PHYS 427: Discussion 1

Jan. 28, 2025

1. ***Binomial distribution.*** You are given a glass box containing a gas of N particles. The particles almost never collide with each other (we say they are approximately *non-interacting*). The particles are moving quickly around the box and bouncing off the walls. When you look closely, you notice that each particle is labeled with a number from 1 to N that distinguishes it from the other particles. You imagine dividing the box into M imaginary parts of equal volume, which we'll call *zones*.

   (a) With a camera, you take a picture of the box. Looking at the picture, you carefully mark down which zone each particle is in. For example, you might find:

particle	1	2	3	$\cdots$	N
found in zone	3	8	1	$\cdots$	3

   This table defines a *microstate* of the system. How many distinct microstates are there? In how many microstates is particle 3 located in the first zone?

   Solution. Think about placing each particle one by one. The particle labeled 1 can go in any of the M zones. The particle labeled 2 can go in any of the M zones. So there are M^2 distinct ways to place the first two particles. Continuing in this way, there are M^N distinct microstates.

   If we know particle 3 is located in the first zone, then the counting looks like $M \cdot M \cdot 1 \cdot M \cdots M = M^{N-1}$.

   Since every microstate is equally likely, the probability is just given by the fraction of microstates in which particle 3 is in the first zone. That is, $p = M^{N-1}/M^N = 1/M$.

   Each microstate is equally probable (make sure this sounds reasonable to you). From your previous answers, calculate the probability that particle 3 will be in the first zone when you take a picture.

   (b) You take many more pictures of the box. In each picture, you count how many particles are in the first zone, without caring about the labels on the particles. Then you take the average of these counts. Without doing any calculation, what do *expect* the average count to be? (We will calculate this in part (f).)

   Solution. *"Never make a calculation until you know the answer. Make an estimate before every calculation, try a simple physical argument (symmetry! invariance! conservation!) before every derivation, guess the answer to every paradox and puzzle. Courage: No one else needs to know what the guess is. Therefore make it quickly, by instinct. A right guess reinforces this instinct. A wrong guess brings the refreshment of surprise. In either case life as a [physicist], however long, is more fun!"* - John Wheeler

   (c) How many microstates have exactly n particles in the first zone (without caring about the labels on the particles)?

   Solution. Here is one way to think about it. Consider the microstates in which the first n particles are located in zone 1. These microstates look like

particle	1	2	$\cdots$	n	n+1	n+2	$\cdots$	N
found in zone	1	1	1	1	any except 1	any except 1	any except 1	any except 1

How many of these microstates are there? Well, there are $M - 1$ choices for the zone of particle $n + 1$ (because it can't be in zone 1), and $M - 1$ choices for the zone of particle $n + 2$, etc. So there are $(M - 1)^{N-n}$ microstates like this.

But it didn't have to be the case that the *first* n particles were in zone 1, just that *some* n particles are in zone 1. There are $\binom{N}{n} = N!/(n!(N - n)!)$ ways of choosing n of the N particles to be in zone 1 (see lecture 1 of week 1 for the proof). Each such choice, by the argument above, has $(M - 1)^{N-n}$ associated microstates. So the answer is $\binom{N}{n} \cdot (M - 1)^{N-n}$.

(d) Using your answer to part (c), show that the probability of finding exactly n particles in the first zone is equal to

$$\binom{N}{n} p^n (1 - p)^{N-n},$$

where

$$\binom{N}{n} \equiv \frac{N!}{n!(N - n)!}$$

is the **binomial coefficient** and $p = 1/M$ is the probability that you found in part (a).

Solution. The reasoning is pretty much the same as in the end of part (a). All microstates are equally likely, so the probability that the system is in a certain macrostate (specified here by n) is equal to the fraction of all microstates that are associated to that macrostate. So we just divide the answer to part (c) by the total number of microstates M^N.

$$\frac{\binom{N}{n}(M - 1)^{N-n}}{M^N} = \frac{\binom{N}{n}(M - 1)^{N-n}}{M^{N-n}M^n} = \frac{\binom{N}{n}(1 - 1/M)^{N-n}}{M^n} = \binom{N}{n} p^n (1 - p)^{N-n}.$$

(e) Explain why the expected (i.e. average) number of particles in the first zone is given by

$$\langle n \rangle = \sum_{n=0}^{N} n \binom{N}{n} p^n (1 - p)^{N-n}.$$

(The first zone isn't special, so this is the expected number of particles in any one zone.)

