \frac{\exp[-E_1 / k_B T]}{\exp[-E_2 / k_B T]} \tag{10}$$

One can immediately generalize this argument to a to a microsystem with many energy levels. The above argument applies to any pair i,j of energy levels, so we must have:

$$\frac{p_i}{p_j} = \frac{\exp[-E_i / k_B T]}{\exp[-E_j / k_B T]} \tag{11}$$

Essentially the lower energy states have a higher probability to be occupied because this implies more energy in Σ_0 (given total energy is conserved) and therefore a larger entropy (and multiplicity) of the *total system*. It follows again that:

$$\boxed{p_i = Z^{-1} \exp[-E_i / k_B T] = Z^{-1} \exp[-\beta E_i]} \tag{12}$$

where again the normalization constant – the canonical partition function - is given by

$$\boxed{Z = \sum_i \exp[-\beta E_i]} \tag{13}$$

This result is absolutely central in statistical mechanics – along with the Boltzmann result that $S = k_B \ln \Omega$, it is the most important result in the whole subject.

   Note that the sum in (13) is over all states *and not energies*. One can transform it to a sum over energies by including a degeneracy factor, g_i, which is the number of states with energy E_i. We then have

$$Z = \sum_i g_i \exp[-\beta E_i] \tag{14}$$

Now let's summarize here what we have found. We see that in the canonical ensemble, the temperature *T* has acquired a much more profound physical meaning. It is telling us how – if the system is in thermodynamic equilibrium - the energy scale over which the probability density falls off as we increase the energy away from the ground state. We see now what the idea of negative temperatures really means – a system with T < 0 has a probability density which is increasing exponentially as we increase *E*.

We see also that the system is only in thermodynamic equilibrium if P(E) has this exponential dependence – any other distribution implies the system is not in equilibrium.

4(b)   Properties of Canonical Partition Function

At first glance the canonical partition function Z(β) in (13) seems to be of no real importance except as a normalizing factor for the canonical probability in (12). However this is deceptive, and we will not see that it is of central importance.

Thermodynamic Quantities: Let us first notice first that if you know Z you can calculate all the thermodynamic functions, intensive and extensive. A simple but key example is the mean energy, given by:

$$\langle E \rangle = \sum_i p_i E_i = Z^{-1} \sum_i E_i \exp[-\beta E_i] = -Z^{-1} \frac{\partial Z}{\partial \beta} = -\frac{\partial \ln Z}{\partial \beta} = k_B T^2 \frac{\partial \ln Z}{\partial T} \tag{15}$$

Rewriting this differently, we can say that the energy U, which appears as an extensive quantity in thermodynamics, is also given directly by the partition function as

$$U = -\left(\frac{\partial \ln Z}{\partial \beta} \right)_V = k_B T^2 \frac{\partial \ln Z}{\partial T} \tag{16}$$

We see that the key point here is that we can extract a thermodynamics quantity like U from a derivative of the *logarithm* of the partition function.

Free Energy function F: Let us now follow this up in a more general way. First, let us recall that the internal energy U is also given by

$$U = \left(\frac{\partial(\beta F)}{\partial \beta} \right)_V \tag{17}$$

Now, equating the RHS of (16) and (17), and then integrating over β (or T), we get $\beta F = -[\ln Z + c]$, so that we have

$$F = -k_B T[\ln Z + c] \tag{18}$$

where the constant c is independent of T (or β). In other words the free energy F and the quantity $-k_B T \ln Z$ agree to within some constant (independent of T) . Now recall the entropy S can also be written in terms of the free energy, and therefore in terms of the partition function Z; one has

$$S = -\left(\frac{\partial F}{\partial T} \right)_V = k_B \left[T \frac{\partial \ln Z}{\partial T} + \ln Z + c \right] = k_B \left[-\beta \frac{\partial \ln Z}{\partial \beta} + \ln Z + c \right] \tag{19}$$

where we have used the simple identity:

$$\frac{\partial}{\partial T} = \frac{\partial \beta}{\partial T} \times \frac{\partial}{\partial \beta} = -\frac{1}{k_B T^2} \times \frac{\partial}{\partial \beta} \tag{20}$$

Now we can find the constant of integration c, since as noted it does not depend on T. Let us therefore evaluate c at $T = 0$. In the limit where $T \to 0$ only the ground state of the many particle system is occupied and thus the partition function $Z \to \Omega_0 \exp[-E_0/k_B T]$, where Ω_0 is the multiplicity of the ground state. (if we are dealing with a single particle partition function then $Z \to g_0 \exp[-E_0/k_B T]$, where g_o is the degeneracy of the ground state). Substituting $Z = \Omega_0 \exp[-E_0/k_B T]$ into Eq. 3 leads to

$$S = k_B (E_0/k_B T + \ln \Omega_0 - E_0/k_B T + c)$$

Implying that $c = 0$. Thus we have the fundamental result that

$$\boxed{F = -k_B T \ln Z} \tag{21}$$

which also immediately gives us the final result for the entropy S, by putting c = 0 in eqtn. (19) above.

We can also easily show that the free energy must be a minimum, starting from the partition function. Consider again our system Σ in contact with a heat bath. It can be in many different macrostates, each with a different energy U and multiplicity Ω. The probability that any particular macrostate is occupied is proportional to the Boltzman factor weighted by the multiplicity i.e. $\Omega \exp[-\beta U]$. However $\Omega = \exp[S/k_B]$, so this probability is proportional to $\exp[S/k_B - \beta U] = \exp[\beta(TS - U)] = \exp[-\beta F]$. Thus the most likely macrostate corresponds to the minimum in F since this probability is just $\exp[-\beta F]$.

