

Section 2

Introduction to Statistical Mechanics

2.1 Introducing entropy

2.1.1 Boltzmann's formula

A very important thermodynamic concept is that of entropy S . Entropy is a function of state, like the internal energy. It measures the relative degree of order (as opposed to disorder) of the system when in this state. An understanding of the meaning of entropy thus requires some appreciation of the way systems can be described microscopically. This connection between thermodynamics and statistical mechanics is enshrined in the formula due to Boltzmann and Planck:

$$S = k \ln \Omega$$

where Ω is the number of microstates accessible to the system (the meaning of that phrase to be explained).

We will first try to explain what a microstate is and how we count them. This leads to a statement of the second law of thermodynamics, which as we shall see has to do with maximising the entropy of an isolated system.

As a thermodynamic function of state, entropy is easy to understand. Entropy *changes* of a system are intimately connected with heat flow into it. In an infinitesimal reversible process $dQ = TdS$; the heat flowing into the system is the product of the increment in entropy and the temperature. Thus while heat is not a function of state entropy is.

2.2 The spin 1/2 paramagnet as a model system

Here we introduce the concepts of *microstate*, *distribution*, *average distribution* and the relationship to *entropy* using a model system. This treatment is intended to complement that of Guenault (chapter 1) which is very clear and which you must read. We choose as our main example a system that is very simple: one in which each particle has only two available quantum states. We use it to illustrate many of the principles of statistical mechanics. (Precisely the same arguments apply when considering the number of atoms in two halves of a box.)

2.2.1 Quantum states of a spin 1/2 paramagnet

A spin $S = 1/2$ has two possible orientations. (These are the two possible values of the projection of the spin on the z axis: $m_s = \pm 1/2$.) Associated with each spin is a magnetic moment which has the two possible values $\pm\mu$. An example of such a system is the nucleus of ^3He which has $S = 1/2$ (this is due to an unpaired neutron; the other isotope of helium ^4He is non-magnetic). Another example is the electron.

In a magnetic field the two states are of different energy, because magnetic moments prefer to line up parallel (rather than antiparallel) with the field.

Thus since $E = -\boldsymbol{\mu} \cdot \mathbf{B}$, the energy of the two states is $E_{\uparrow} = -\mu B$, $E_{\downarrow} = \mu B$. (Here the arrows refer to the direction of the nuclear magnetic moment: $\uparrow$ means that the moment is parallel to the field and $\downarrow$ means antiparallel. There is potential confusion because for ^3He the moment points opposite to the spin!).

2.2.2 The notion of a microstate

So much for a single particle. But we are interested in a system consisting of a large number of such particles, N . A microscopic description would necessitate specifying the state of *each* particle.

In a localised assembly of such particles, each particle has only two possible quantum states $\uparrow$ or $\downarrow$. (By localised we mean the particles are fixed in position, like in a solid. We can label them if we like by giving a number to each position. In this way we can tell them apart; they are *distinguishable*). In a magnetic field the energy of each particle has only two possible values. (This is an example of a two level system). You can't get much simpler than that!

Now we have set up the system, let's explain what we mean by a microstate and enumerate them. First consider the system in zero field. Then the states $\uparrow$ and $\downarrow$ have the same energy.

We specify a microstate by giving the quantum state of each particle, whether it is $\uparrow$ or $\downarrow$.

For $N = 10$ spins

$\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow$	
$\downarrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow$	
$\uparrow\downarrow\downarrow\downarrow\downarrow\downarrow\uparrow\uparrow\uparrow\uparrow$	are possible microstates.

That's it!

2.2.3 Counting the microstates

What is the *total number* of such microstates (accessible to the system). This is called Ω .

Well for each particle the spin can be $\uparrow$ or $\downarrow$ (there is no restriction on this in zero field). Two possibilities for each particle gives 2^{10} arrangements (we have merely given three examples) for $N = 10$. For N particles $\Omega = 2^N$.

2.2.4 Distribution of particles among states

This list of possible microstates is far too detailed to handle. What's more we don't *need* all this detail to calculate the properties of the system. For example the total magnetic moment of all the particles is $M = (N_{\uparrow} - N_{\downarrow})\mu$; it just depends on the relative number of up and down moments and not on the detail of *which* are up or down.

So we collect together all those microstates with the same number of up moments and down moments. Since the relative number of ups and downs is constrained by $N_{\uparrow} + N_{\downarrow} = N$ (the moments *must* be *either* up or down), such a state can be characterised by one number:

$$m = N_{\uparrow} - N_{\downarrow}.$$

This number m tells us the *distribution* of N particles among possible states:

$$N_{\downarrow} = (N - m)/2, \quad N_{\uparrow} = (N + m)/2.$$

Now we ask the question: "how many microstates are there consistent with a given distribution (given value of m ; given values of $N_{\uparrow}, N_{\downarrow}$). Call this $t(m)$. (This is Guenault's notation).

Look at $N = 10$.

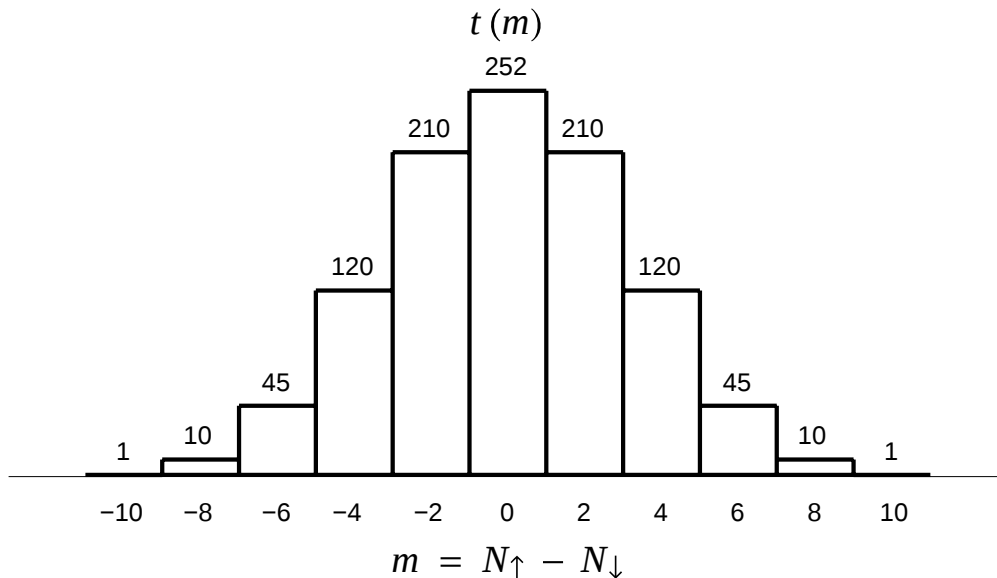
For $m = 10$ (all spins up) and $m = -10$ (all spins down), that's easy! — $t = 1$.

Now for $m = 8$ (one moment down) $t(m = 8) = 10$. There are ten ways of reversing one moment.

The general result is $t(m) = N! / N_{\uparrow}! N_{\downarrow}!$. This may be familiar to you as a binomial coefficient. It is the number of ways of dividing N objects into one set of $N_{\uparrow}$ identical objects and $N_{\downarrow}$ different identical objects (red ball and blue balls or tossing an unbiased coin and getting heads or tails). In terms of m the function $t(m)$ may be written

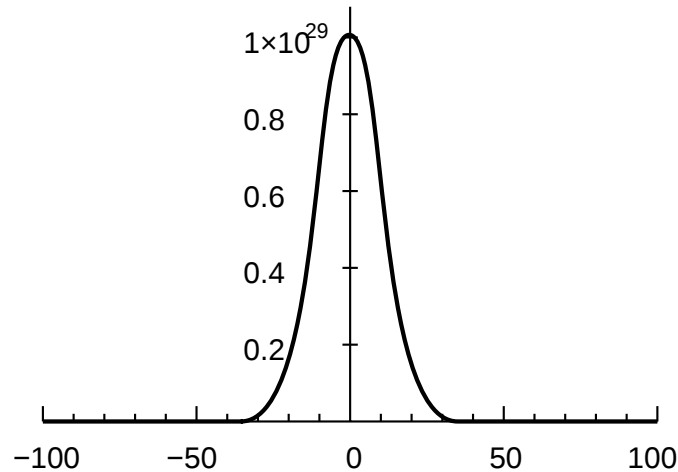
$$t(m) = \frac{N!}{\left(\frac{N-m}{2}\right)! \left(\frac{N+m}{2}\right)!}.$$

If you plot the function $t(m)$ it is peaked at $m = 0$, ($N_{\uparrow} = N_{\downarrow} = N/2$). Here we illustrate that for $N = 10$. The sharpness of the distribution increases with the size of the system N , since the standard deviation goes as $\sqrt{N}$.