Solution. Consider taking $N_{\text{pics}} \gg 1$ pictures. Let a_n denote the number of pictures in which there are exactly n particles in the first zone. Among all of the pictures, the total number of particles counted the first zone is $a_1 + 2 \cdot a_2 + \cdots + N \cdot a_n$. The average number of particles in the first zone in any one picture, by the definition of the average, is then

$$\frac{a_1 + 2 \cdot a_2 + \cdots + N \cdot a_N}{N_{\text{pics}}} = \sum_{n=1}^{N} n \frac{a_n}{N_{\text{pics}}}.$$

When $N_{\text{pics}} \gg 1$, $\frac{a_n}{N_{\text{pics}}}$ approaches the probability of finding exactly n particles in the first zone, which we computed in the previous part.

Another perfectly good answer to this question is "because that's the definition of the expected value". But this shows *why* that is the definition.

(f) Compute the sum in part (e). Does it agree with your expectations from part (b)? *Hint: There are multiple ways to do this. A slick way involves taking a derivative of both sides of the **binomial identity**,*

$$(p + q)^N = \sum_{n=0}^{N} \binom{N}{n} p^n q^{N-n},$$

which holds for any p and q. [If it's unfamiliar, prove it by multiplying out the left-hand side.]

Solution. Let's do it the slick way first. Consider the operator $p \frac{\partial}{\partial p}$, where q is held fixed in the partial derivative. Apply this operator to both sides of the binomial identity to get

$$Np(p + q)^{N-1} = \sum_{n=0}^{N} n \binom{N}{n} p^n q^{N-n}.$$

2

Now evaluate this equation for $q = 1 - p$. We find

$$Np = \sum_{n=1}^{N} n \binom{N}{n} p^n (1-p)^{N-n},$$

and we're done. To do it the hard way, we will evaluate the sum by brute force.

$$
\begin{aligned}
\sum_{n=0}^{N} n \binom{N}{n} p^n (1-p)^{N-n} &= \sum_{n=1}^{N} n \binom{N}{n} p^n (1-p)^{N-n} \\
&= \sum_{n=1}^{N} n \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n} \\
&= \sum_{n=1}^{N} \frac{N!}{(n-1)!(N-n)!} p^n (1-p)^{N-n} \\
&= \sum_{m=0}^{N-1} \frac{N!}{m!(N-m-1)!} p^{m+1} (1-p)^{N-m-1} \\
&= Np \sum_{m=0}^{N-1} \frac{(N-1)!}{m!([N-1]-m)!} p^m (1-p)^{[N-1]-m}
\end{aligned}
$$

In the first equality we noticed that the $n = 0$ term is zero anyway, so we can start the summation at $n = 1$. In the fourth equality, we changed the summation variable to $m = n - 1$. In the last equality we rearranged things so it's clear that the sum is equal to the right-hand side of the binomial identity, but with $N - 1$ instead of N and with $q = 1 - p$. From the left-hand side of the binomial identity, this sum is just equal to 1, and we're done.

2. **Definition of temperature.** In this problem, we will motivate the definition of temperature by considering two systems in thermal contact. It's the same thing you saw in lecture today, so if you're feeling comfortable with that, you can skip this problem.