Entropy function S: Using our result for Z we can also re-derive a key result that we established in the last chapter, using arguments based on the microcanonical distribution. (compare eqtn. (3.14) of the last chapter). Let us, using eqtn. (13), try to rewrite the entropy in terms of the probabilities $P_i = Z^{-1} \exp[-\beta E_i]$ for the system to be in the ith microstate. The manipulations are straightforward, and we have

$$\sum_i P_i \ln P_i = \sum_i Z^{-1} \exp[-\beta E_i][-\beta E_i - \ln Z \;]$$

$$= -\beta Z^{-1}\left(\sum_i E_i \exp[-\beta E_i]\right) - Z^{-1}\ln Z \sum_i \exp[-\beta E_i]$$

$$= \beta Z^{-1}\frac{\partial Z}{\partial \beta} - \ln Z$$

$$= \beta \frac{\partial \ln Z}{\partial \beta} - \ln Z$$

Comparing this with Eqtn. (19) after setting $c = 0$ we immediately get the result

$$S = -k_B \sum_i P_i \ln P_i \qquad (22)$$

Which is consistent with what we found in eqtn. (3.14) of the last chapter, but where we now, in the canonical distribution, have an explicit form for the probabilities (in eqtns. (6) and/or (12).

Note this is result for the entropy has been derived in a more general way than in the last chapter, since it applies to systems in contact with a heat bath where the probabilities for a given microstate to be occupied are not all equal. In fact, we can easily see that the microcanonical result in eqtn. (3.14) is just a special case of the result here. If our system is isolated, then all microstates will have the same energy and thus the same Boltzman factor. In this case $P_i = 1/\Omega$, as we saw in the last chapter, and then eqtn. (22) reduces to the expression for an isolated system, ie:

$$S = -k_B \sum_i \frac{1}{\Omega}\ln\frac{1}{\Omega} = k_B \Omega^{-1}\ln\Omega \sum_i \; = k_B \ln\Omega \qquad (23)$$

which is just eqtn. (3.13) of the last chapter.
   One can continue in this vein – using our knowledge of thermodynamics and how different extensive and intensive variables are related to the free energy, etc., we can clearly derive a very large number of expressions for thermodynamic functions and their derivatives in terms of the partition function.

Energy Fluctuations: Up to now we have simply been relating results for equilibrium thermodynamic quantities to the partition function and its derivatives. However we can step outside the framework of ordinary thermodynamics in a very useful way, by considering fluctuations about equilibrium quantities.

As an example, consider a central system in thermal equilibrium with a heat reservoir. The temperature is well defined, according to the way we have defined the canonical ensemble, but the energy of the sub-system will fluctuate about the mean value $<E> = U$

as energy is exchanged between the central system and the reservoir. We can define the magnitude of these fluctuations by looking at their variance, ie., by looking at the mean squared deviation in energy:

$$\langle \Delta E^2 \rangle = \langle E^2 \rangle - \langle E \rangle^2 \tag{24}$$

Let us compute this – we have

$$\langle E \rangle = \sum_i p_i E_i = Z^{-1} \sum_i E_i \exp[-\beta E_i] = -Z^{-1} \frac{\partial Z}{\partial \beta} \qquad \text{as before, and}$$

$$\langle E^2 \rangle = \sum_i p_i E_i^2 = Z^{-1} \sum_i E_i^2 \exp[-\beta E_i] = Z^{-1} \frac{\partial^2 Z}{\partial \beta^2} \tag{25}$$

Thus the mean squared deviation in energy is given by

$$< \Delta E^2 > \; = Z^{-1} \frac{\partial^2 Z}{\partial \beta^2} - Z^{-2} \left[\frac{\partial Z}{\partial \beta} \right]^2 = -\frac{\partial \langle E \rangle}{\partial \beta} \tag{26}$$

Note that this is just the specific heat at constant volume – this we have derived the very interesting result that

$$< \Delta E^2 > \; = \; k_B T^2 \, C_V(\text{T}) \tag{27}$$

Just as we did above for thermodynamic quantities, we extend this idea to can look at the fluctuations of many different intensive and extensive variables.

Multiple sub-Systems: All of the above is valid for one single central system coupled to a thermal bath. We can easily generalize to many *identical but distinguishable* sub-systems. As a very simple example, consider N non-interacting spin ½ particles on lattice in a magnetic field. Then we can write the partition functions for one, two, and three spins respectively, as

$$Z_1 = \exp[-\beta E_\uparrow] + \exp[-\beta E_\downarrow]$$

$$Z_2 = \exp[-2\beta E_\uparrow] + 2\exp[-\beta(E_\uparrow + E_\downarrow)] + \exp[-2\beta E_\downarrow] = Z_1^2$$

$$Z_3 = \exp[-3\beta E_\uparrow] + 3\exp[-\beta(2E_\uparrow + E_\downarrow)] + 3\exp[-\beta(E_\uparrow + 2E_\downarrow)] + \exp[-3\beta E_\downarrow] = Z_1^3$$

where we note that the expansion coefficients are binomial coefficients – this iterative process can be continued indefinitely.

In general it is clear that the partition function for N non-interacting *distinguishable* sub-systems will just be

$$Z_N = Z_1^N \tag{28}$$

where Z_1 is the partition for one of the sub-systems.

To show this, it is sufficient to show it is true for a pair of sub-systems A and B; ie, that $Z_{A+B} = Z_A Z_B$, where Z_A is the partition system for one sub-system (for, e.g. N_A particles) and Z_B is the partition function for the other sub-system (for, eg., N_B particles). The only condition required here is that the two systems are distinguishable, and that they are independent (so that they are not interacting with other, or correlated in any way). Otherwise it is irrelevant what the 2 sub-systems are made of, or whether there are internal interactions in each one. Then we have

$$Z_A Z_B = \sum_j \exp[-\beta E_j] \sum_\alpha \exp[-\beta \mathcal{E}_\alpha] \rightarrow \sum_{j,\alpha} \exp[-\beta(E_j + \mathcal{E}_\alpha)] = Z_{A+B} \tag{29}$$

where j and α label the (in general quite different) eigenstates of each sub-system respectively. It is obvious that the same argument applies when we have a set of N different non-interacting sub-systems. Notice that whenever we have a set of non-interacting sub-systems, we find that thermodynamic quantities add, as they should. Thus, eg., the mean total energy is given by

$$U = \langle E_{tot} \rangle = -\frac{\partial \ln Z_1^N}{\partial \beta} = -N \frac{\partial \ln Z_1}{\partial \beta} = N \langle E_1 \rangle \tag{30}$$

as expected.