Plot of the function $t(m)$ for 10 spins

When $N = 100$ the function is already much narrower:



Plot of the function $t(m)$ for 100 spins

When N is large, then $t(m)$ approaches the normal (Gaussian) function

$$t(m) \approx 2^N \exp - \frac{m^2}{2N}.$$

This may be shown using Stirling's approximation (Guenault, Appendix 2). Note that the RMS width of the function is $\sqrt{N}$.

2.2.5 The average distribution and the most probable distribution

The physical significance of this result derives from the fundamental assumption of statistical physics that each of these microstates is equally likely. It follows that $t(m)$ is the statistical weight of the distribution m (recall m determines $N_{\uparrow}$ and $N_{\downarrow}$), that is the relative probability of that distribution occurring.

Hence we can work out the *average distribution*; in this example this is just the average value of m . The probability of a particular value of m is just $t(m)/\Omega$ i.e. the number of microstates with that value of m divided by the total number of microstates ($\Omega = \sum_m t(m)$). So the average value of m is $\sum_m m t(m)/\Omega$.

In this example because the distribution function $t(m)$ is symmetrical it is clear that $m_{\text{av}} = 0$. Also the value of m for which $t(m)$ is maximum is $m = 0$.

So on average for a system of spins in zero field $m = 0$: there are equal numbers of up and down spins. This average value is also the most probable value. For large N the function $t(m)$ is very strongly peaked at $m = 0$; there are far more microstates with $m = 0$ than any other value. If we observe a system of spins as a function of time for most of the time we will find m to be at or near $m = 0$. The observed m will fluctuate about $m = 0$, the (relative) extent of these fluctuations decreases with N . States far away from $m = 0$ such as $m = N$ (all spins up) are highly unlikely; the probability of observing that state is $1/\Omega = 1/2^N$ since there is only one such state out of a total 2^N .

(Note: According to the Gibbs method of ensembles we represent such a system by an ensemble of Ω systems, each in a definite microstate and one for every microstate. The thermodynamic properties are obtained by an average over the ensemble. The equivalence of the ensemble average to the time average for the system is a subtle point and is the subject of the ergodic hypothesis.)

It is worthwhile to note that the model treated here is applicable to a range of different phenomena. For example one can consider the number of particles in two halves of a container, and examine the density fluctuations. Each particle can be either side of the imaginary division so that the distribution of density in either half would follow the same distribution as derived above.

End of lecture 4

2.3 Entropy and the second law of thermodynamics

2.3.1 Order and entropy

Suppose now that somehow or other we could set up the spins, in zero magnetic field, such that they were all in an up state $m = N$. This is not a state of internal thermodynamic equilibrium; it is a highly ordered state. We know the equilibrium state is random with equal numbers of up and down spins. It also has a large number of microstates associated with it, so it is far the most probable state. The initial state of the system will therefore spontaneously evolve towards the more likely, more disordered $m = 0$ equilibrium state. Once there the fluctuations about $m = 0$ will be small, as we have seen, and the probability of the system returning fleetingly to the initial ordered state is infinitesimally small. The tendency of isolated systems to evolve in the direction of increasing disorder thus follows from (i) disordered distributions having a larger number of microstates (mathematical fact) (ii) all microstates are equally likely (physical hypothesis).

The thermodynamic quantity intimately linked to this disorder is the entropy. Entropy is a function of state and is defined for a system in thermodynamic equilibrium. (It is possible to introduce a

generalised entropy for systems not in equilibrium but let's not complicate the issue). Entropy was defined by Boltzmann and Planck as

$$S = k \ln \Omega$$

where Ω is the total number of microstates accessible to a system. Thus for our example of spins $\Omega = 2^N$ so that $S = Nk \ln 2$. Here k is Boltzmann's constant.

2.3.2 The second law of thermodynamics

The second law of thermodynamics is the law of increasing entropy. *During any real process the entropy of an isolated system always increases. In the state of equilibrium the entropy attains its maximum value.*

This law may be illustrated by asking what happens when a constraint is removed on an isolated composite system. Is the number of microstates of the final equilibrium state be smaller, the same or bigger? We expect the system to evolve towards a more probable state. Hence the number of accessible microstates of the final state must be *greater or the same* and the entropy increases or stays the same. A nice simple example is given by Guenault on p12-13. The mixing of two different gases is another.

For a composite system $\Omega = \Omega_1 \Omega_2$. If the two systems are allowed to exchange energy, then in the final equilibrium state the total number of accessible microstates Ω^* is always greater.

The macroscopic description of this is that heat flows until the temperature of the two systems is the same. This leads us, in the next section, to a definition of statistical temperature as

$$\frac{1}{T} = \left. \frac{\partial S}{\partial E} \right|_{V,N}.$$

2.3.3 Thermal interaction between systems and temperature

Consider two isolated systems of given volume, number of particles, total energy

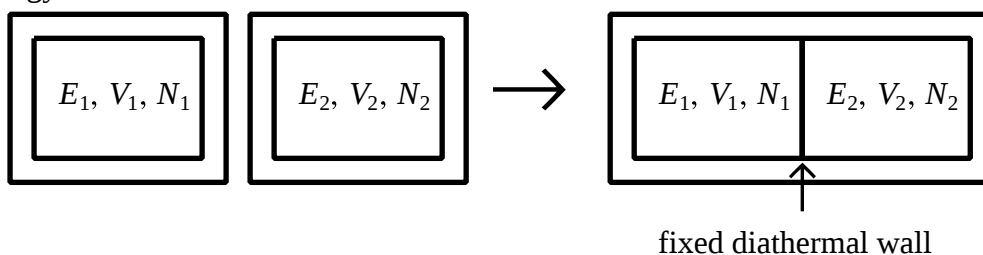
When separated and in equilibrium they will individually have

$$\Omega_1 = \Omega_1(E_1, V_1, N_1) \text{ and } \Omega_2 = \Omega_2(E_2, V_2, N_2)$$

microstates.

The total number of microstates for the combined system is $\Omega = \Omega_1 \Omega_2$.

Now suppose the two systems are brought into contact through a diathermal wall, so they can now exchange energy.



Thermal interaction

The composite system is isolated, so its total energy is constant. So while the systems exchange energy (E_1 and E_2 can vary) we must keep $E_1 + E_2 = E_0 = \text{const.}$ And since V_1, N_1, V_2, N_2 all

remain fixed, they can be ignored (and kept constant in any differentiation).

Our problem is this: after the two systems are brought together what will be the equilibrium state? We know that they will exchange energy, and they will do this so as to maximise the total number of microstates for the composite system.

The systems will exchange energy so as to maximise Ω . Writing

$$\Omega = \Omega_1(E) \Omega_2(E_0 - E)$$

we allow the systems to vary E so that Ω is a maximum:

$$\frac{\partial \Omega}{\partial E} = \frac{\partial \Omega_1}{\partial E} \Omega_2 - \Omega_1 \frac{\partial \Omega_2}{\partial E} = 0$$

or

$$\frac{1}{\Omega_1} \frac{\partial \Omega_1}{\partial E} = \frac{1}{\Omega_2} \frac{\partial \Omega_2}{\partial E}$$

or

$$\frac{\partial \ln \Omega_1}{\partial E} = \frac{\partial \ln \Omega_2}{\partial E}.$$

But from the definition of entropy, $S = k \ln \Omega$, we see this means that the equilibrium state is characterised by

$$\frac{\partial S_1}{\partial E} = \frac{\partial S_2}{\partial E}.$$

In other words, when the systems have reached equilibrium the quantity $\partial S / \partial E$ of system 1 is equal to $\partial S / \partial E$ of system 2. Since we have defined temperature to be that quantity which is the same for two systems in thermal equilibrium, then it is clear that $\partial S / \partial E$ (or $\partial \ln \Omega / \partial E$) must be related (somehow) to the temperature.

We define statistical temperature as

$$\frac{1}{T} = \left. \frac{\partial S}{\partial E} \right|_{V,N}$$

(recall that V and N are kept constant in the process) With this definition it turns out that statistical temperature is identical to absolute temperature.

It follows from this definition and the law of increasing entropy that *heat must flow from high temperature to low temperature* (this is actually the Clausius statement of the second law of thermodynamics. We will discuss all the various formulations of this law a bit later on.)

