

Example—A Spin System

In the last lecture, we discussed the binomial distribution. Now, I would like to add a little physical content by considering a spin system. Actually this will be a model for a paramagnetic material. This system is also a trivial subcase of the Ising model.

We'll consider a large number, N , of identical spin $1/2$ systems. As you know, if you pick an axis, and measure the component of angular momentum of a spin $1/2$ system along that axis, you can get only two answers: $+\hbar/2$ and $-\hbar/2$. If there's charge involved, then there's a magnetic moment, \mathbf{m} , parallel or antiparallel to the angular momentum. If there's a magnetic field, \mathbf{B} , then this defines an axis and the energy $\mathbf{m} \cdot \mathbf{B}$ of the spin system in the magnetic field can be either $-mB$ if the magnetic moment is parallel to the field or $+mB$ if the magnetic moment is anti-parallel to the field. To save some writing, let $E = mB > 0$ so the energy of an individual system is $\pm E$.

In this model, we are considering only the energies of the magnetic dipoles in an external magnetic field. We are ignoring all other interactions and sources of energy. For example, we are ignoring magnetic interactions between the individual systems, which means we are dealing with a paramagnetic material, not a ferromagnetic material. Also, we are ignoring diamagnetic effects—effects caused by induced magnetic moments when the field is established. Generally, if there is a permanent dipole moment \mathbf{m} , paramagnetic effects dominate diamagnetic effects.

Of course, there must be some interactions of our magnets with each other or with the outside world or there would be no way for them to change their energies and come to equilibrium. What we're assuming is that these interactions are there, but just so small that we don't need to count them when we add up the energy. (Of course the smaller they are, the longer it will take for equilibrium to be established..)

Our goal here is to work out expressions for the energy, entropy, temperature, in terms of the number of parallel and antiparallel magnetic moments.

If there is no magnetic field, then there is nothing to pick out any direction, and we expect that any given magnetic moment or spin is equally likely to be parallel or antiparallel to any direction we pick. So the probability of parallel should be the same as the probability of antiparallel should be $1/2$: $p = 1 - p = q = 1/2$. If we turn on the magnetic field, we expect that more magnets will line up parallel to the field than antiparallel ($p > q$) so that the entire system has a lower total energy than it would have with equal numbers of magnets parallel and antiparallel.

If we didn't know anything about thermal effects, we'd say that *all* the magnets should align with the field in order to get the lowest total energy. But we do know something about thermal effects. What we know is that these magnets are exchanging energy with each other and the rest of the world, so a magnet that is parallel to the field, having energy

$-E$, might receive energy $+2E$ and align antiparallel to the field with energy $+E$. It will stay antiparallel until it can give up the energy $2E$ to a different magnet or to the outside world. The strengths of the interactions determine how rapidly equilibrium is approached (a subject we will skip for the time being), but the temperature sets an energy scale and determines how likely it is that chunks of energy of size $2E$ are available.

So suppose that n of the magnets are parallel to the field and $N - n$ are antiparallel. K&K define the “spin excess”, as the number parallel minus the number antiparallel, $2s = n - (N - n) = 2n - N$ or $n = s + N/2$. The energy of the entire system is then

$$U(n) = -nE + (N - n)E = -(2n - N)E = -2sE .$$

The entropy is the log of the number of ways our system can have this amount of energy and this is just the binomial coefficient.

$$\sigma(n) = \log \binom{N}{n} = \log \frac{N!}{(N/2 + s)!(N/2 - s)!} .$$

To put this in the context of our previous discussion of entropy and energy, note that we talked about determining the entropy as a function of energy, volume, and number of particles. In this case, the volume doesn't enter and we're not changing the number of particles (or systems) N . At the moment, we are not writing the entropy as an explicit function of the energy. Instead, the two equations above are parametric equations for the entropy and energy.

To find the temperature, we need $\partial\sigma/\partial U$. In our formulation, the entropy and energy are functions of a discrete variable, not a continuous variable. No problem! We'll just send one magnet from parallel to anti-parallel. This will make a change in energy, ΔU , and a change in entropy, $\Delta\sigma$ and we simply take the ratio as the approximation to the partial derivative. So,

$$\begin{aligned} \Delta U &= U(n - 1) - U(n) = 2E , \\ \Delta\sigma &= \sigma(n - 1) - \sigma(n) \\ &= \log \binom{N}{n - 1} - \log \binom{N}{n} \\ &= \log \left[\frac{N!}{(n - 1)!(N - n + 1)!} \frac{n!(N - n)!}{N!} \right] \\ &= \log \frac{n}{N - n + 1} \\ &= \log \frac{n}{N - n} \quad 1 \text{ can't matter if } N - n \sim N_0 \\ &= \log \frac{N/2 + s}{N/2 - s} , \end{aligned}$$

where the last line expresses the result in terms of the spin excess. Throwing away the 1 is OK, provided we are not at zero temperature where $n = N$.

