Readings: (K&K 2, Schroeder 6)

This week we will be following Chapter 2 of Kittel and Kroemer, which uses a microcanonical approach (or Boltzmann entropy approach) to relate entropy to temperature. This is an alternative derivation to the Gibbs approach we used last week, and it can be helpful to have seen both. In a few ways the Boltzmann approach is conceptually simpler, while there are a number of *other* ways in which the Gibbs approach is simpler.

0.0.1 Fundamental assumption

The difference between these two approaches is in what is considered the fundamental assumption. In the Gibbs entropy approach we assumed that the entropy was a "nice" function of the probabilities of microstates, which gave us the Gibbs formula. From there, we could maximize the entropy to find the probabilities under some set of constraints.

The Boltzmann approach makes what is perhaps simpler assumption, which is that if only microstates with a given energy are permitted, then all of the microstates with that energy are equally probable. (This scenario with all microstates having the same energy is the **microcanonical ensemble**.) Thus the macrostate with the most corresponding microstates will be most probable macrostate. The number of microstates corresponding to a given macrostate is called the **multiplicity** $g(E, V)$. In this approach, multiplicity (which did not show up last week!) becomes a fundamentally important quantity, since the macrostate with the highest multiplicity is the most probable macrostate.

Outline of the week One or two topics per day:

1. Quick version showing the conclusions we will reach.
2. Finding the multiplicity of a paramagnet (Chapter 1).
3. *(Probably skipped in class)* Combining two non-interacting systems; defining temperature.

0.1 Quick version

This quick version will tell you all the essential physics results for the week, without proof. The beauty of statistical mechanics (whether following the text or using the information-theory approach of last week) is that you don't *actually* need to take on either faith or experiment the connection between the statistical theory and the empirical definitions used in thermodynamics.

0.1.1 Entropy

The multiplicity sounds sort of like entropy (since it is maximized), but the multiplicity is *not* extensive (*nor* intensive), because the number of microstates for two identical systems taken together is the *square* of the number of microstates available to one of the single systems. This naturally leads to the Boltzmann definition of the entropy, which is

$$S(E, V) = k_B \ln g(E, V). \tag{1}$$

The logarithm converts the multiplicity into an extensive quantity, in a way that is directly analogous to the logarithm that appears in the Gibbs entropy.

For large systems (e.g. systems composed of $\sim 10^{23}$ particles\$), the *most probable* configuration is essentially the same as *any remotely probable* configuration. This comes about due for the same reason that if you flip 10^{23} coins, you will get $5 \times 10^{22} \pm 10^{12}$ heads. On an absolute scale, that's a lot of uncertainty in the number of heads that would show up, but on a fractional scale, you're pretty accurate if you assume that 50% of the flips will be heads.

0.1.2 Temperature

From *Energy and Entropy* (and last week), you will remember that $dU = TdS - pdV$, which tells us that $T = \left(\frac{\partial U}{\partial S}\right)_V$. If we assume that only states with one particular energy E have a non-zero probability of being occupied, then $U = E$, i.e. the thermodynamic internal energy is the same as the energy of any allowed microstate. Then we can replace U with E and conclude that

$$T = \left(\frac{\partial E}{\partial S}\right)_V \tag{2}$$

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_V \tag{3}$$

$$= \left(\frac{\partial k_B \ln g(E,V)}{\partial E}\right)_V \tag{4}$$

$$= k_B \frac{1}{g}\left(\frac{\partial g}{\partial E}\right)_V \tag{5}$$

From this perspective, it looks like our job is to learn to solve for $g(E)$ and from that to find $S(E)$, and once we have done those tasks we will know the temperature (and soon everything else).

Differentiable multiplicity The above assumes that $g(E)$ is a differentiable function, which means that the number of microstates must be a continuous function of energy! This highlights one of the distinctions between the microcanonical approach and our previous (cannonical) Gibbs approach.

In reality, we know from quantum mechanics that any system of finite size has a finite number of eigenstates within any given energy range, and thus $g(E)$ cannot be either continuous or differentiable. Boltzmann, of course, did not know this, and assumed that there were an infinite number of microstates possible within any energy range, and would strictly speaking interpret $g(E)$ in terms of a volume of phase space.