Let us prove this. For a composite system $\Omega = \Omega_1 \Omega_2$ so that $S = S_1 + S_2$ as we have seen. According to the law of increasing entropy $\Delta S \geq 0$, so that:

$$\Delta S = \left(\frac{\partial S_1}{\partial E} - \frac{\partial S_2}{\partial E} \right) \Delta E_1 \geq 0$$

or, using our definition of statistical temperature:

$$\Delta S = \left(\frac{1}{T_1} - \frac{1}{T_2} \right) \Delta E_1 \geq 0.$$

this means that

$$E_1 \text{ increases if } T_2 > T_1$$

$$E_1 \text{ decreases if } T_2 < T_1$$

so energy flows from systems at higher temperatures to systems at lower temperatures.

End of lecture 5

2.4 More on the $S=1/2$ paramagnet

2.4.1 Energy distribution of the $S=1/2$ paramagnet

Having motivated this definition of temperature we can now return to our simple model system and solve it to find $N_\uparrow$ and $N_\downarrow$ as a function of magnetic field B and temperature T . Previously we had considered the behaviour in zero magnetic field only and we had not discussed temperature.

As discussed already in a magnetic field the energies of the two possible spin states are

$$E_\uparrow = -\mu B \text{ and } E_\downarrow = \mu B.$$

We shall write these energies as $\mp\epsilon$.

Now for an isolated system, with total energy E , number N , and volume V fixed, it turns out that $N_\uparrow$ and $N_\downarrow$ are *uniquely determined* (since the energy depends on $N_\uparrow - N_\downarrow$). Thus $m = N_\uparrow - N_\downarrow$ is *fixed* and therefore it exhibits no fluctuations (as it did in zero field about the average value $m = 0$).

This follows since we must have both

$$E = (N_\downarrow - N_\uparrow) \epsilon$$

$$N = (N_\downarrow + N_\uparrow)$$

These two equations are solved to give

$$N_\downarrow = (N + E/\epsilon)/2$$

$$N_\uparrow = (N - E/\epsilon)/2.$$

All microstates have the same distribution of particles between the two energy levels $N_\uparrow$ and $N_\downarrow$.

But there are still things we would like to know about this system; in particular we know nothing about the temperature. We can approach this from the entropy in our now-familiar way.

The total number of microstates Ω is given by

$$\Omega = \frac{N!}{N_\uparrow! N_\downarrow!}.$$

So we can calculate the entropy via $S = k \ln \Omega$:

$$S = k \{ \ln N! - \ln N_\uparrow! - \ln N_\downarrow! \}$$

Now we use Stirling's approximation (Guenault Appendix 2) which says

$$\ln x! \approx x \ln x - x;$$

this is important you should remember this.

Hence

$$S = k \{ N \ln N - N_{\uparrow} \ln N_{\uparrow} - N_{\downarrow} \ln N_{\downarrow} \},$$

where we have used $N_{\downarrow} + N_{\uparrow} = N$. Into this we substitute for $N_{\uparrow}$ and $N_{\downarrow}$ since we need S in terms of E for differentiation to get temperature.

$$S = k \left[N \ln N - \frac{1}{2} (N + E/\varepsilon) \ln \left\{ \frac{1}{2} (N + E/\varepsilon) \right\} - \frac{1}{2} (N - E/\varepsilon) \ln \left\{ \frac{1}{2} (N - E/\varepsilon) \right\} \right].$$

Now we use the definition of statistical temperature: $1/T = \partial S / \partial E|_{N,V}$ to obtain the temperature:

$$\frac{1}{T} = \frac{k}{2\varepsilon} \ln \left\{ \frac{N - E/\varepsilon}{N + E/\varepsilon} \right\}.$$

(You can check this differentiation with *Maple* or *Mathematica* if you wish!)

Recalling the expressions for $N_{\uparrow}$ and $N_{\downarrow}$, the temperature expression can be written:

$$\frac{1}{T} = \frac{k}{2\varepsilon} \ln \left\{ \frac{N_{\uparrow}}{N_{\downarrow}} \right\},$$

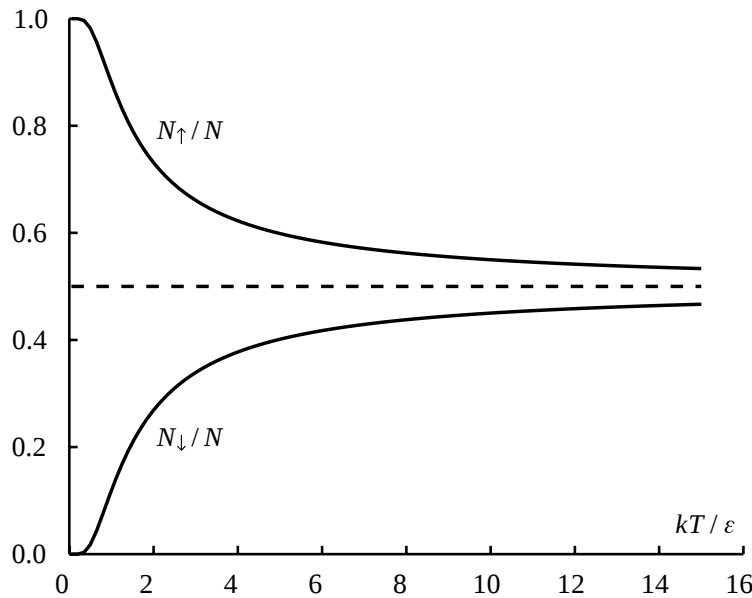
which can be inverted to give the ratio of the populations as

$$\frac{N_{\uparrow}}{N_{\downarrow}} = \exp -2\varepsilon / kT.$$

Or, since $N_{\downarrow} + N_{\uparrow} = N$ we find

$$\frac{N_{\uparrow}}{N} = \frac{\exp -\varepsilon / kT}{\exp +\varepsilon / kT + \exp -\varepsilon / kT}$$

$$\frac{N_{\downarrow}}{N} = \frac{\exp +\varepsilon / kT}{\exp +\varepsilon / kT + \exp -\varepsilon / kT}$$



Temperature variation of up and down populations

This is our final result. It tells us the fraction of particles in each of the two energy states as a function of temperature. This is the distribution function. On inspection you see that it can be written rather concisely as

$$n(\varepsilon) = N \frac{\exp -\varepsilon / kT}{z}$$

where the quantity z

$$z = \exp -\varepsilon / kT + \exp +\varepsilon / kT$$

is called the (single particle) *partition function*. It is the sum over the possible states of the factor $\exp -\varepsilon / kT$. This distribution among energy levels is an example of the Boltzmann distribution.

This distribution applies generally (to distinguishable particles) where there is a whole set of possible energies available and not merely two as we have considered thus far. This problem we treat in the next section. In general

$$z = \sum_{\text{states}} \exp -\varepsilon / kT,$$

where the sum is taken over all the possible energies of a single particle. In the present example there are only two energy levels.

In the general case $n(\varepsilon)$ is the *average* number of particles in the state of energy ε (In the present special case of only two levels $n(\varepsilon_1)$ and $n(\varepsilon_2)$ are uniquely determined). To remind you, an example of counting the microstates and evaluating the average distribution for a system of a few particles with a large number of available energies is given in Guenault chapter 1. The fluctuations in $n(\varepsilon)$ about this average value are small for a sufficiently large number of particles. (The $1/\sqrt{N}$ factor)

2.4.2 Magnetisation of $S = 1/2$ magnet — Curie's law

We can now obtain an expression for the magnetisation of the spin $1/2$ paramagnet in terms of temperature and applied magnetic field. The magnetisation (total magnetic moment) is given by:

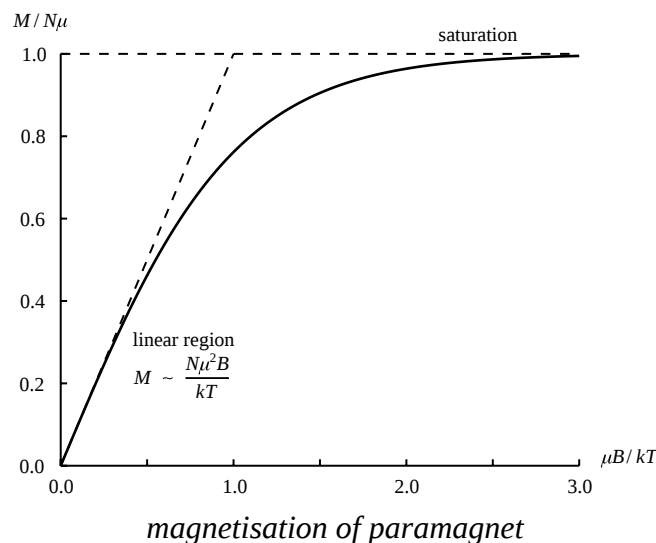
$$M = (N_{\uparrow} - N_{\downarrow})\mu.$$