Think back to problem 1(c). If we know the number n of particles in the first zone, we will say that we know the **macrostate**. What we calculated in problem 1(d) is called the **multiplicity** $\Omega(n)$ of the macrostate. Below is a sketch of $\Omega(n)$ when the total number of particles N is very large. From the plot, we see that the *vast majority* of microstates belong to a very small range of macrostates, namely those for which n is very close to the expected value $\langle n \rangle$. The **fundamental assumption of statistical mechanics** states that, for a closed system in thermal equilibrium, all microstates are equally likely. Therefore, the most likely *macrostate* by far is $n \approx \langle n \rangle$. ***Lesson: in equilibrium, the system settles into the macrostate with the highest multiplicity Ω!!!***
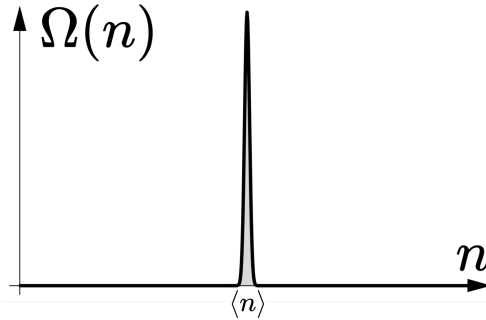
Ok. Now forget about the zones, and focus your attention on the total energy of all the particles in the box. Someone gives you another box containing another gas. You connect the two boxes with an

energy-permeable barrier. This barrier is simply a rubber membrane (a drumhead). Gas particles from box 1 can strike the membrane, transferring vibrational energy into the membrane. The vibrating membrane can then strike gas particles from box 2, transferring kinetic energy into box 2. In this way, energy can be exchanged between the boxes, but it cannot be created or destroyed. The particles themselves cannot pass through the membrane.
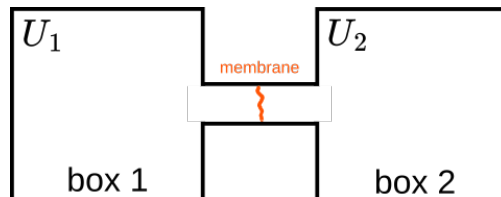
If we know the energy U_1 of the gas in box 1, we will say we know the macrostate of box 1. Similarly, U_2 specifies the macrostate of box 2. The multiplicity functions for the two boxes are denoted $\Omega_1(U_1)$ and $\Omega_2(U_2)$. (They are not necessarily the same function, because the boxes may be a different size and may contain different amounts of gas.)

(a) Show that, when the entire system (box 1 + box 2) settles into thermal equilibrium, we will have

$$\frac{\partial \ln \Omega_1}{\partial U_1} = \frac{\partial \ln \Omega_2}{\partial U_2}. \tag{1}$$

Solution. When box 1 has energy U_1 and box 2 has energy U_2, the total number if microstates of the combined system (box 1 + box 2) is $\Omega = \Omega_1(U_1)\Omega_2(U_2)$. That's because, for each microstate of box 1, box 2 can be any of its $\Omega_2(U_2)$ microstates. The states of the two boxes, by assumption, are completely independent, i.e. knowing the state of box 1 tells you nothing about the state of box 2. (This is not exactly true, because the gases do interact via the membrane. However, as always, we assume the interaction is weak. This is certainly true if the gases are much larger than the membrane.)

Also, the total energy $U = U_1 + U_2$ is conserved, i.e. it remains constant. So we can write $\Omega = \Omega_1(U_1)\Omega_2(U - U_1)$. From the "lesson" at the beginning of this question, which is really the second law of thermodynamics for the closed system (box 1 + box 2), we assert that Ω is maximal in thermal equilibrium. That is, the boxes will transfer energy (changing U_1) until Ω reaches a maximum. So we set $d\Omega/dU_1 = 0$.

$$\begin{aligned} 0 &= \frac{d}{dU_1}[\Omega_1(U_1)\Omega_2(U - U_1)] \\ &= \Omega_1'(U_1)\Omega_2(U - U_1) - \Omega_1(U_1)\Omega_2'(U - U_1), \end{aligned}$$

where the prime denotes differentiation of the function with respect to its argument, and we have used the chain rule. Reinserting $U - U_1 = U_2$ and rearranging, we find

$$\frac{1}{\Omega_1(U_1)}\Omega_1'(U_1) = \frac{1}{\Omega_2(U_2)}\Omega_2'(U_2),$$

or, reverting to the usual notation,

$$\frac{1}{\Omega_1(U_1)}\frac{\partial \Omega_1}{\partial U_1} = \frac{1}{\Omega_2(U_2)}\frac{\partial \Omega_2}{\partial U_2}.$$

The reason we use partials here is to remind ourselves that the volume and particle number are fixed. Finally, using $(d/dx)\ln x = (\ln x)/x$, we get the desired final answer.