Note the key implication of all of this. It is that we can analyze the thermodynamics of the N-particle system, if the particles are not interacting, by only considering the thermodynamics of the single-particle sub-units. We don't have to worry about the N-particle density of states any more, only about the far simpler 1-particle density of states.

The question of what happens if the sub-systems are interacting with each other is one that can be developed in great detail. It is clear that the results in (28)-(30) will no longer be valid, because both the energies and the occupation probabilities in a given sub-system will depend on the state of the other sub-systems. If the interactions are weak then one can use perturbation theory to treat the effect of the interactions. The general effect of interactions – which are always present – is the central theme of most of condensed matter physics, as well as of gas and plasma theory, and of statistical mechanics.

4(c)   Some Simple Examples

The examples given here illustrate the kind of thing we are talking about. None of them are complicated because they involve non-interacting sub-systems; the only one with

significant algebra is the set of non-interacting SHO systems. However, the discussion of them using the canonical ensemble is much more physically realistic than that using the microcanonical ensemble.

(i) 2-level Atom: The simplest possible system is of course s single 2-level system (TLS), or qubit. This is sometimes called the "2-level atom" by people in atomic physics. However, we can complicate the problem a little by having non-zero degeneracy for the 2 levels, so that it involves more than 2 states (ie., it is no longer a spin-1/2 or qubit).

   We consider an atom with a ground state degeneracy g_0 and first excited state degeneracy $g_1 > g_0$, and an energy gap E_0 between the 2 levels. Ignoring higher excited states, the probability for the atom in thermal contact with a heat bath to be in the excited state is:

$$p_1 = \frac{g_1 \exp[-\beta E_0]}{g_0 + g_1 \exp[-\beta E_0]} \quad = \quad g_1 \exp[-\beta E_0] / Z(\beta) \tag{31}$$

whereas the probability to be in the ground state is

$$p_0 = g_0 / Z(\beta) \tag{32}$$

which tends to unity as $T \to 0$. Notice how if we have a higher degeneracy in the excited state, this increases the probability for excitation out of the ground state at finite T.

(ii) N non-interacting Qubits: We can now immediately process to discuss the partition function for a spin ½ magnet with N spins in a magnetic field, ie., for N qubits. The partition function for a single qubit is just a special case of that for the 2-level atom, with $g_0 = g_1 = 1$, and we thus have

$$Z_1 = \exp[\beta \mu B] + \exp[-\beta \mu B]$$

and
$$Z = Z_1^{N} \tag{33}$$

From this result we can immediately derive the internal energy U and heat capacity C_V for the system; for the energy we get

$$U = -NZ_1^{-1} \frac{\partial Z_1}{\partial \beta} = -N\mu B \frac{\exp[\beta \mu B] - \exp[-\beta \mu B]}{\exp[\beta \mu B] + \exp[-\beta \mu B]} = -N\mu B \tanh(\beta \mu B) \tag{34}$$

and for the heat capacity we find

$$C(T) = \left(\frac{\partial U}{\partial T}\right)_B = -k_B \beta^2 \left(\frac{\partial U}{\partial \beta}\right)_B = k_B N (\mu B \beta)^2 \sec h^2 (\beta \mu B) \tag{35}$$

It is actually very useful to see what these results look like as a function of the temperature T, by plotting them out. One should avoid the tendency to just throw this onto a computer – you learn much more by looking at the asymptotic behavour analytically, and then interpolating.

Starting with the energy U, we see that at high T where $\beta \mu B << 1$ and $\tanh(x) \approx x$ we find a Curie law:

$$U \cong -N\beta(\mu B)^2 = -N\frac{(\mu B)^2}{k_B T}; \qquad \text{(for } k_B T >> \mu B \text{)} \tag{36}$$

whereas at low T (large x) $\tanh(x) \approx 1$ so that

$$U \approx -N\mu B \qquad \text{(for } k_B T << \mu B \text{)} \tag{37}$$

Both of these results are easy to understand physically. At high T, the up and down states are almost equally occupied (at infinite temperature, they are equally occupied), and the deviation of U from zero is just caused by the tiny difference between the population of up and down states. At low T, the system polarizes in the field, so that the energy is just that of a set of aligned spins.

If we then interpolate between these 2 asymptotic results, we find the figure shown:
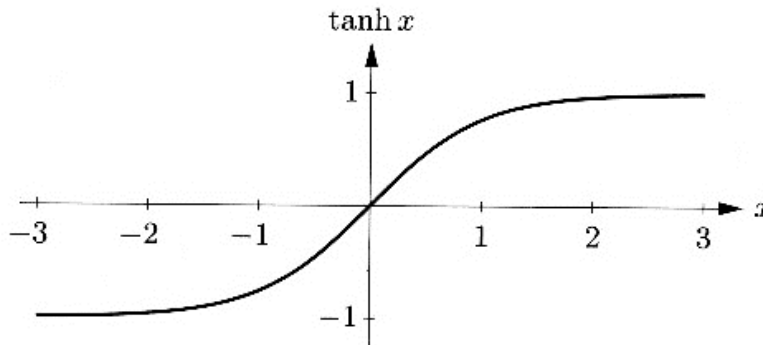

Note that we could have plotted this figure simply by interpolating smoothly between the 2 asymptotic results, without knowing what a tanh function looks like.