We have expressions for $N_{\uparrow}$ and $N_{\downarrow}$, giving the expression for M as

$$\begin{aligned} M &= N\mu \frac{\exp -\varepsilon / kT - \exp +\varepsilon / kT}{\exp +\varepsilon / kT + \exp -\varepsilon / kT} \\ &= -N\mu \tanh \varepsilon / kT \end{aligned}$$

or, since $\varepsilon = -\mu B$, the magnetisation in terms of the magnetic field is

$$M = N\mu \tanh \left(\frac{\mu B}{kT} \right).$$



The general behaviour of magnetisation on magnetic field is nonlinear. At low fields the magnetisation starts linearly but at higher fields it saturates as all the moments become aligned. The low field, linear behaviour may be found by expanding the tanh:

$$M = \frac{N\mu^2}{kT} B.$$

The magnetisation has the general form

$$M = C \frac{B}{T},$$

proportional to B and inversely proportional to T . This behaviour is referred to as Curie's law and the constant $C = N\mu^2/k$ is called the Curie constant.

2.4.3 Entropy of $S = 1/2$ magnet

We can see immediately that for the spin system the entropy is a function of the magnetic field. In particular at large B/T all the moments are parallel to the field. There is then only one possible microstate and the entropy is zero. But it is quite easy to determine the entropy at all B/T from $S = k \ln \Omega$.

As we have seen already

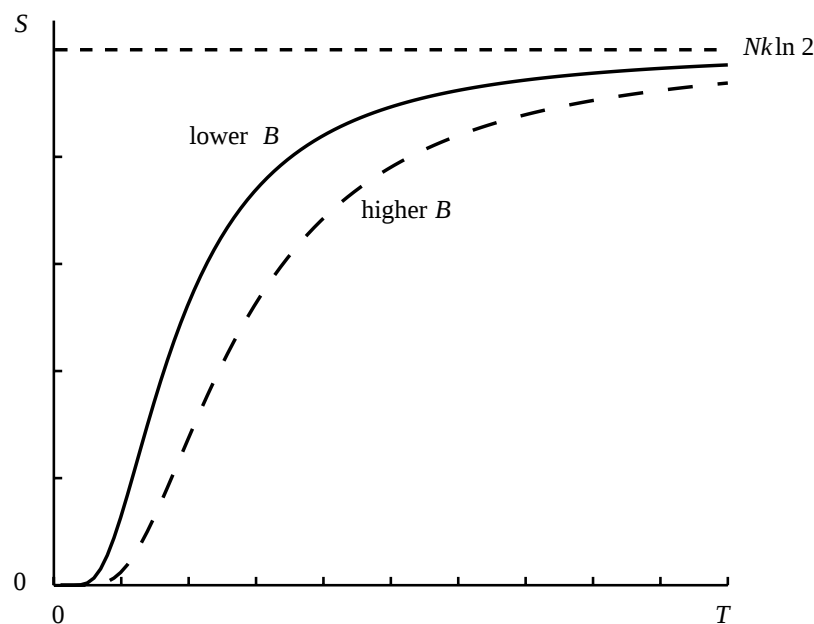
$$S/k = N \ln N - N_{\uparrow} \ln N_{\uparrow} - N_{\downarrow} \ln N_{\downarrow}.$$

Since $N = N_{\downarrow} + N_{\uparrow}$, this can be written

$$S/k = -N_{\uparrow} \ln N_{\uparrow}/N - N_{\downarrow} \ln N_{\downarrow}/N.$$

And substituting for $N_{\uparrow}$ and $N_{\downarrow}$ we then get

$$S = \frac{Nk}{\exp(-2\mu B/kT) + 1} \ln [\exp(-2\mu B/kT) + 1] + \frac{Nk}{\exp(2\mu B/kT) + 1} \ln [\exp(2\mu B/kT) + 1]$$



Entropy of a spin 1/2 paramagnet

- 1 At low temperatures $kT \ll 2\mu B$. Then the first term is negligible and the 1s may be ignored. In this case $S \approx Nk (2\mu B / kT) \exp(-2\mu B / kT)$. Clearly $S \rightarrow 0$ as $T \rightarrow 0$, in agreement with our earlier argument.
- 2 At high temperatures $kT \gg 2\mu B$. Then both terms are equal and we find $S = Nk \ln 2$. There are equal numbers of up and down spins: maximum disorder.
- 3 The entropy is a function of B/T and increasing B at constant T reduces S ; the spins tend to line up; the system becomes more disordered.

End of lecture 6

2.5 The Boltzmann distribution

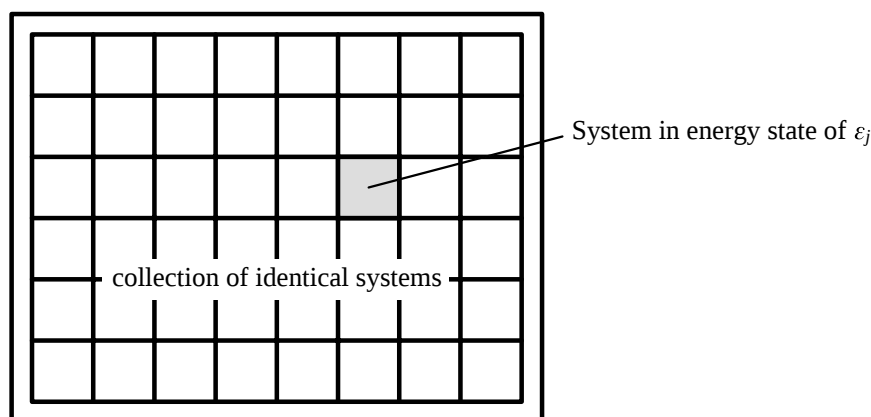
2.5.1 Thermal interaction with the rest of the world – using a Gibbs ensemble

For an isolated system all microstates are equally likely; this is our Fundamental Postulate. But what about a non-isolated system? What can we say about the probabilities of microstates of such a system? Here the probability of a microstate will depend on its energy and on properties of the surroundings.

A real system will have its temperature determined by its environment. That is, it is not isolated, its energy is not fixed; it will fluctuate (but the relative fluctuations are very small because of the $1/\sqrt{N}$ rule).

All we really know about is isolated systems; here all quantum states are equally probable. So to examine the properties of a non-isolated system we shall adopt a trick. We will take many copies of our system. These will be allowed to exchange (heat) energy with each other, but the entire collection will be isolated from the outside world. This collection of systems is called an *ensemble*.

The point now is that we now know that all microstates of the composite ensemble are equally likely. Of the N individual elements of the ensemble, there will be n_i in the microstate of energy ϵ_i , so the probability of an element being in this microstate will be n_i / N .



a Gibbs ensemble of systems used to calculate $\{n_i\}$

There are many distributions $\{n_i\}$ which are possible for this ensemble. In particular we know that because the whole ensemble is isolated then the number of elements and the total energy are fixed. In other words any distribution $\{n_i\}$ must satisfy the requirements

$$\sum_i n_i = N$$

$$\sum_i n_i \varepsilon_i = E.$$

By analogy with the case of the spin 1/2 paramagnet, we will denote the number of microstates of the ensemble corresponding to a given distribution $\{n_i\}$ by $t(\{n_i\})$. Then the most likely distribution is that for which $t(\{n_i\})$ is maximised.

2.5.2 Most likely distribution

The value of $t(\{n_i\})$ is given by

$$t(\{n_i\}) = \frac{N!}{\prod_i n_i!}$$

since there are N elements all together and there are n_i in the i^{th} state. So the most probable distribution is found by varying the various n_i to give the maximum value for t . It is actually more convenient (and mathematically equivalent) to maximise the logarithm of t .