4

You have found a quantity which is the same for two systems in *thermal contact* when they reach thermal equilibrium. That's exactly the sort of behavior we expect of the quantity called "temperature". The **temperature** T_1 of box 1 is *defined* as

$$\frac{1}{T_1} \equiv k_B \frac{\partial \ln \Omega_1}{\partial U_1}, \tag{2}$$

where $k_B \approx 1.38 \times 10^{-23} J/K$ is called Boltzmann's constant, introduced for historical reasons. (In a perfect world, we would set k_B to 1 and measure temperature in the same units as energy.) We define the **entropy** of box 1 as

$$S_1 \equiv k_B \ln \Omega_1, \tag{3}$$

so that

$$\frac{1}{T_1} = \frac{\partial S_1}{\partial U_1}. \tag{4}$$

(b) Suppose that initially $T_1 > T_2 > 0$. Show that box 1 will lose energy (and box 2 will gain energy) as the combined system approaches equilibrium. We say heat flows from hot to cold.

Solution. Here is a graphical solution. By assumption we have $\partial S_2/\partial U_2 > \partial S_1/\partial U_1 > 0$ initially. That means, in the vicinity of the initial energies U_1 and U_2, the slope of the graph of $S_2(U_2)$ is steeper than that of $S_1(U_1)$, and both slopes are positive. Now imagine that box 1 gains a tiny bit of energy. Then box 2 must lose that same amount of energy. From the graphs of $S_2(U_2)$ and $S_1(U_1)$, you can see that box 1 will gain some entropy and box 2 will lose some entropy. But, because the slopes are different, *the entropy gained by box 1 is less than the entropy lost by box 2*, so the total entropy of the universe decreases. This contradicts the second law of thermodynamics. On the other hand, if box 1 loses a little bit of energy (and box 2 gains it), the total entropy will increase, so this will happen spontaneously.

If we had foolishly defined temperature as $T = \partial S/\partial U$ instead, then we would have to say it's hot outside on a snowy day and cold on a sunny day. We would have to say "it's not possible to increase the temperature above absolute infinity". We would get strange looks, but there wouldn't be anything wrong with such a definition.

3. ***Paramagnet.*** Consider a system of N identical non-interacting spin-1/2 particles that are are fixed in place, e.g. on the sites of a crystal lattice. Each particle can be in one of two states: spin pointed along $+\hat{\boldsymbol{z}}$ ("spin up", denoted $\uparrow$) or spin pointed along $-\hat{\boldsymbol{z}}$ ("spin down", denoted $\downarrow$). In a uniform external magnetic field $\boldsymbol{B} = B\hat{\boldsymbol{z}}$, a spin can lower its energy by aligning with the field. The energy of a particle is $-\mu B$ if it's spin-up and $+\mu B$ if it's spin-down. Here μ is the magnetic moment, which characterizes the strength of the interaction between the spin and the magnetic field.

Treat the collection of spins as a closed system in thermal equilibrium.

(a) Let $N_\uparrow$ denote the total number of particles that are spin-up. Write the total energy U of the system in terms of $N_\uparrow$ and the other variables defined above.

Solution. From the problem statement, we deduce $U = -\mu B \cdot N_\uparrow + \mu B \cdot N_\downarrow = \mu B(N_\downarrow - N_\uparrow)$. Then, inserting $N_\downarrow = N - N_\uparrow$, we find $U = \mu B(N - 2N_\uparrow)$.