The resolution to this conundrum is to invoke large numbers, and to assume that we are averaging $g(E)$ over a range of energies in which there are many, many states. For real materials with $N \approx 10^{23}$, this assumption is pretty valid. Much of this chapter will involve learning to work with this large N assumption, and to use it to extract physically meaningful results. In the Gibbs approach this large N assumption was not needed.

As Kittel discusses towards the end of the chapter, we only really need to know $g(E)$ up to some constant factor, since a constant factor in g becomes a constant additive change in S, which doesn't have any physical impact.

The "real" $g(E)$ is a smoothed average over a range of energies. In practice, doing this can be confusing, and so we tend to focus on systems where the energy is always an integer multiple of some constant. Thus a focus on spins in a magnetic field, and harmonic oscillators.

0.2 Multiplicity of a paramagnet

So now the question becomes how to find the number of microstates that correspond to a given energy $g(E)$. Once we have this in an analytically tractable form, we can everything else we might care for (with effort). This is essentially a counting problem, and much of what you need is introduced in Chapter 1. We will spend some class time going over one example of computing the multiplicity. Consider a paramagnetic system consisting of spin $\frac{1}{2}$ particles that can be either up or down. Each spin has a magnetic moment in the $\hat{z}$ direction of $\pm m$, and we are interested in the total magnetic moment μ_{tot}, which is the sum of all the individual magnetic moments. Note that the magnetization M used in electromagnetism is just the total magnetic moment of the material divided by its volume.

$$M \equiv \frac{\mu_{tot}}{V} \tag{6}$$

Note It is confusingly common to refer to the total magnetic moment as the magnetization. Given either a numerical value or an expression, it's usually easy to tell what you've got by checking the dimensions.

Small Group Question Work out how many ways a system of 4 spins can have any possible magnetization of enumerating all the microstates corresponding to each magnetization.

Now find a mathematical expression that will tell you the multiplicity of a system with an even number N spins and just one $\uparrow$ spin. Then find the multiplicity for two $\uparrow$ spins, and for three $\uparrow$ spins.

Now find a mathematical expression that will tell you the multiplicity of a system with an even number N spins and total magnetic moment $\mu_{tot} = 2sm$ where s is an integer. We call s the **spin excess**, since $N_\uparrow = \frac{1}{2}N + s$. Alternatively, you could write your expression in terms of the number of up spins $N_\uparrow$ and the number of down spins $N_\downarrow$.

Answer We can enumerate all spin microstates:

$\mu_{tot} = -4m$ $\downarrow\downarrow\downarrow\downarrow$ g=1

$\mu_{tot} = -2m$ $\downarrow\downarrow\downarrow\uparrow$ $\downarrow\downarrow\uparrow\downarrow$ $\downarrow\uparrow\downarrow\downarrow$ $\uparrow\downarrow\downarrow\downarrow$ g=4

$\mu_{tot} = 0$ $\downarrow\downarrow\uparrow\uparrow$ $\downarrow\uparrow\uparrow\downarrow$ $\uparrow\uparrow\downarrow\downarrow$ $\uparrow\downarrow\downarrow\uparrow$ $\uparrow\downarrow\uparrow\downarrow$ $\downarrow\uparrow\downarrow\uparrow$ g=6

$\mu_{tot} = 2m$ $\uparrow\uparrow\uparrow\downarrow$ $\uparrow\uparrow\downarrow\uparrow$ $\uparrow\downarrow\uparrow\uparrow$ $\downarrow\uparrow\uparrow\uparrow$ g=4

$\mu_{tot} = 4m$ $\uparrow\uparrow\uparrow\uparrow$ g=1

To generalize this to $g(N, s)$, we need to come up with a systematic way to count the states that have the same spin excess s. Clearly if $s = \pm N/2$, $g = 1$, since that means that all the spins are pointed the same way, and there is only one way to do that.

$$g(N, s = \pm\frac{1}{2}N) = 1 \tag{7}$$

Now if we have just one spin going the other way, there are going to be N ways we could manage that:

$$g\left(N, s = \pm\left(\frac{1}{2}N - 1\right)\right) = N \tag{8}$$

Now when we go to flip it so we have two spins up, there will be $N-1$ ways to flip the second spin. But then, when we do this we will end up counting every possibility twice, which means that we will need to divide by two.