Turning now to the heat capacity in (35), we note at high temperatures the heat capacity falls as $1/T^2$ because the energy is changing very slowly (the system is almost completely disordered, and it is hard to add any more entropy). At low T it falls again,

this time because the system is tending towards saturation of the polarization. Thus a peak must occur in between – a detailed analysis shows that this peak occurs when $k_B T / \Delta E = 0.42$ where $\Delta E = 2\mu B$ is the energy level separation. This peak is called a *Schottky anomaly*, and it is characteristic of the thermal behavior of all two level systems.

Plotting the behaviour shows the shape of the Schottky anomaly (see below):
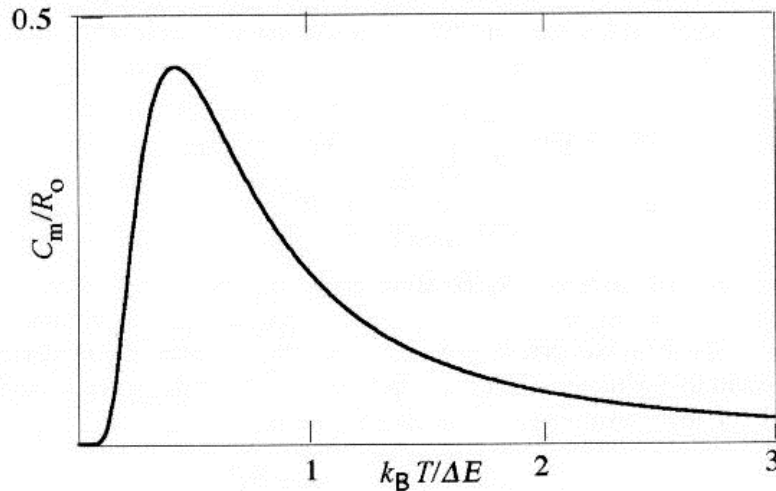
In any real solid, one sees a Schottky anomaly in the low-T specific heat, unless the system is both pure and without disorder. This is because in most solids, the impurities will include spin impurities, behaving like 2-level (or more generally 2S+1-level) systems; and because the disorder will include local regions in which an atom or ion can move back and forth between 2 different positions, thereby creating a 2-level system. As we will see below and in later chapters, the other low-energy excitations in the sample give a quite different form for the specific heat, and the two different contributions add together.

(iii) Zipper Model for DNA unwinding: W consider a toy model for a DNA molecule, which is essentially a "zipper" with N links – we can separate the 2 chains involved in the DNA molecule by pulling them apart, just like a zipper. In real DNA, N will be huge (a DNA molecule can be many cm in length, with n in the hundreds of millions). In the figure we show a very small example., with 5 open links and 5 closed links.
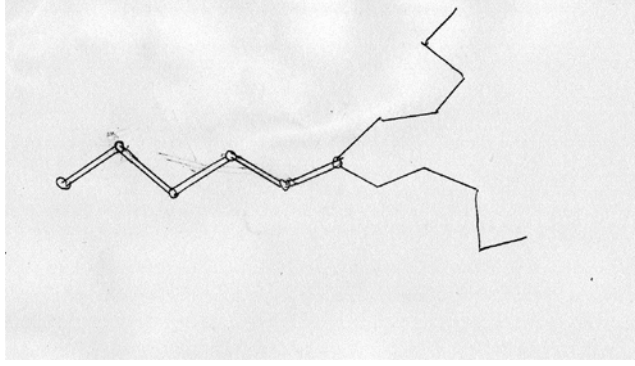
Now, let us assume that opening one link cost an energy E_0 but increases the degeneracy by a factor g_0 due to the increased number of configurations available to the system compared to the situation when the link is closed (these assumptions make physical sense – if there was no binding energy E_0 for links, the double chain would never form in the first place). However the entropy of the unzipped system is clearly greater – there are far more available configurations – and so the system will eventually unzip as we raise the temperature.

Let us suppose that the completely closed zipper has energy $E = 0$. If there are $f(N)$ microstates for a closed zipper, then there will $f(N)g_0$ states with energy , having one link open, there will be $f(N)g_0^2$ states with energy $2E_0$ having to two links open; and so on.

The value of the "degeneracy" g_0 will depend on the allowed configurations for the open links. Thus, for example, if the angle between adjacent links is restricted to be either 0 or 180 degrees only, then $g_0 = 2$, and $f(N) = 2^N$. This is because for the closed chain, each double link in the chain has 2 configurations available, but for open links, there are 4 configuations available (each of the 2 individual chains has 2 possible orientations). Thus the closed zipper has 2^N possible configurations, whereas a zipper with r links open has 2^{N+r} possible configurations . Note that since $f(N)$ is independent of r it does not affect the probability $p(r) = Z^{-1} g_0^r \exp[-r\beta E_0]$ of having r links open.

We can now write down the partition function – we have:

$$Z = \sum_{r=0}^{N} g_0^r \exp[-r\beta E_0] = \sum_{r=0}^{N} [g_0 \exp(-\beta E_0)]^r \tag{38}$$

which when summed just gives

$$Z = \frac{1 - g_0^{N+1} \exp[-(N+1)\beta E_0]}{1 - g_0 \exp[-\beta E_0]} \tag{39}$$

leading to a mean energy of

$$\langle E \rangle = \langle r \rangle E_0 = -\frac{\partial \ln Z}{\partial \beta} \tag{40}$$

(a) **Case 1**: assume that the degeneracy $g_0 = 1$; then the only different between the zipped and unzipped chain is that the latter has higher energy. Let's assume also that the chain is very long, so that $NE_0 \gg k_B T \gg E_0$. Then we have

$$Z \cong \frac{1}{\beta E_0} \qquad \ln Z \cong -\ln \beta - \ln E_0, \qquad (41)$$

so that $\langle E \rangle = k_B T$ and $\qquad \qquad \langle r \rangle = k_B T / E_0 \qquad (42)$

(b) **Case 2**: Now assume $g_0 > 1$. Note that the quantity $g_0 \exp[-\beta E_0]$ must necessarily be less than 1; otherwise $p(r) = Z^{-1} g_0^r \exp[-r\beta E_0]$ increases with r and the partition function $Z = \sum_{r=0}^{N} [g_0 \exp(-\beta E_0)]^r$ diverges as $N \rightarrow \infty$. Physically this simply means that for high enough temperature, there is no stable value of r and the zipper will unwind completely.