The temperature is then

$$\tau = \frac{\Delta U}{\Delta \sigma} = \frac{2E}{\log(N/2 + s)/(N/2 - s)}.$$

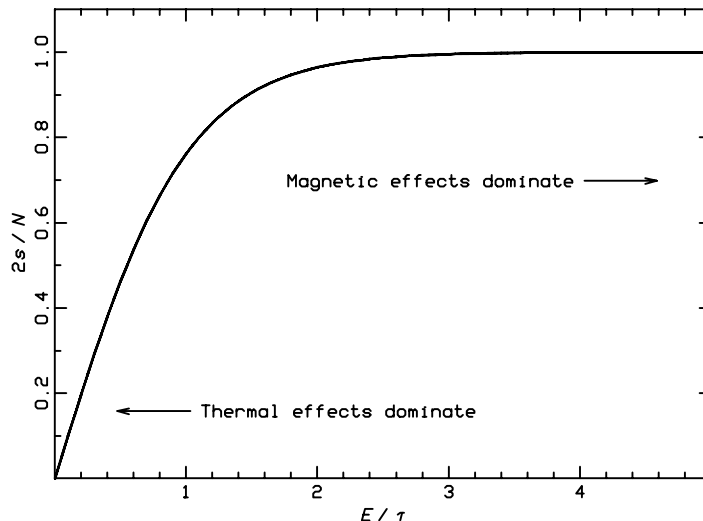
At this point it's convenient to solve for s . We have

$$\frac{N/2 + s}{N/2 - s} = e^{2E/\tau},$$

and with a little algebra

$$\frac{2s}{N} = \tanh \frac{E}{\tau}.$$

The plot shows this function—fractional spin excess versus E/τ . To the left, thermal



energy dominates magnetic energy and the net alignment is small. To the right, magnetic energy dominates thermal energy and the alignment is large. Just what we expected!

Suppose the situation is such that E/τ is large. Then the magnets are all aligned. Now turn off the magnetic field, leaving the magnets aligned. What happens? The system is no longer in equilibrium. It absorbs energy and entropy from its surroundings, cooling the surroundings. This technique is actually used in low temperature experiments. It's called *adiabatic demagnetization*. Demagnetization refers to removing the external magnetic field and adiabatic refers to doing it gently enough to leave the magnets aligned.

The Boltzmann Factor

An additional comment on probabilities: When the spin excess is $2s$, the probabilities of parallel or antiparallel alignment are:

$$p = \frac{1}{2} + \frac{s}{N}, \quad q = \frac{1}{2} - \frac{s}{N}.$$

The ratio of the probabilities is

$$\frac{q}{p} = \frac{1 - 2s/N}{1 + 2s/N} = e^{-2E/\tau}.$$

This is a general result. The relative probability that a system is in two states with an energy difference ΔE is just

$$\frac{\text{Probability of high energy state}}{\text{Probability of low energy state}} = e^{-\Delta E/\tau} = e^{-\Delta E/kT}.$$

This is called the Boltzmann factor. As we've already mentioned, this says that energies $\lesssim kT$ are "easy" to come by, while energies $> kT$ are hard to come by! The temperature sets the scale of the relevant energies.

The Gaussian Distribution

We've discussed two discrete probability distributions, the binomial distribution and (in the homework) the Poisson distribution. As an example of a continuous distribution, we'll consider the Gaussian (or normal) distribution. It is a function of one continuous variable and occurs throughout the sciences.

The reason the Gaussian distribution is so prevalent is that under very general conditions, the distribution of a random variable which is the sum of a large number of independent, identically distributed random variables, approaches the Gaussian distribution as the number of random variables in the sum goes to infinity. This result is called the central limit theorem and is proven in probability courses.