$$g\left(N, s = \pm\left(\frac{1}{2}N - 2\right)\right) = N(N-1)/2 \tag{9}$$

When we get to adding the third $\uparrow$ spin, we'll have $N-2$ spins to flip. But now we have to be even more careful, since for the same three up-spins, we have several ways to reach that microstate. In fact, we will need to divide by 6, or 3×2 to get the correct answer (as we can check for our four-spin example).

$$g\left(N, s = \pm\left(\frac{1}{2}N - 3\right)\right) = \frac{N(N-1)(N-2)}{3!} \tag{10}$$

At this stage we can start to see the pattern, which comes out to

$$g\left(N, s\right) = \frac{N!}{\left(\frac{1}{2}N + s\right)!\left(\frac{1}{2}N - s\right)!} \tag{11}$$

$$= \frac{N!}{N_\uparrow! N_\downarrow!} \tag{12}$$

0.2.1 Stirling's approximation

As you can see, we now have a bunch of factorials. Once we compute the entropy, we will have a bunch of logarithms of factorials.

$$N! = \prod_{i=1}^{N} i \tag{13}$$

$$\ln N! = \ln\left(\prod_{i=1}^{N} i\right) \tag{14}$$

$$= \sum_{i=1}^{N} \ln i \tag{15}$$

So you can see that the log of a factorial is a sum of logs. When the number of things being summed is large, we can approximate this sum with an integral. This may feel like a funny business, particularly for those of you who took my computational class, where we frequently used sums to approximate integrals!

But the approximation can go both ways. In this case, if we approximate the integral as a sum we can find an analytic expression for the factorial:

$$\ln N! = \sum_{i=1}^{N} \ln i \tag{16}$$

$$\approx \int_1^N \ln x dx \tag{17}$$

$$= x \ln x - x|_1^N \tag{18}$$

$$= N \ln N - N + 1 \tag{19}$$

At this point, we should recognize that the 1 that we see is much smaller than the other two terms, and is actually likely to be wrong. Importantly, there is a larger error being made here, which we can see if we zoom into the upper end of our integral. We are missing $\frac{1}{2} \ln N!$ The reason is that our integral went precisely to N, but if we imagine a midpoint rule picture (or trapezoidal rule) we are missing half of that last point. This gives us:

$$\ln N! \approx \left(N + \frac{1}{2}\right) \ln N - N \tag{20}$$

We could find the constant term correctly (it is not 1), but that is more work, and even the $\frac{1}{2}$ above is usually omitted when using Stirling's approximation, since it is much smaller than the others when $N \gg 0$

0.3 Entropy of our spins

I'm going to use a different approach than the text to find the entropy of this spin system when there are many spins and the spin excess is relatively small.

$$S = k \ln g\left(N, s\right) \tag{21}$$

$$= k \ln \left(\frac{N!}{\left(\frac{1}{2}N + s\right)! \left(\frac{1}{2}N - s\right)!} \right) \tag{22}$$

$$= k \ln \left(\frac{N!}{N_\uparrow! N_\downarrow!} \right) \tag{23}$$

$$= k \ln \left(\frac{N!}{(h+s)!\,(h-s)!} \right) \tag{24}$$

At this point I'm going to define for convenience $h \equiv \frac{1}{2}N$, just to avoid writing so many $\frac{1}{2}$. I'm also going to focus on the s dependence of the entropy.

$$\frac{S}{k} = \ln\left(N!\right) - \ln\left(N_\uparrow!\right) - \ln\left(N_\downarrow!\right) \tag{25}$$

$$= \ln N! - \ln(h+s)! - \ln(h-s)! \tag{26}$$

$$= \ln N! - \sum_{i=1}^{h+s} \ln i - \sum_{i=1}^{h-s} \ln i \tag{27}$$

At the last step, I wrote the log of the factorial as a sum of logs. This is still looking pretty hairy. So let's now consider the *difference* between the entropy with s and the entropy when $s = 0$ (which I will call here S_0 for compactness and convenience).