Thus we arrive at the conclusion that when $g_0 > 1$, there must be a transition between zipped and fully unzipped behavour, at a critical temperature defined by

$$g_0 \exp[-E_0 / k_B T_c] = 1. \qquad \text{so that} \qquad k_B T_c = E_0/\ln g_0 \qquad (43)$$

At higher temperatures the DNA strand will unwind, independently of its length.

This model is too crude to analyze real DNA, but it does give a very useful result which is actually true – above a critical temperature, a molecule like DNA will dissociate, ie., unzip. Transitions of this kind are very common in biological molecules; another example is the unfolding of proteins molecules, which below a critical temperature will fold up into complex shapes.

(iv) Einstein Model of Phonons: In any solid, the relative positions of the ions, atoms, or molecules will move relative to each other – this movement involves a large number of degrees of freedom which, when quantized, we call phonons. The simplest cases that can be analyzed are either crystalline lattices (where the ionic or molecular constituents are arranged in a regular lattice), or a simple liquid (which is translationally uniform). In either of these cases the phonons can be characterized by simple quantum numbers – if the phonons can move around in the crystal or liquid, these quantum numbers label the momentum k of the phonon, and their frequencies $\omega(k)$ depend on k. In such a system we refer to the phonons as lattice normal modes or liquid normal modes, and $\omega(k)$ is called the dispersion relation for these modes.

I will not give a complete discussion of phonons here. In solids the phonon spectrum is comprised of 2 different sorts of modes. The first is the so-called "acoustic modes", which correspond classically to sound waves of different wavelength, propagating

accross the entire system, and whose frequency $\omega(k)$ goes to zero as $k \rightarrow 0$. (ie., in the limit where the wavelength of the acoustic mode goes to infinity – such modes are called `long wavelength' modes). These will be discussed later on in the course. Their dispersion relation simply reads $\omega(k) = ck$, where c is the sound velocity.

However there is also another set phonon of modes in many solids, which do not propagate through the system, but remain localized around a particular region. These are local vibrations, which depending on what type of system one is dealing with, go under the name of 'vibrons', or 'optical phonons', or 'local phonons'.

One can see intuitively what these modes might look like by considering a large molecule. The entire molecule can distort – long-wavelength flexural modes, or twisting modes, or just straightforward longitudinal compression modes. But one can also have local oscillations in which, eg., a H atom stuck to a C atom vibrates wobbles around on its C-H bond, without these oscillations being transmitted to the rest of the molecule – this is a vibron mode.

In a very simple molecule, like O2, NH3, or HCl, one can isolate the individual local vibrational modes, and they hardly couple to each other. One then gets a very simple set of localized oscillations in the system, each having a different frequency. Thus, eg., in HCl, we only have one vibrational mode, in which the distance between the H and Cl atoms oscillates. The spectrum is then very simple – as we see in the figure. There is no wavelength associated with such modes – they are simply harmonic oscillator modes, with the usual discrete quantum number n, associated with the oscillator level.
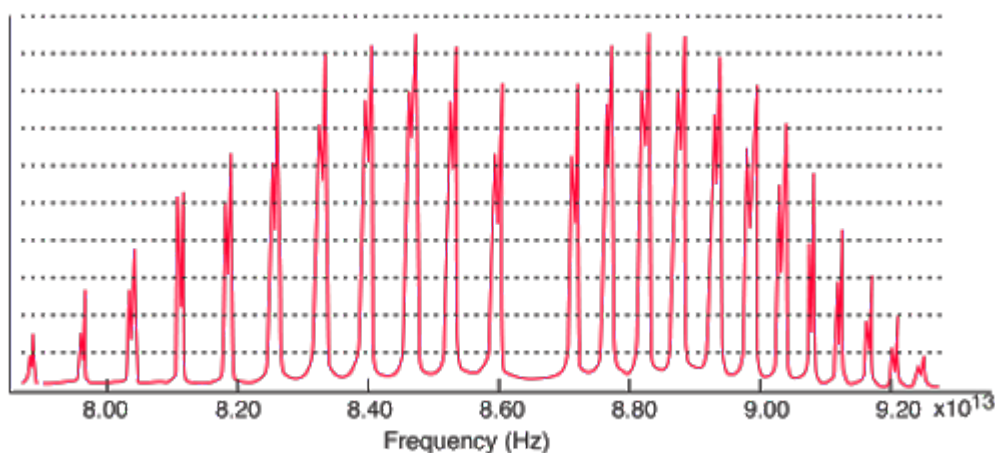

Vibrational spectrum of HCl – there is only a single frequency involved.

However in a more complicated molecule, there will be lots of different kinds of vibrational mode, in different parts of the molecule, all having different frequencies. Some of these will involve vibrations in the distances between ions, others will involve A kind of `wobbling' or bending of bonds between 2 ions, others will be more complicated twisting of configurations involving 3 or more ions. For a sufficiently small