The distribution depends on two parameters, the mean, μ , (not the chemical potential!) and the standard deviation, σ (not the entropy!). The probability density is

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

You should be able to show that

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1,$$

$$\langle x \rangle = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \mu ,$$

$$\text{var}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} (x-\mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \sigma^2 .$$

Appendix A of K&K might be useful if you have trouble with these integrals. One can always recenter so that x is measured from μ and rescale so that x is measured in units of σ . Then the density takes the dimensionless form,

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} .$$

Sometimes you might need to integrate this density over a finite (rather than infinite) range. Two related functions are of interest, the error function

$$\text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt = 2 \frac{1}{\sqrt{2\pi}} \int_0^{\sqrt{2}z} e^{-x^2/2} dx ,$$

and the complementary error function

$$\text{erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-t^2} dt = 2 \frac{1}{\sqrt{2\pi}} \int_{\sqrt{2}z}^\infty e^{-x^2/2} dx ,$$

where the first expression (involving t) is the typical definition, and the second (obtained by changing variables $t = x/\sqrt{2}$) rewrites the definition in terms of the Gaussian probability density. Note that $\text{erf}(0) = 0$, $\text{erf}(\infty) = 1$, and $\text{erf}(z) + \text{erfc}(z) = 1$.

The Gaussian density is just the “bell” curve, peaked in the middle, with small tails. The error function gives the probability associated with a range in x at the middle of the curve, while the complementary error function gives probabilities associated with the tails of the distribution. In general, you have to look these up in tables, or have a fancy calculator that can generate them. As an example, you might hear someone at a research talk say, “I’ve obtained a marginal two-sigma result.” What this means is that the signal that was detected was only 2σ larger than no signal. A noise effect this large or larger will happen with probability

$$\frac{1}{\sqrt{2\pi}} \int_2^\infty e^{-x^2/2} dx = \frac{1}{2} \text{erfc} \frac{2}{\sqrt{2}} = 0.023 .$$

That is, more than 2 percent of the time, noise will give a 2σ result just by chance. This is why 2σ is marginal.

We’re straying a bit from thermal physics, so let’s get back on track. One of the reasons for bringing up a Gaussian distribution is that many other distributions approach

a Gaussian distribution when large numbers are involved. (The central limit theorem might have something to do with this!) For example, the binomial distribution. When the numbers are large, we can replace the discrete distribution in n with a continuous distribution. The advantage is that it is often easier to work with a continuous function.

In particular, the probability of a spin excess, s , is

$$p_s = \frac{N!}{(N/2 + s)!(N/2 - s)!} p^{N/2+s} q^{N/2-s}.$$

We need to do something with the factorials. In K&K, Appendix A, Stirling's approximation is derived. For large N ,

$$N! \sim \sqrt{2\pi N} N^N e^{-N}.$$

With this, we have

$$\begin{aligned} p_s &\sim \sqrt{\frac{2\pi N}{2\pi(N/2 + s)2\pi(N/2 - s)}} \frac{N^N}{(N/2 + s)^{(N/2+s)}(N/2 - s)^{(N/2-s)}} p^{N/2+s} q^{N/2-s} \\ &= \sqrt{\frac{1}{2\pi N(1/2 + s/N)(1/2 - s/N)}} \frac{p^{N/2+s} q^{N/2-s}}{(1/2 + s/N)^{(N/2+s)}(1/2 - s/N)^{(N/2-s)}} \\ &= \sqrt{\frac{1}{2\pi N(1/2 + s/N)(1/2 - s/N)}} \sqrt{\frac{pq}{pq}} \frac{p^{N/2+s} q^{N/2-s}}{(1/2 + s/N)^{(N/2+s)}(1/2 - s/N)^{(N/2-s)}} \\ &= \sqrt{\frac{1}{2\pi Npq}} \left(\frac{p}{1/2 + s/N}\right)^{(N/2+s+1/2)} \left(\frac{q}{1/2 - s/N}\right)^{(N/2-s+1/2)}. \end{aligned}$$

Recall that the variance of the binomial distribution is Npq , so things are starting to look promising. Also, we are working under the assumption that we are dealing with large numbers. This means that s cannot be close to $\pm N/2$. If it were, then we would have a small number of aligned, or a small number of anti-aligned magnets. So, in the exponents in the last line, $N/2 \pm s$ is a large number and we can ignore the $1/2$. Then

$$p_s = \sqrt{\frac{1}{2\pi Npq}} \left(\frac{p}{1/2 + s/N}\right)^{(N/2+s)} \left(\frac{q}{1/2 - s/N}\right)^{(N/2-s)}.$$

This is a sharply peaked function. We expect the peak to be centered at $s = s_0 = \langle s \rangle = \langle n \rangle - N/2 = Np - N/2 = N(p - 1/2)$. We want to expand this function about its maximum. Actually, it will be easier to locate the peak and expand the function, if we work with its logarithm.