$$\frac{S(s) - S_0}{k_B} = -\sum_{i=1}^{h+s} \ln i - \sum_{i=1}^{h-s} \ln i + \sum_{i=1}^{h} \ln i + \sum_{i=1}^{h} \ln i \tag{28}$$

$$= -\sum_{i=h+1}^{h+s} \ln i + \sum_{j=h-s+1}^{h} \ln j \tag{29}$$

where I have changed the sums to account for the difference between the sums with s and those without. At this stage, our indices are starting to feel a little inconvenient given the short range we are summing over, so let's redefine our index ov summation so the sums will run up to s. In preparation for this, at the last step, I renamed one of my dummy indexes.

$$i = h + k \qquad\qquad j = h + 1 - k \tag{30}$$

With these indexes, each sum can go from $k = 1$ to $k = s$, which will enable us to combine our sums into one.

$$\frac{S - S_0}{k} = -\sum_{k=1}^{s} \ln(h + k) + \sum_{k=1}^{s} \ln(h + 1 - k) \tag{31}$$

$$= \sum_{k=1}^{s} (\ln(h + 1 - k) - \ln(h + k)) \tag{32}$$

At this point, if you're anything like me, you're thinking "I could turn that difference of logs into a log of a ratio!" Sadly, this doesn't turn out to help us. Instead, we are going to start trying to get the h out of the way in preparation for taking the limit as $s \ll h$.

$$\frac{S - S_0}{k} = \sum_{k=1}^{s} \ln h + \ln\left(1 - \frac{k-1}{h}\right) - \ln h - \ln\left(1 + \frac{k}{h}\right) \tag{33}$$

$$= \sum_{k=1}^{s} \left(\ln\left(1 - \frac{k-1}{h}\right) - \ln\left(1 + \frac{k}{h}\right)\right) \tag{34}$$

It is now time to make our first approximation: we assume $s \ll N$, which means that $s \ll h$. That enables us to simplify these logarithms drastically! ☺

$$\frac{S - S_0}{k} \approx \sum_{k=1}^{s} \left(-\frac{k-1}{h} - \frac{k}{h}\right) \tag{35}$$

$$= -\frac{2}{h} \sum_{k=1}^{s} \left(k - \tfrac{1}{2}\right) \tag{36}$$

$$= -\frac{4}{N} \sum_{k=1}^{s} \left(k - \tfrac{1}{2}\right) \tag{37}$$

Missing /var/www/paradigms_media_2/media/activity_media/sum-gives-square.pdf

Now we have this sum to solve. You can find this sum either geometrically or with calculus. The calculus involves turning the sum into an integral. As you can see in the figure, the integral

$$\int_0^s x dx = \tfrac{1}{2}s^2 \tag{38}$$

has the same value as the sum, since the area under the orange curve (which is the sum) is equal to the area under the blue curve (which is the integral).

The geometric way to solve this looks visually very much the same as the integral picture, but instead of computing the area from the straight line, we cut the stair-step area "half" and fit the two pieces together such that they form a rectangle with width $s/2$ and height s.

Missing /var/www/paradigms_media_2/media/activity_media/sum-geometrically.pdf

Figure 1: Sum geometrically

Taken together, this tells us that when $s \ll N$

$$S(N,s) \approx S(N,s=0) - k\frac{4}{N}\frac{s^2}{2} \tag{39}$$

$$= S(N,s=0) - k\frac{2s^2}{N} \tag{40}$$

This means that the multiplicity is gaussian:

$$S = k \ln g \tag{41}$$

$$g(N,s) = e^{\frac{S(N,s)}{k}} \tag{42}$$

$$= e^{\frac{S(N,s=0)}{k} - \frac{2s^2}{N}} \tag{43}$$

$$= g(N,s=0)e^{-\frac{2s^2}{N}} \tag{44}$$

Thus the multiplicity (and thus probability) is peaked at $s = 0$ as a gaussian with width $\sim \sqrt{N}$. This tells us that the width of the peak *increases* as we increase N. However, the excess spin *per particle decreases* as $\sim \frac{1}{\sqrt{N}}$. So that means that our fractional polarization becomes far more sharply peaked as we increase N.