(b) Compute the entropy S of the system as a function of U, N and B in the limit where N, $N_\uparrow$ and $N_\downarrow \equiv N - N_\uparrow$ are each much greater than 1. *Hint: First find the multiplicity Ω in terms of N, $N_\uparrow$ and $N_\downarrow$, then compute S using Stirling's approximation to simplify: $\ln(q!) \approx q \ln q - q$ when $q \gg 1$. Finally, use your answer to part (a) to express $N_\uparrow$ and $N_\downarrow$ in terms of U, N and B. The answer will not be pretty.*

Solution. In the macrostates where a certain number $N_\uparrow$ of the particles are spin-up, the multiplicity is the number of ways of choosing $N_\uparrow$ particles out of a set of N, i.e. $\Omega = \binom{N}{N_\uparrow} = \frac{N!}{N_\uparrow! N_\downarrow!}$. Using Stirling (which is justified because $N, N_\uparrow, N_\downarrow \gg 1$ by assumption, we have

$$\ln \Omega = \ln N! - \ln N_\uparrow! - \ln N_\downarrow!$$

$$= N \ln N - N - [N_\uparrow \ln N_\uparrow - N_\uparrow] - [N_\downarrow \ln N_\downarrow - N_\downarrow]$$
$$= N \ln N - N_\uparrow \ln N_\uparrow - N_\downarrow \ln N_\downarrow.$$

From (a) we have $N_\uparrow = \frac{N}{2} - \frac{U}{2\mu B}$ and $N_\downarrow = N - N_\uparrow = \frac{N}{2} + \frac{U}{2\mu B}$. Inserting these gives

$$S/k_B = \ln \Omega = N \ln N - \left(\frac{N}{2} - \frac{U}{2\mu B} \right) \ln \left(\frac{N}{2} - \frac{U}{2\mu B} \right) - \left(\frac{N}{2} + \frac{U}{2\mu B} \right) \ln \left(\frac{N}{2} + \frac{U}{2\mu B} \right).$$

(c) Compute the energy of the system as a function of its temperature T defined by $T^{-1} \equiv (\partial S/\partial U)_{N,B}$.
Hint: to simplify your work, recall the definition of the hyperbolic tangent, $\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}$.
**Solution.**

$$\frac{1}{k_B T} = \left(\frac{\partial \ln \Omega}{\partial U} \right)_{N,B}$$

$$= \frac{1}{2\mu B} \ln \left(\frac{N}{2} - \frac{U}{2\mu B} \right) - \left(\frac{N}{2} - \frac{U}{2\mu B} \right) \left(-\frac{1}{2\mu B} \right) \frac{1}{\frac{N}{2} - \frac{U}{2\mu B}}$$
$$- \frac{1}{2\mu B} \ln \left(\frac{N}{2} + \frac{U}{2\mu B} \right) - \left(\frac{N}{2} + \frac{U}{2\mu B} \right) \left(\frac{1}{2\mu B} \right) \frac{1}{\frac{N}{2} + \frac{U}{2\mu B}}$$

$$= \frac{1}{2\mu B} \ln \left(\frac{\frac{N}{2} - \frac{U}{2\mu B}}{\frac{N}{2} + \frac{U}{2\mu B}} \right)$$

$$= \frac{1}{2\mu B} \ln \left(\frac{1 - \frac{U}{N\mu B}}{1 + \frac{U}{N\mu B}} \right)$$

Now we solve this equation for U, starting by multiplying both sides by $2\mu B$ and taking the exponential of both sides.

$$\exp \left(\frac{2\mu B}{k_B T} \right) = \frac{1 - \frac{U}{N\mu B}}{1 + \frac{U}{N\mu B}}$$

$$\exp \left(\frac{2\mu B}{k_B T} \right) \left(1 + \frac{U}{N\mu B} \right) = 1 - \frac{U}{N\mu B}$$

$$\frac{U}{N\mu B} \left(\exp \left(\frac{2\mu B}{k_B T} \right) + 1 \right) = 1 - \exp \left(\frac{2\mu B}{k_B T} \right)$$

$$U = N\mu B \frac{1 - \exp \left(\frac{2\mu B}{k_B T} \right)}{1 + \exp \left(\frac{2\mu B}{k_B T} \right)}$$