The maximum in $\ln t$ corresponds to the place where its differential is zero:

$$d \ln t(\{n_i\}) = \frac{\partial \ln t}{\partial n_1} dn_1 + \frac{\partial \ln t}{\partial n_2} dn_2 + \dots + \frac{\partial \ln t}{\partial n_i} dn_i + \dots = 0$$

If the n_i could be varied independently then we would have the following set of equations

$$\frac{\partial \ln t(\{n_i\})}{\partial n_i} = 0, \quad i = 1, 2, 3, \dots$$

Unfortunately the n_i can not all be varied independently because all possible distributions $\{n_i\}$ must satisfy the two requirements

$$\sum_i n_i = N$$

$$\sum_i n_i \varepsilon_i = E;$$

there are two constraints on the distribution $\{n_i\}$. These may be incorporated in the following way. Since the constraints imply that

$$d \sum_i n_i = 0$$

and

$$d \sum_i n_i \varepsilon_i = 0,$$

we can consider the differential of the expression

$$\ln t(\{n_i\}) - \alpha \sum_i n_i - \beta \sum_i n_i \varepsilon_i$$

where α and β are, at this stage undetermined. This gives us two extra degrees of freedom so that we can maximise this by varying all n_i independently. In other words, the maximum in $\ln t$ is also

specified by

$$d \left\{ \ln t(\{n_i\}) - \alpha \sum_i n_i - \beta \sum_i n_i \varepsilon_i \right\} = 0$$

Now we have recovered the two lost degrees of freedom so that the n_i can be varied independently. But then the multipliers α and β are no longer free variables and their values must be found from the constraints fixing N and E . (This is called the method of Lagrange's undetermined multipliers.)

The maximisation can now be performed by setting the N partial derivatives to zero

$$\frac{\partial}{\partial n_i} \left[\ln t(\{n_i\}) - \alpha \sum_i n_i - \beta \sum_i n_i \varepsilon_i \right] = 0, \quad i = 1, 2, 3, \dots$$

We use Stirling's approximation for $\ln t$, given by

$$\ln t(\{n_i\}) = \left(\sum_i n_i \right) \ln \left(\sum_i n_i \right) - \sum_i n_i \ln n_i.$$

so that upon evaluating the N partial derivatives we obtain

$$\ln \sum_i n_i - \ln n_j - \alpha - \beta \varepsilon_j = 0$$

which has solution

$$n_j = N e^{-\alpha} e^{-\beta \varepsilon_j}.$$

These are the values of n_j which maximise t subject to N and E being fixed. This gives us the most probable distribution $\{n_i\}$.

The probability that a system will be in the state j is then found by dividing by N :

$$p_j = e^{-\alpha} e^{-\beta \varepsilon_j}.$$

So the remaining step, then, is to find the constants α and β .

2.5.3 What are α and β ?

For the present we shall sidestep the question of α by appealing to the normalisation of the probability distribution. Since we must have

$$\sum_j p_j = 1$$

it follows that we can express the probabilities as

$$p_j = \frac{e^{-\beta \varepsilon_j}}{Z}$$

where the normalisation constant Z is given by

$$Z = \sum_j e^{-\beta \varepsilon_j}.$$

The constant Z will turn out to be of central importance; from this all thermodynamic properties can be found. It is called the partition function, and it is given the symbol Z because of its name in German: *zustandsumme* (sum over states).

But what is this β ? For the spin 1/2 paramagnet we had many similar expressions with $1/kT$ there, where we have β here. We shall now show that this identification is correct. We will approach this from our definition of temperature:

$$\frac{1}{T} = \left. \frac{\partial S}{\partial E} \right|_{V,N}.$$

Now the fundamental expression for entropy is $S = k \ln \Omega$, where Ω is the total number of microstates of the ensemble. This is a little difficult to obtain. However we know $t(\{n_i\})$, the number of microstates corresponding to a given distribution $\{n_i\}$. And this t is a very sharply peaked function. Therefore we make negligible error in approximating Ω by the value of t corresponding to the most probable distribution. In other words, we take

$$S = k \ln \left(\frac{N!}{\prod_i n_i!} \right)$$

where

$$n_j = N e^{-\alpha} e^{-\beta \epsilon_j}.$$

Using Stirling's approximation we have

$$\begin{aligned} S &= k \left\{ N \ln N - \sum_i n_i \ln n_i \right\} \\ &= k \{ N \ln N + \alpha N + \beta E \}. \end{aligned}$$

From this we immediately identify

$$\frac{1}{T} = \left. \frac{\partial S}{\partial E} \right|_{V,N} = k\beta$$

so that

$$\beta = \frac{1}{kT}$$

as we asserted.

The probability of a (non-isolated) system being found in the i^{th} microstate is then

$$p_i \propto e^{-\epsilon_i/kT}$$

or

$$p_i = \frac{e^{-\epsilon_i/kT}}{Z}$$

This is known as the Boltzmann distribution or the Boltzmann factor. It is a key result. Feynman says "This fundamental law is the summit of statistical mechanics, and the entire subject is either a slide-down from the summit, as the principle is applied to various cases, or the climb-up to where the fundamental law is derived and the concepts of thermal equilibrium and temperature clarified".

End of lecture 7

2.5.4 Link between the partition function and thermodynamic variables

All the important thermodynamic variables for a system can be derived from Z and its derivatives. We can see this from the expression for entropy. We have

$$S = k \{N \ln N + \alpha N + \beta E\}.$$

And since $e^{-\alpha} = N/Z$ we have, on taking logarithms

$$\alpha = -\ln N + \ln Z$$

so that

$$\begin{aligned} S &= k \{N \ln N - N \ln N + N \ln Z + E/kT\} \\ &= Nk \ln Z + E/T. \end{aligned}$$

Here both S and E refer to the ensemble of N elements. So for the entropy and internal energy of a single system

$$S = k \ln Z + E/T,$$

which, upon rearrangement, can be written

$$E - TS = -k \ln Z.$$

Now the thermodynamic function

$$F = E - TS$$

is known as the Helmholtz free energy, or the Helmholtz potential. We then have the memorable result

$$F = -kT \ln Z.$$

2.5.5 Finding Thermodynamic Variables

A host of thermodynamic variables can be obtained from the partition function. This is seen from the differential of the free energy. Since

$$dE = TdS - pdV$$

it follows that

$$dF = -SdT - pdV.$$

We can then identify the various partial derivatives:

$$\begin{aligned} S &= - \left. \frac{\partial F}{\partial T} \right|_V = kT \left. \frac{\partial \ln Z}{\partial T} \right|_V + k \ln Z \\ p &= - \left. \frac{\partial F}{\partial V} \right|_T = kT \left. \frac{\partial \ln Z}{\partial V} \right|_T. \end{aligned}$$

Since $E = F + TS$ we can then express the internal energy as

$$E = kT^2 \left. \frac{\partial \ln Z}{\partial T} \right|_V.$$

Thus we see that once the partition function is evaluated by summing over the states, all relevant thermodynamic variables can be obtained by differentiating Z .

It is instructive to examine this expression for the internal energy further. This will also confirm the identification of this function of state as the actual energy content of the system. If p_j is the probability of the system being in the eigenstate corresponding to energy ε_j then the mean energy of the system may be expressed as

$$\begin{aligned}
E &= \sum_j \varepsilon_j p_j \\
&= \frac{1}{Z} \sum_j \varepsilon_j e^{-\beta \varepsilon_j}
\end{aligned}$$

where Z is the previously-defined partition function, and it is convenient here to work in terms of β rather than converting to $1/kT$.

In examining the sum we note that

$$\varepsilon_j e^{-\beta \varepsilon_j} = -\frac{\partial}{\partial \beta} e^{-\beta \varepsilon_j}$$

so that the expression for E may be written, after interchanging the differentiation and the summation,

$$E = -\frac{1}{Z} \frac{\partial}{\partial \beta} \sum_j e^{-\beta \varepsilon_j}.$$

But the sum here is just the partition function, so that

$$\begin{aligned}
E &= -\frac{1}{Z} \frac{\partial Z}{\partial \beta} \\
&= -\frac{\partial}{\partial \beta} \ln Z.
\end{aligned}$$

And since $\beta = 1/kT$, this is equivalent to our previous expression

$$E = kT^2 \left. \frac{\partial \ln Z}{\partial T} \right|_v,$$

however the mathematical manipulations are often more convenient in terms of β .

2.5.6 Summary of methodology

We have seen that the partition function provides a means of calculating all thermodynamic information from the microscopic description of a system. We summarise this procedure as follows:

- 1 Write down the possible energies for the system.
- 2 Evaluate the partition function for the system.
- 3 The Helmholtz free energy then follows from this.
- 4 All thermodynamic variables follow from differentiation of the Helmholtz free energy.

2.6 Localised systems

2.6.1 Working in terms of the partition function of a single particle

The partition function methodology described in the previous sections is general and very powerful. However a considerable simplification follows if one is considering systems made up of noninteracting distinguishable or localised particles. The point about *localised* particles is that even though they may be indistinguishable, perhaps a solid made up of argon atoms, the particles can be distinguished by their positions.