0.4 Thermal contact (probably skipped in class)

Suppose we put two systems in contact with one another. This means that energy can flow from one system to the other. We assume, however, that the contact between the two systems is weak enough that their energy eigenstates are unaffected. This is a bit of a contradiction you'll need to get used to: we treat our systems as non-interacting, but assume there is some energy transfer between them. The reasoning is that the interaction between them is very small, so that we can treat each system separately, but energy can still flow.

We ask the question: "How much energy will each system end up with after we wait for things to settle down?" The answer to this question is that energy will settle down in the way that maximizes the number of microstates.

Let us consider two simple systems: a 2-spin paramagnet, and a 4-spin paramagnet.

System A A system of 3 spins each with energy ± 1. This system has the following multiplicity found from Pascal's triangle:

1			
1	1		
1	2	1	
1	3	3	1
-3	-1	1	3

System B A system of 4 spins each with energy ± 1. This system has the following multiplicity found from Pascal's triangle:

1				
1	1			
1	2	1		
1	3	3	1	
1	4	6	4	1
-4	-2	0	2	4

Question What is the total number of microstates when you consider systems A and B together as a combined system? Answer

We need to multiply the numbers of microstates for each system separately, because for each microstate of A, it is possible to have B be in any of its microstates. So the total is $2^3 2^4 = 128$.

Since we have two separate systems here, it is meaningful to ask what the probability is for system A to have energy E_A, given that the combined system has energy E_{AB}.

Small group question What is the multiplicity of the combined system if the energy is 3, i.e. $g_{AB}(E_{AB} = 3)$?

Answer To solve this, we just need to multiply the multiplicities of the two systems and add up all the energy possibilities that total 3:

$$g_{AB}(E_{AB} = 0) = g_A(-1)g_B(4) + g_A(1)g_B(2) + g_A(3)g_B(0) \tag{45}$$
$$= 3 \cdot 1 + 3 \cdot 4 + 1 \cdot 6 \tag{46}$$
$$= 21 \tag{47}$$

Small group question What is the probability that system A has energy 1, if the combined energy is 3?

Answer To solve this, we just need to multiply the multiplicities of the two systems, which we already found and divide by the total number of microstates:

$$P(E_A = 1 | E_{AB} = 3) = \frac{g_A(1)g_B(2)}{g_{AB}(3)} \tag{48}$$
$$= \frac{3 \cdot 4}{21} \tag{49}$$
$$= \frac{4}{7} \tag{50}$$

which shows that this is the most probable distribution of energy between the two subsystems.

Given that these two systems are able to exchange energy, they ought to have the same temperature. To find the most probable energy partition between the two systems, we need to find the partition that maximizes the multiplicity of the combined system:

$$g_{AB}(E_A) = g_A(E_A)g_B(E_{AB} - E_A) \tag{51}$$

$$0 = \frac{dg_{AB}}{dE_A} \tag{52}$$

$$= g'_A g_B - g'_B g_A \tag{53}$$

$$\frac{g'_A}{g_A} = \frac{g'_B}{g_B} \tag{54}$$

$$\frac{1}{g_A(E_A)}\frac{\partial g_A(E_A)}{\partial E_A} = \frac{1}{g_B(E_B)}\frac{\partial g_B(E_B)}{\partial E_B} \tag{55}$$

This tells us that the "thing that becomes equal" when the two systems are in thermal contact is this strange ratio of the derivative of the multiplicity with respect to energy divided by the multiplicity itself. You may be able to recognize this as what is called a logarithmic derivative.

$$\frac{\partial}{\partial E_A}\ln(g_A(E_A)) = \frac{1}{g_A(E_A)}\frac{\partial g_A(E_A)}{\partial E_A} \tag{56}$$

thus we can conclude that when two systems are in thermal contact, the thing that equalizes is

$$\beta \equiv \left(\frac{\partial \ln g}{\partial E}\right)_V \tag{57}$$

At this stage, we haven't shown that $\beta = \frac{1}{kT}$, but we have shown that it should be a function of T, since T is also a thing that is equalized when two systems are in thermal contact.

By dimensional reasoning, you can recognize that this *could* be $\frac{1}{kT}$, and we're just going to leave this at that.