$$U = N\mu B \frac{\exp \left(-\frac{\mu B}{k_B T} \right) - \exp \left(\frac{\mu B}{k_B T} \right)}{\exp \left(-\frac{\mu B}{k_B T} \right) + \exp \left(\frac{\mu B}{k_B T} \right)}$$

$$U = -N\mu B \tanh \left(\frac{\mu B}{k_B T} \right)$$

(d) Sketch a graph of $U(T)$. Conclude that at very low temperature, the vast majority of spins are aligned parallel to the applied magnetic field – this feature of the model is known as **paramagnetism**. It's also seen in more realistic models of spin systems and in real magnetic materials.

Solution. Assume $B > 0$, so the magnetic field points along $+\hat{z}$. (you may have already been assuming this, but everything we did before also applies to the case $B < 0$). Note that for positive temperatures we have $U < 0$, which should make sense—more spins want to be in the lower-energy state, which has negative energy per spin.

$$U = -N\mu B \tanh(\mu B/k_B T)$$
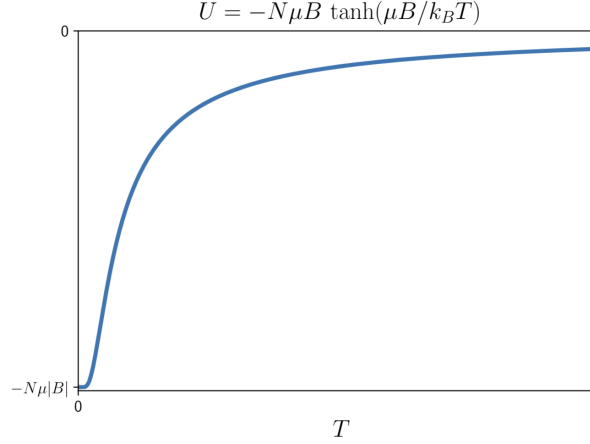
Note that $\tanh(x) \approx x$ when $x \ll 1$ and $\tanh(x) \approx 1$ when $x \gg 1$. If the temperature is very small, i.e. $\mu B/k_B T \gg 1$, then $U \approx -N\mu B$. That is, $N_\uparrow$ is very large and almost all of the spins are aligned with the magnetic field. This effect—where the spins in a material align with an applied magnetic field—is known as paramagnetism. On the other hand, if the temperature is large, then $U \approx -N\mu^2 B^2/k_B T$ is very close to zero, hence there are only slightly more spins pointing parallel to the magnetic field than anti-parallel. This means the paramagnetic effect is suppressed at high temperatures.

This system also supports negative temperatures, corresponding to positive U, in which case most of the spins are *down*. Using reasoning similar to problem 2(b), you can show that energy spontaneously flows from a negative temperature system to a positive temperature system. So in a sense, negative temperature is hotter than hot. ("What's cooler than being cool? Ice cold." - a conjecture famously made by André 3000.)

(e) A typical spin magnetic moment is about 10^{-4} electronvolts per tesla. Estimate the ratio $N_\uparrow/N$ in a field of 1 tesla at room temperature. *Hint: No calculator needed! Room temperature is about 300 kelvin, and room temperature times k_B is about $1/40 = 0.025$ electronvolts. These are useful to memorize. Also note $\tanh x \approx x$ when $x \ll 1$.*

Solution. From the answers to (a) and (c), we have

$$\mu B(N - 2N_\uparrow) = -N\mu B \tanh\left(\frac{\mu B}{k_B T}\right)$$

or

$$\frac{N_\uparrow}{N} = \frac{1 + \tanh(x)}{2},$$

where $x = \mu B/k_B T \approx (10^{-4} \cdot 1)/(1/40) = 4 \times 10^{-3} \ll 1$. Using $\tanh x \approx x$ gives $N_\uparrow/N \approx 0.5 + 0.004$. Even in relatively strong magnetic fields of one tesla, paramagnetism involves a pretty small fraction of the spins. You could increase this fraction by increasing x, which could be accomplished by cooling the system.