Since the partition function for distinguishable systems is the product of the partition function for each system, if we have an assembly of N localised identical particles, and if the partition function for a single such particle is z , then the partition function for the assembly is

$$Z = z^N.$$

It then follows that the Helmholtz free energy for the assembly is

$$\begin{aligned} F &= -kT \ln Z \\ &= -NkT \ln z \end{aligned}$$

in other words the free energy is N times the free energy contribution of a single particle; the free energy is an extensive quantity, as expected.

This allows an important simplification to the general methodology outlined above. For localised systems we need only consider the energy levels of a *single particle*. We then evaluate the partition function z of a single particle and then use the relation $F = -NkT \ln z$ in order to find the thermodynamic properties of the system.

We will consider two examples of this, one familiar and one new.

End of lecture 8

2.6.2 Using the partition function I — the $S = 1/2$ paramagnet (again)

Step 1) *Write down the possible energies for the system*

The assembly of magnetic moments μ are placed in a magnetic field B . The spin $1/2$ has two quantum states, which we label by $\uparrow$ and $\downarrow$. The two energy levels are then

$$\varepsilon_{\uparrow} = -\mu B, \quad \varepsilon_{\downarrow} = \mu B.$$

Step 2) *Evaluate the partition function for the system*

We require to find the partition function for a single spin. This is

$$\begin{aligned} z &= \sum_{\text{states } i} e^{-\varepsilon_i/kT} \\ &= e^{\mu B/kT} + e^{-\mu B/kT}. \end{aligned}$$

This time we shall obtain the results in terms of hyperbolic functions rather than exponentials — for variety. The partition function is expressed as the hyperbolic cosine:

$$z = 2 \cosh \mu B / kT.$$

Step 3) *The Helmholtz free energy then follows from this*

Here we use the relation

$$\begin{aligned} F &= -NkT \ln z \\ &= -NkT \ln \{2 \cosh \mu B / kT\}. \end{aligned}$$

Step 4) *All thermodynamic variables follow from differentiation of the Helmholtz free energy*

Before proceeding with this step we must pause to consider the performance of magnetic work. Here we don't have pressure and volume as our work variables; we have magnetisation M and magnetic field B . The expression for magnetic work is

$$dW = -MdB,$$

so comparing this with our familiar $-pdV$, we see that when dealing with magnetic systems we must make the identification

$$p \leftrightarrow M$$

$$V \leftrightarrow B.$$

The internal energy differential of a magnetic system is then

$$dE = TdS - MdB$$

and, of more immediate importance, the Helmholtz free energy $E - TS$ has the differential

$$dF = -SdT - MdB.$$

We can then identify the various partial derivatives:

$$S = - \left. \frac{\partial F}{\partial T} \right|_B = NkT \left. \frac{\partial \ln z}{\partial T} \right|_B + Nk \ln z,$$

$$M = - \left. \frac{\partial F}{\partial B} \right|_T = NkT \left. \frac{\partial \ln z}{\partial B} \right|_T.$$

Upon differentiation we then obtain

$$S = -\frac{N\mu B}{T} \tanh \mu B / kT + Nk \ln \{2 \cosh \mu B / kT\},$$

$$M = N\mu \tanh \mu B / kT.$$

The internal energy is best obtained now from

$$E = F + TS$$

$$= -NkT \ln z + TS$$

$$= -NkT \ln \{2 \cosh \mu B / kT\} + T \left[-\frac{N\mu B}{T} \tanh \mu B / kT + Nk \ln \{2 \cosh \mu B / kT\} \right],$$

giving

$$E = -N\mu B \tanh \mu B / kT.$$

We note that the internal energy can be written as $E = -MB$ as expected.

By differentiating the internal energy with respect to temperature we obtain the thermal capacity at constant field

$$C_B = Nk \left(\frac{\mu B}{kT} \right)^2 \operatorname{sech}^2 \mu B / kT.$$

Some of these expressions were derived before, from the microcanonical approach; you should check that the exponential expressions there are equivalent to the hyperbolic expressions here.

We plot the magnetisation as a function of *inverse* temperature. Recall that at high temperatures we have a linear region, where Curie's law holds. At low temperatures the magnetisation saturates as all the moments become aligned against the applied field.

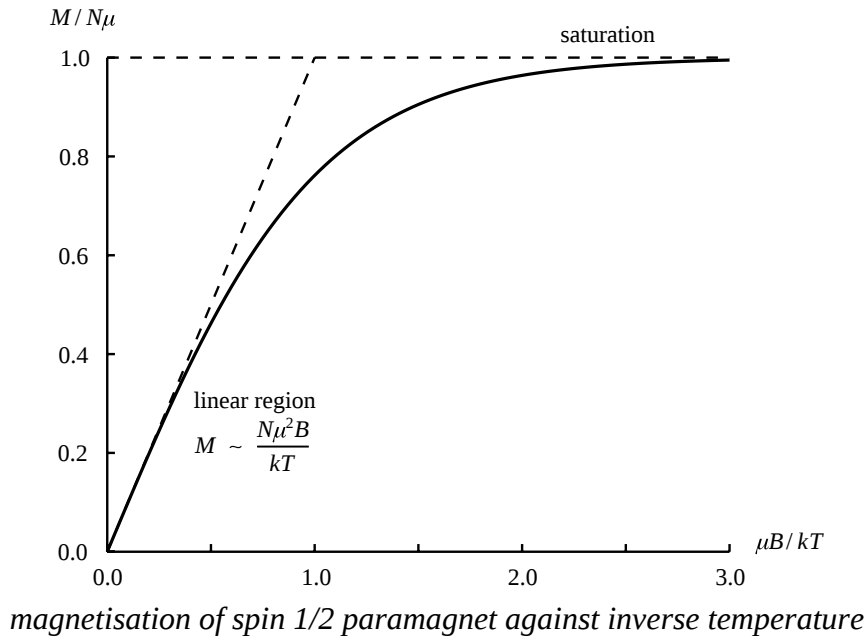
Incidentally, we note that the expression

$$M = N\mu \tanh \mu B / kT$$

is the equation of state for the paramagnet. This is a nonlinear equation. But recall that we found linear behaviour in the high T/B region where

$$M = \frac{N\mu^2 B}{kT},$$

just as it is in the high temperature region that the ideal gas equation of state $p = NkT/V$ is valid.



Next we plot the entropy, internal energy and the thermal capacity as a function of temperature. Observe that they all go to zero as $T \rightarrow 0$.

At low temperatures the entropy goes to zero, as expected. And at high temperatures the entropy goes to the classical two-state value of $k \ln 2$ per particle

$$S \sim 2Nk \left(\frac{\mu B}{kT} \right) e^{-\frac{2\mu B}{kT}} \quad T \rightarrow 0$$