molecule these different modes will still be largely independent of each other, and so we can imagine that the spectrum of the system will look like that of a set of oscillators, all with different frequencies.
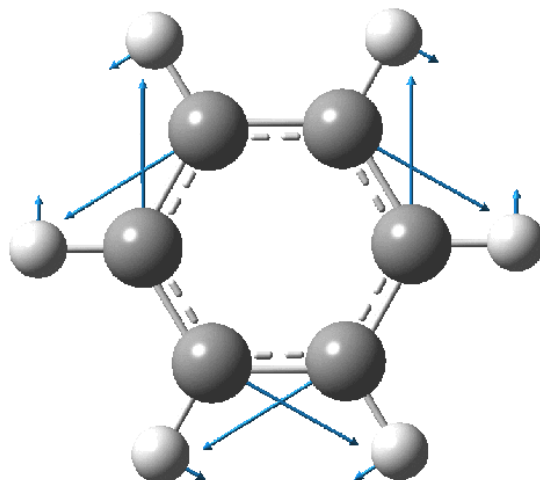
**Vibrational modes ("Vibrons") for Benzene molecule**
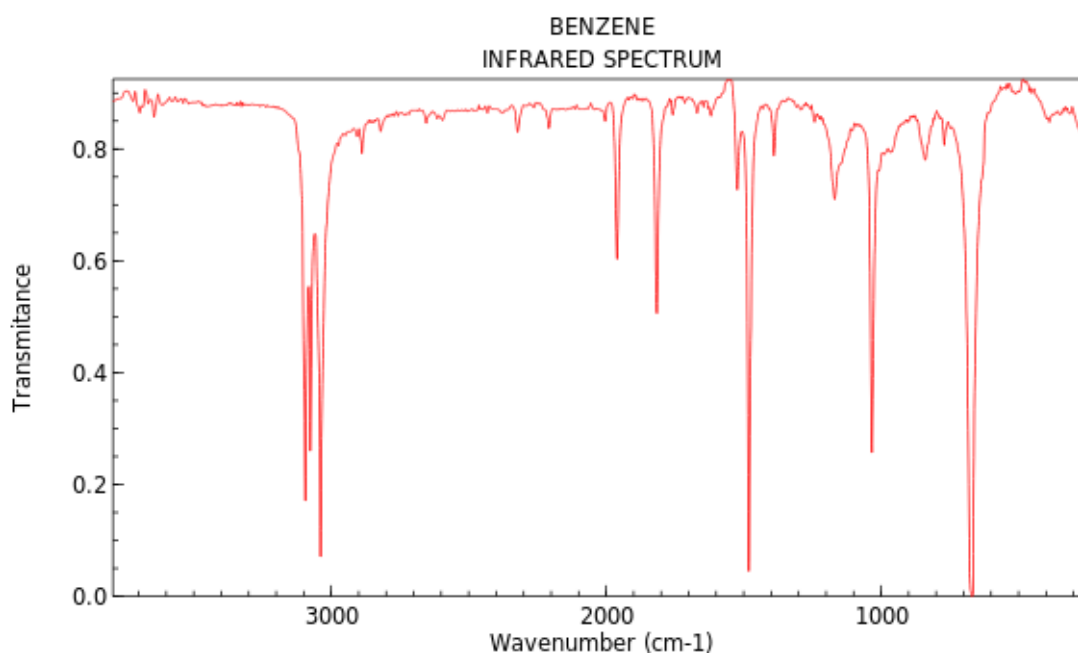
Absorption spectrum of benzene in the IR range.

A nice example of this is provided by the benzene molecule C_6H_6, where things simplify greatly because of the symmetry of the system – many of the different modes then have the same frequency, because they are just the same oscillation, taking place in different parts of the molecule.

It will be clear that in benzene, the different oscillations can't be entirely independent – setting one off will slowly excite the others. Another way of thinking about this is to say that the eigenmodes of the system are not exactly localized, but can travel slowly around the system. We then say that the localized vibrons weakly **hybridize** with each other, to produce spatially extended modes. These modes then have the generic name of **optical modes**, since their frequencies are typically in the optical range.

Now consider a macroscopic crystal, made up of periodically repeated identical cells, each of which may contain one or more atoms. There are of course really simple examples like Na, or He, or Al, in which a single atom occupies a unit cell. There are then only acoustic phonon modes, in which a disturbance of one atoms sets off a wave excitation through the system. However in a crystal like SiO2 (quartz), or NaCl (salt), one has 2 ions per unit cell, and one can have quasi-localized oscillations in which one of the two ions oscillates relative to the other. Clearly the unit cell can be very large (as in, eg., crystals of DNA molecules). In any case, one then expects both acoustic modes and optical modes.

Even in simple systems like crystalline Ge, shown in the figure below, one can have both acoustic and optical modes (and each set can be divided into longitudinal and transverse oscillations, labeled TO and LO for the transverse and longitudinal optical modes, and TA and LA for the acoustic modes). This can happen because Ge crystals have a unit lattice cell with two Ge atoms, not one, and the two atoms are in different local environments (the electric fields around each of them is different). The optical modes are then relative oscillations between each such pair of Ge atoms, inside a unit cell.

Phonon frequency distribution in Ge
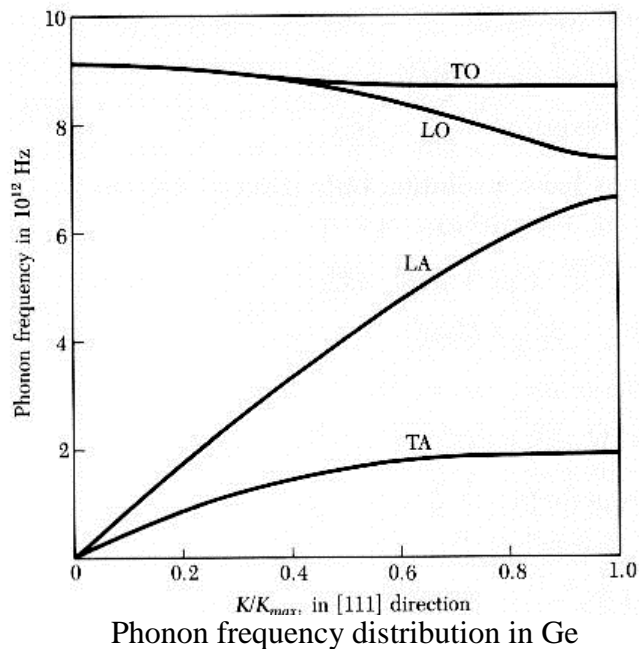
If the optical modes in Ge were genuinely localized, ie., if they only involved local oscillations of atoms, then they would all have the same frequency. However, in the

same way as described above, these 'in-cell' oscillations weakly couple to each other across the crystal, producing the optical modes shown at the top of the graph. As noted previously, the acoustic mode energy goes to zero as $k \rightarrow 0$.