$$\log p_s = A + \left(\frac{N}{2} + s\right) \left[\log p - \log \left(\frac{1}{2} + \frac{s}{N}\right) \right] + \left(\frac{N}{2} - s\right) \left[\log q - \log \left(\frac{1}{2} - \frac{s}{N}\right) \right],$$

where

$$A = \frac{1}{2} \log \left(\frac{1}{2\pi Npq} \right).$$

To locate the maximum of this function, we take the derivative and set it to 0

$$\frac{d \log p_s}{ds} = \log p - \log \left(\frac{1}{2} + \frac{s}{N} \right) - 1 - \log q + \log \left(\frac{1}{2} - \frac{s}{N} \right) + 1.$$

We note that this expression is 0 when $s/N = p - 1/2$, just as we expected. So this is the point about which we'll expand the logarithm. The next term in a Taylor expansion requires the second derivative

$$\begin{aligned} \frac{d^2 \log p_s}{ds^2} &= -\frac{1}{N/2 + s} - \frac{1}{N/2 - s} \\ &= -\frac{1}{Np} - \frac{1}{Nq} = -\frac{1}{Npq}, \end{aligned}$$

where, in the last line, we substituted the value of s at the maximum. We can expand the logarithm as

$$\log p_s = A - \frac{1}{2} \frac{1}{Npq} (s - s_0)^2 + \dots$$

where $s_0 = N(p - 1/2)$ is the value of s at the maximum. Finally, we let $\sigma^2 = Npq$, exponentiate the logarithm, and obtain,

$$p(s) \sim \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(s - s_0)^2/2\sigma^2},$$

where the notation has been changed to indicate a continuous variable rather than a discrete variable. You might worry about this last step. In particular, we have a discrete probability that we just converted into a probability density. In fact, $p(s) ds$ is the probability that the variable is in the range $s \rightarrow s + ds$. In the discrete case, the spacing between values of s is unity, so we require,

$$p(s)((s + 1) - s) = p_s,$$

which leads to $p(s) = p_s$. Had there been a different spacing there would be a different factor relating the discrete and continuous expressions.

All this was a lot of work to demonstrate in some detail that for large N (and not too large s), the binomial distribution describing our paramagnetic system goes over to the Gaussian distribution. Of course, expanding the logarithm to second order guarantees a Gaussian!

In practice, you would not go to all this trouble to do the conversion. The way you would actually do the conversion is to notice that large numbers are involved, so the

distribution must be Gaussian. Then all you need to know are the mean and variance which you calculate from the binomial distribution or however you can. Then you just write down the Gaussian distribution with the correct mean and variance.

Returning to our paramagnetic system, we found earlier that the mean value of the spin excess is

$$s_0 = \frac{N}{2} \tanh \frac{E}{\tau} .$$

We can use the Gaussian approximation provided s is not too large compared to $N/2$ which means $E \lesssim \tau$. In this case, a little algebra shows that the variance is

$$\sigma^2 = Npq = N \left(\frac{1}{2} \operatorname{sech} \frac{E}{\tau} \right)^2 .$$

For given E/τ , the actual s fluctuates about the mean s_0 with a spread proportional to \sqrt{N} and a fractional spread proportional to $1/\sqrt{N}$. A typical system has $N \sim N_0$, so the fractional spread is of order 10^{-12} and the actual s is always very close to s_0 .

While we're at it, it's also interesting to apply Stirling's approximation to calculate the entropy of our paramagnetic system. Recalling Stirling's approximation for large N ,

$$N! \sim \sqrt{2\pi N} N^N e^{-N} .$$

Taking the logarithm, we have

$$\log N! \sim \frac{1}{2} \log 2\pi + \frac{1}{2} \log N + N \log N - N .$$

The first two terms can be ignored in comparison with the last two, so

$$\log N! \sim N \log N - N .$$

Suppose our spin system has $s_0 \approx 0$. Then the entropy is

$$\begin{aligned} \sigma &\approx \log \frac{N!}{(N/2)!(N/2)!} \\ &\sim N \log N - N - 2((N/2) \log(N/2) - (N/2)) \\ &= N \log N - N \log(N/2) \\ &= N \log 2 \\ &= 4.2 \times 10^{23} \quad (\text{fundamental units}) \\ &= 5.8 \times 10^7 \text{ erg K}^{-1} \quad (\text{conventional units}), \end{aligned}$$

where the last two lines assume one mole of magnets.