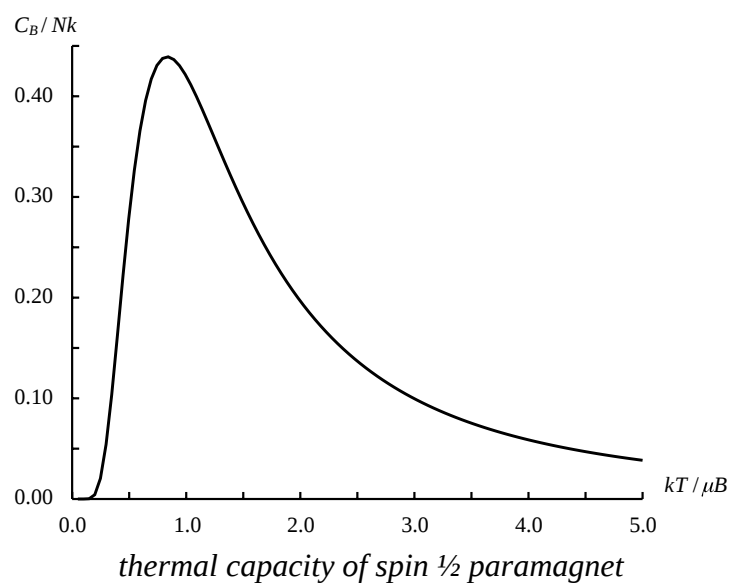
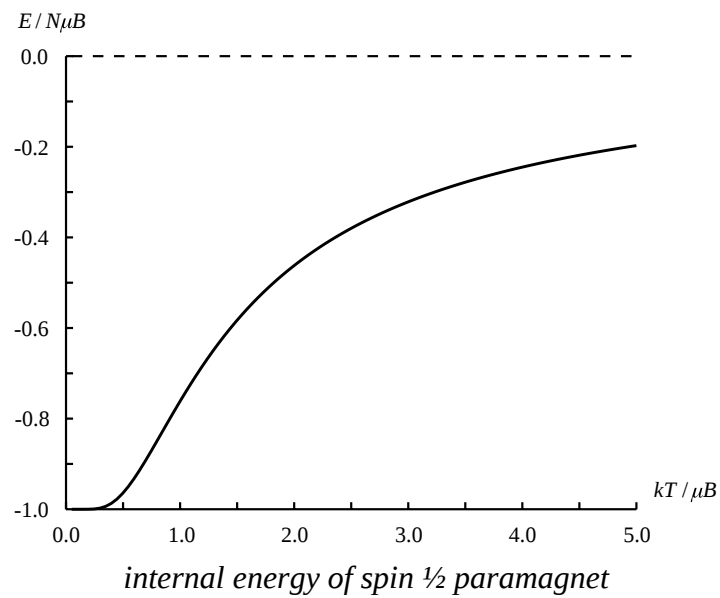
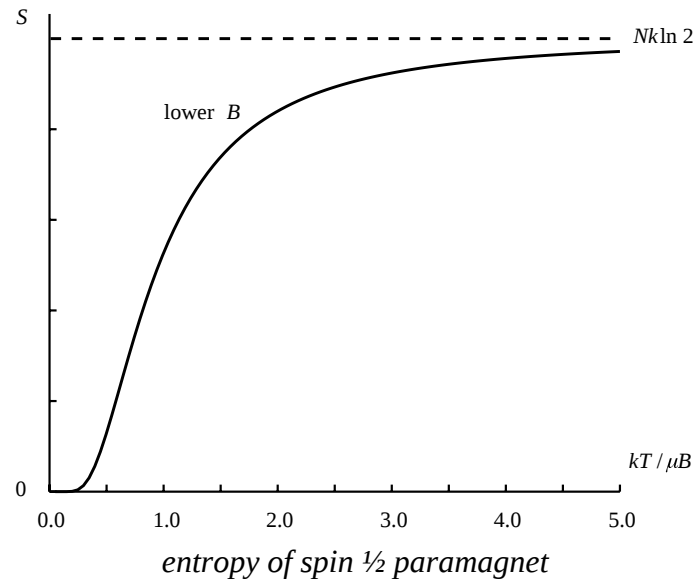
$$S \sim Nk \ln 2 - 4Nk \left(\frac{\mu B}{kT} \right)^2 \quad T \rightarrow \infty.$$

The thermal capacity is particularly important as it is readily measurable. It exhibits a maximum of $C_B \sim 0.44Nk$ at a temperature of $T \sim 0.83\mu B/k$. This is known as a Schottky anomaly. Ordinarily the thermal capacity of a substance increases with temperature, saturating at a constant value at high temperatures. Spin systems are unusual in that the energy states of a spin are finite and therefore *bounded from above*. The system then has a maximum entropy. As the entropy increases towards this maximum value it becomes increasingly difficult to pump in more heat energy.

At both high and low temperatures the thermal capacity goes to zero, as may be seen by expanding the expression for C_B :

$$C_B \sim 4Nk \left(\frac{\mu B}{kT} \right)^2 e^{-\frac{2\mu B}{kT}} \quad T \rightarrow 0$$

$$C_B \sim 2Nk \left(\frac{\mu B}{kT} \right)^2 \quad T \rightarrow \infty.$$



End of lecture 9

2.6.3 Using the partition function II — the Einstein model of a solid

One of the challenges faced by Einstein was the explanation of why the thermal capacity of solids tended to zero at low temperatures. He was concerned with nonmagnetic insulators, and he had an inkling that the explanation was something to do with quantum mechanics. The thermal excitations in the solid are due to the vibrations of the atoms. Einstein constructed a simple model of this which was partially successful in explaining the thermal capacity.

In Einstein's model each atom of the solid was regarded as a simple harmonic oscillator vibrating in the potential energy minimum produced by its neighbours. Each atom sees a similar potential, so they all oscillate at the same frequency; let us call this $\omega / 2\pi$. And since each atom can vibrate in three independent directions, the solid of N atoms is modelled as a collection of $3N$ identical harmonic oscillators.

We shall follow the procedures outlined above.

Step 1) *Write down the possible energies for the system*

The energy states of the harmonic oscillator form an infinite 'ladder' of equally spaced levels; you should be familiar with this from your quantum mechanics course.

$$\begin{array}{rcl}
 \vdots & & \vdots \\
 4 & \text{-----} & \frac{9}{2}\hbar\omega \\
 3 & \text{-----} & \frac{7}{2}\hbar\omega \\
 2 & \text{-----} & \frac{5}{2}\hbar\omega \\
 1 & \text{-----} & \frac{3}{2}\hbar\omega \\
 j = 0 & \text{-----} & \frac{1}{2}\hbar\omega
 \end{array}
 \quad \begin{array}{l}
 \text{single particle energies:} \\
 \varepsilon_j = \left(j + \frac{1}{2}\right)\hbar\omega \quad j = 0, 1, 2, 3, \dots
 \end{array}$$

Step 2) *Evaluate the partition function for the system*

We require to find the partition function for a single harmonic oscillator. This is

$$\begin{aligned}
 z &= \sum_{j=0}^{\infty} e^{-\varepsilon_j/kT} \\
 &= \sum_{j=0}^{\infty} \exp\left\{-\left(j + \frac{1}{2}\right)\frac{\hbar\omega}{kT}\right\}.
 \end{aligned}$$

To proceed, let us define a characteristic temperature θ , related to the oscillator frequency by

$$\theta = \hbar\omega / k.$$

Then the partition function may be written as

$$\begin{aligned}
 z &= \sum_{j=0}^{\infty} \exp\left\{-\left(j + \frac{1}{2}\right)\frac{\theta}{T}\right\} \\
 &= e^{-\theta/2T} \sum_{j=0}^{\infty} (e^{-\theta/T})^j
 \end{aligned}$$

and we observe the sum here to be a (convergent) geometric progression. You should recall the result

$$\sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + \dots = \frac{1}{1 - x}.$$

If you don't remember this result then multiply the power series by $1 - x$ and check the sum.

The harmonic oscillator partition function is then given by

$$z = \frac{e^{-\theta/2T}}{1 - e^{-\theta/T}}.$$

Step 3) *The Helmholtz free energy then follows from this*

Here we use the relation

$$\begin{aligned} F &= -3NkT \ln z \\ &= \frac{3N}{2}k\theta + 3NkT \ln \{1 - e^{-\theta/T}\}. \end{aligned}$$

Step 4) *All thermodynamic variables follow from differentiation of the Helmholtz free energy*

Here we have no explicit volume dependence, indicating that the equilibrium solid is at zero pressure. If an external pressure were applied then the vibration frequency $\omega / 2\pi$ would be a function of the interparticle spacing or, equivalently, the volume. This would give a volume dependence to the free energy from which the pressure could be found. However we shall ignore this.

The thermodynamic variables we are interested in are then entropy, internal energy and thermal capacity. The entropy is

$$\begin{aligned} S &= - \left. \frac{\partial F}{\partial T} \right|_V \\ &= 3Nk \left\{ \ln \left(\frac{e^{\theta/T}}{e^{\theta/T} - 1} \right) + \frac{\theta / T}{(e^{\theta/T} - 1)} \right\}. \end{aligned}$$

In the low temperature limit S goes to zero, as one would expect. The limiting low temperature behaviour is

$$S \sim 3Nk \frac{\theta}{T} e^{-\theta/T} \quad T \rightarrow 0.$$

At high temperatures the entropy tends towards a logarithmic increase

$$S \sim 3Nk \ln \left(\frac{T}{\theta} \right) \quad T \rightarrow \infty.$$

The internal energy is found from

$$E = -3N \frac{\partial}{\partial \beta} \ln z$$

where we write $\ln z$ as

$$\ln z = -\frac{\beta \hbar \omega}{2} - \ln \{1 - e^{-\beta \hbar \omega}\}.$$

Then on differentiation we find

$$E = 3N \left\{ \frac{\hbar \omega}{2} + \frac{\hbar \omega}{e^{\beta \hbar \omega} - 1} \right\}.$$

The first term represents the contribution to the internal energy from the zero point oscillations; it is present even at $T = 0$.