Even before quantum mechanics was discovered, Einstein realized – within the context of the old quantum theory - that a quantized theory of lattice vibrations could have such modes, and he formulated a very simple model for them, in which a system with N atoms is assumed to have *3N local modes of vibration* (3 for each atoms, because we are in 3 dimensions), all with the same frequency. Einstein imagined that each atom oscillated ints own 3-d harmonic potential well. This model then give s a simple picture of uncoupled vibrons.

From the SM point of view this amounts to a collection of *3N* weakly interacting harmonic oscillators, all with the same frequency. It is then easy to analyze. We assume that each oscillator has the same set of eigenenergies, given by

$$E_n = (n + 1/2)h\nu \tag{44}$$

In which ν is the frequency of each oscillator, and where we see that each oscillator energy now involves Planck's constant (here we write $h = h/2\pi$, where h is Planck's constant).

Now the partition function for a single oscillator is clearly just

$$Z = \sum_n \exp[-\beta E_n] = \exp[-\beta h\nu/2] \sum_n (\exp[-\beta h\nu])^n \tag{45}$$

which is easily evaluated to give

$$Z_1(T) \quad = \quad \tfrac{1}{2} \operatorname{cosech}(h\nu/2k_BT) \tag{46}$$

where we write Z_1 to indicate that we deal with a single oscillator. It then follows that the energy and specific heat of the single oscillator mode will be given by

$$U_1(T) \quad = \quad \tfrac{1}{2} h\nu \coth(h\nu/2k_BT) \qquad \text{(single oscillator)} \tag{47}$$

$$C_V(T) \quad = \quad k_B(h\nu/2k_BT)^2 \operatorname{cosech}^2(h\nu/2k_BT) \qquad \text{(single oscillator)} \tag{48}$$

From this we then deduce the corresponding behaviour for the full Einstein crystal; the partition function in (46) is raised to the power *3N*, and the energy and specific heat (which come from ln Z) are multiplied by a factor *3N*. One can, using the same techniques as discussed above, evaluate all the other thermodynamic quantities for the Einstein solid, and then continue on to evaluate the size of fluctuations, etc.

In a later chapter we will discuss acoustic phonons, and phonons in real solids.

(v) N distinguishable particles in a Box: We want to look now at distinguishable particles in some closed container. We already looked at in Chapter 2, for a single particle in a box. We will assume that the box can have dimension 1, 2, or 3, and that it has equal sides L (so that, eg., the 3D box has sides $L \times L \times L$ and volume L^3).

Rather than deal with all the discrete states for the particle, we will go to the limit of very large L, so that the distribution of levels becomes quasi-continuous. Recall that in chapter 2 we defined what I called the "N-particle density of states", which we will write here as

$$\mathcal{N}(E) = \sum_n g(N, n)\ \delta(E - E_n^{(N)}) \qquad (49)$$

where $g(N,n)$ is the degeneracy of the n-th level of the N-particle state, with energy $E_n^{(N)}$ (in other words, $g(N,n) = \Omega(N,n)$, the multiplicity of states). Recall also that the number of these levels rises fantastically as N increases, and so does their density of states (recall, eg., eqtns (2.11), (2.15), and (2.19)).

In what follows we will first look at the density of states of a single particle in a box (in either 1-d, 2d, or 3-d); we then go to N particles in the box, but because the particles are all independent, and the partition function then factorizes into 1-particle contributions, we will find as before that we can stick with the 1-particle density of states. Following convention, we will call the 1-particle density of states $g(E)$, so that

$$g(E) = \sum_n g(N, n)\ \delta(E - E_n^{(N)})\ |_{N=1} \qquad (50)$$

We are going to find the partition function for a particle in a D-dimensional box, in which we let the box size become very large. In this case we will go over to a continuous density of states. In this case the probability $p(E)\,dE$ for a particle to have its energy in the range $(E, E + dE)$, for a system of volume $V = L^D$, is now given in the canonical distribution by

$$p(E)dE = Z^{-1}Vg(E)\exp[-\beta E]dE \qquad (51)$$

where the canonical partition function goes over to

$$Z = \sum_i g_i \exp[-\beta E_i] \approx V\int_0^\infty g(E)\exp[-\beta E]dE \qquad (52)$$

which defines the 1-particle density of states $g(E)$ for this problem.

g(E) **in a Box:** Let's start with a 1-d system. We use eqtn. (2.1) for the particle spectrum. To find the density of states we first note that the separation in wave-vector k between allowed states is $\Delta k = \pi/L$, and since

$$k = \frac{(2mE)^{1/2}}{\hbar} \qquad (53)$$

It then follows that

$$g_{1D}(E) \quad = \quad 2\pi \, (m/2E)^{1/2} \, /\pi h \tag{54}$$

In 2 dimensions we have a 2-d wave vector $\vec{\kappa} = (\kappa_x, \kappa_y)$, with allowed energy levels at energies of

$$E_{nm} = \frac{\hbar^2(\kappa_x^2 + \kappa_y^2)}{2m} = \frac{(\pi\hbar)^2}{2mL^2}\left(n^2 + l^2\right) \tag{55}$$

where n and l are positive integers. Each state occupies an area of κ-space of π^2/L^2, so that the total number $N(k)$ of states with a magnitude $|\vec{\kappa}| = \sqrt{\kappa_x^2 + \kappa_y^2}$ less than some given k is the allowed area in κ-space ($\pi k^2/4$) divided by the area occupied by each state (π^2/L^2), ie., we have

$$N(k) = \frac{\pi k^2}{4}\frac{L^2}{\pi^2} = \frac{2mE}{\hbar^2}\frac{L^2}{4\pi} \tag{56}$$

and since the total number of states per unit area in k-space is $N(k)/L^2$, we just get the density of states as the derivative of this, ie., we have

$$g_{2D}(E) = \frac{m}{2\pi\hbar^2} \tag{57}$$

which is independent of energy (this is a key point in 2-d electronic devices).