At high temperatures the variation of the internal energy is

$$E \sim 3NkT \left\{ 1 + \frac{1}{12} \left(\frac{\theta}{T} \right)^2 - \dots \right\}.$$

The first term is the classical equipartition part, which is independent of the vibration frequency.

Turning now to the thermal capacity, we have

$$\begin{aligned} C_V &= \left. \frac{\partial E}{\partial T} \right|_V, \\ &= 3Nk\theta \frac{\partial}{\partial T} \left\{ \frac{1}{e^{\theta/T} - 1} \right\}. \end{aligned}$$

Upon differentiation this then gives

$$C_V = 3Nk \left(\frac{\theta}{T} \right)^2 \frac{e^{\theta/T}}{(e^{\theta/T} - 1)^2}.$$

At high temperatures the variation of the thermal capacity is

$$C_V \sim 3Nk - \frac{1}{4}Nk \left(\frac{\theta}{T} \right)^2 + \dots \quad T \gg \theta$$

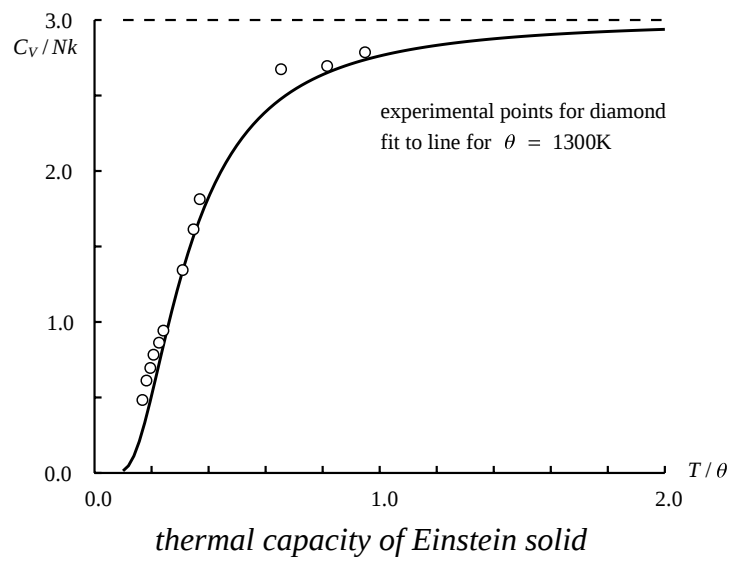
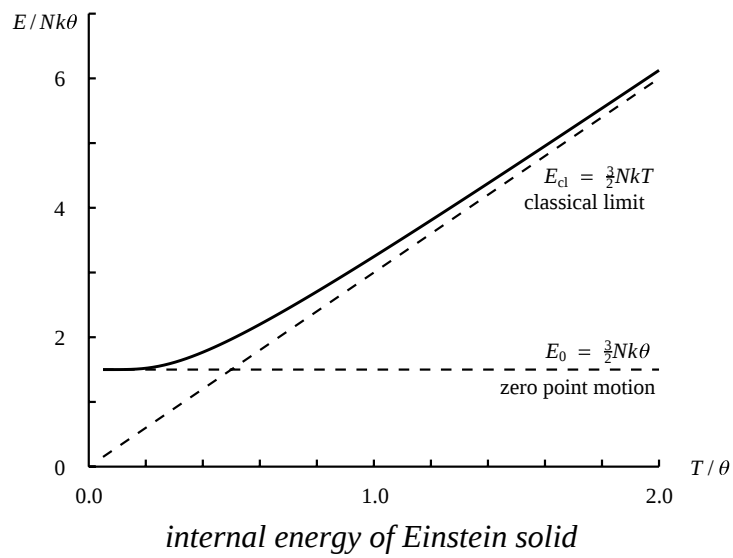
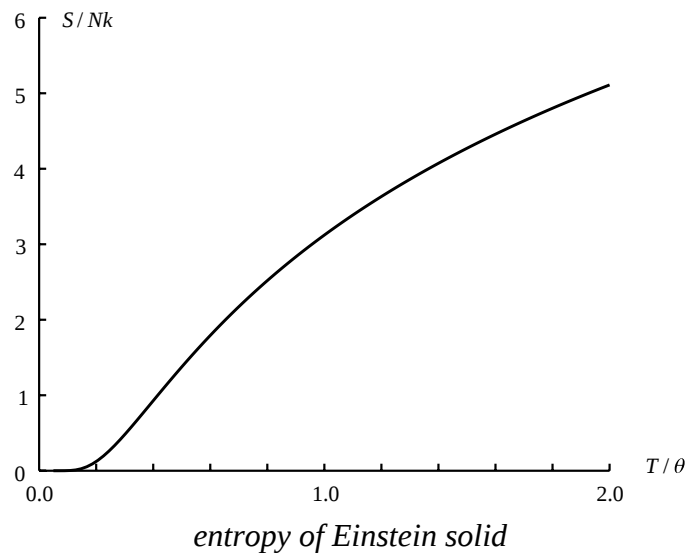
where the first term is the constant equipartition value. At low temperatures we find

$$C_V \sim 3Nk \left(\frac{\theta}{T} \right)^2 e^{-\theta/T} \quad T \ll \theta.$$

We see that this model does indeed predict the decrease of C_V to zero as $T \rightarrow 0$, which was Einstein's challenge. The decrease in heat capacity below the classical temperature-independent value is seen to arise from the quantisation of the energy levels. However the observed low temperature behaviour of such thermal capacities is a simple T^3 variation rather than the more complicated variation predicted from this model.

The explanation for this is that the atoms do not all oscillate independently. The coupling of their motion leads to normal mode waves propagating in the solid with a continuous range of frequencies. This was explained by the Debye model of solids.

The Einstein model introduces a single new parameter, the frequency $\omega / 2\pi$ of the oscillations or, equivalently, the characteristic temperature $\theta = \hbar\omega / k$. And the prediction is that C_V is a *universal* function of T / θ . To the extent that the model has some validity, each substance should have its own value of θ ; as the C_V figure shows, the value for diamond is $\sim 1300\text{K}$. This is an example of a *law of corresponding states*: When the temperatures are scaled appropriately all substances should exhibit the same behaviour.



Finally we give the calculated properties of the Einstein model in terms of hyperbolic functions; you should check these.

The partition function for simple harmonic oscillator can be written as

$$z = \frac{1}{2} \operatorname{cosech}\left(\frac{\theta}{2T}\right),$$

so that the Helmholtz free energy for the solid comprising $3N$ such oscillators is

$$F = 3NkT \ln\left[2 \sinh\left(\frac{\theta}{2T}\right)\right].$$

And from this the various properties follow:

$$S = 3Nk \left\{ \left(\frac{\theta}{2T}\right) \coth\left(\frac{\theta}{2T}\right) - \ln\left[2 \sinh\left(\frac{\theta}{2T}\right)\right] \right\}$$

$$E = \frac{3}{2}Nk\theta \coth\left(\frac{\theta}{2T}\right)$$

$$C_V = 3Nk \left(\frac{\theta}{2T}\right)^2 \operatorname{cosech}^2\left(\frac{\theta}{2T}\right).$$

End of lecture 10

Equation of state for the Einstein solid

We have no equation of state for this model so far. No $p - V$ relation has been found, since there was no volume-dependence in the energy levels, and thus in the partition function and everything which followed from that. This deficiency can be rectified by recognising that the Einstein frequency, or equivalently the temperature θ may vary with volume. Then the pressure may be found from

$$\begin{aligned} p &= - \left. \frac{\partial F}{\partial V} \right|_T \\ &= - \frac{3Nk}{2} \coth\left(\frac{\theta}{2T}\right) \frac{d\theta}{dV}. \end{aligned}$$

We note that the equilibrium state of the solid when no pressure is applied, corresponds to the vanishing of $d\theta/dV$. In fact θ will be a *minimum* at the equilibrium volume V_0 (why?). For small changes of volume from this equilibrium we may then write

$$\theta = \theta_0 + \alpha \left(\frac{V - V_0}{V_0} \right)^2$$

so that

$$\frac{d\theta}{dV} = \frac{2\alpha}{V_0^2} (V - V_0).$$

Then the equation of state for the solid is

$$p = \frac{3Nk\alpha}{V_0^2} (V_0 - V) \coth\left(\frac{\theta_0}{2T}\right).$$

The high-temperature limit of this is

$$p = \frac{6Nk\alpha}{V_0^2 \theta} (V_0 - V) T$$

and the low-temperature, temperature-independent limit is

$$p = 3Nk\alpha \frac{(V_0 - V)}{V_0^2}.$$