Without going through the details, we can see that in 3 dimensions a box of unit volume will have $N(k) = k^3/6\pi^2$ states up to momentum k, and converting to energy as above, and again differentiating with respect to E, we get the 3d density of states as

$$g_{3D}(E) = \frac{1}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} E^{1/2} \tag{58}$$

which increases like $E^{1/2}$. We can plot these as shown:
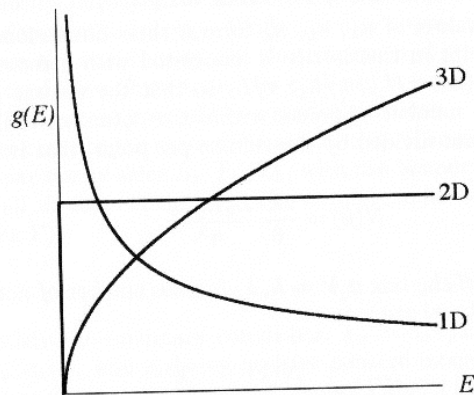
1-particle density of states for particle in 1, 2, and 3 dimensions.

Partition Function: We can now find the partition function using (52). Let's just look at the 3D system; e then immediately have, using (58), the result for the single particle partition function:

$$Z_1 = V \int_0^\infty g_{3D}(E) \exp[-\beta E] dE = \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int_0^\infty E^{1/2} \exp[-\beta E] dE$$

$$= \frac{V}{4\pi^2} \left(\frac{2m}{\beta\hbar^2} \right)^{3/2} \int_0^\infty x^{1/2} \exp[-x] dx \tag{59}$$

And if we evaluate the integral we get

$$Z_1 = \frac{V}{4\pi^2} \left(\frac{2m}{\beta\hbar^2} \right)^{3/2} \frac{\pi^{1/2}}{2} = V \left(\frac{m}{2\pi\beta\hbar^2} \right)^{3/2} = V \left(\frac{m}{2\pi\hbar^2} \right)^{3/2} \beta^{-3/2} \tag{60}$$

and this gives a mean energy per particle of:

$$\langle E \rangle = -Z^{-1} \frac{\partial Z}{\partial \beta} = \frac{3}{2\beta} = \frac{3}{2} k_B T \tag{61}$$

Now let us turn to the set of N particles. By hypothesis the particles are distinguishable, and we will moreover assume that an arbitrary number of them can occupy the same momentum state (since they do noit interact, and they are particles of zero spatial extent, this makes sense). Accordingly we just have $Z_N = Z_I{}^N$, as usual. We can then write the total energy for a set of N atoms in the 3D box as

$$U = \frac{3}{2} N k_B T \tag{62}$$

with associated specific heat

$$C_V = \frac{3}{2} N k_B \tag{63}$$

These last 2 results are interesting – an important chapter in the history of physics is tied up with them, because they illustrate is a very simple way the principle of "*equipartion of energy*". The thermal energy in the system is equally shared amongst all the different translational degrees of freedom, with an amount $k_B T/2$ for each one. By examining the specific heat we can then determine experimentally how many degrees of freedom there are at a given temperature – in the gas of indistinguishable particles we look at here, this number does not depend on T at all. The discovery in experiments over a century ago that the specific heat does vary with T in real solids was a key ingredient in the discovery of quantum mechanics.

One can interpret the 1-particle partition function in various ways. One interesting way is to define a quantity called the "quantum volume", which is the volume below which discrete quantum effects start to become important. For the 3D gas here this is defined as

$$V_q \equiv \left(\frac{2\pi\beta\hbar^2}{m} \right)^{3/2} \tag{64}$$

Thus it is closely related to the thermal de Broglie wavelength, which is

$$\lambda = \frac{h}{(3mk_BT)^{1/2}} \tag{65}$$

since we see that $V_q \sim (0.7\,\lambda)^3$. We can then write the partition function as

$$Z_1 = V/V_q = V\rho_q \tag{66}$$

where the quantity $\rho_q \equiv 1/V_q$ is sometimes called the quantum concentration. Now the key point here is that at temperature T, the uncertainty in energy caused by quantum fluctuations over an energy range k_BT is associated with a position uncertainty, or thermal wavepacket size, of order the thermal de Broglie wavelength λ; the partition function, for a system of unit volume, is just the number of thermal de Broglie cells one can fit into this volume.

One should note a counter-intuitive aspect of the result for Z here. This is that the partition function (60) or (66) is not exponential in volume, but linear in it, so that the free energy of the system looks like

$$F = -Nk_BT[3/2 \ln \rho_q + \ln V] \tag{67}$$

and is logarithmic in V (but linear in N). This seems paradoxical – we saw previously that if we "double a system", then the free energy doubles, and the new partition function is a product of the 2 partition functions for each, ie., it is squared.

However we note here that we are not in fact doubling the system if we increase the volume by a factor of two – we are instead spreading out the same number N of particles in twice the volume. This is not the same – we still have only N particles, not $2N$ particles, and the energy levels and states are not the product over those in each volume either. They are still the same single-particle levels, although in a volume $2V$ their energies have decreased (as has the spacing between the eigenstates).

To see that there is nothing wrong with the result (67), let's note first that the energy in (62) and the free energy in (67) are still additive in N. However, even more convincing an argument comes if we calculate the pressure in the system. This is given by

$$p = -\left(\frac{\partial F}{\partial V} \right)_T \rightarrow \left(\frac{Nk_BT}{V} \right) \tag{68}$$

which is just the usual ideal gas law $pV = RT$, with $R = Nk_B$.