Reading assignment

Schroeder, section 2.4.

Recap of lecture 6

- Constituent states, microstates, and macrostates.
- Goal is calculation of probabilities of macrostates.
- If microstates are equally probable (usually assumed), the probability of a macrostate is proportional to the number Ω of microstates belonging to that macrostate. Normalization by dividing by the total number of microstates gives the probability.
- Toy problem—two-state paramagnet:

$$\Omega(N_\uparrow) = \binom{N}{N_\uparrow} = \frac{N!}{N_\uparrow! N_\downarrow!} \,.$$

- Toy problem—Einstein model for lattice vibrations:

$$\Omega(N, q) = \frac{(N-1+q)!}{q!\,(N-1)!} = \binom{N-1+q}{q} \,.$$

Homework

HW Problem

Schroeder problem 2.7, p. 55.

Our current view

We've seen that macrostates of an isolated system having more microstates (higher multiplicity in Schroeder's lingo) are statistically more likely to occur than those having fewer. As time progresses, transitions occur among accessible microstates (those satisfying the constraints on the system, such as total energy), so we expect measurements to observe the most probable macrostates.

But it's a bit difficult to see just how to apply this idea to an isolated system. For example, for an isolated system of harmonic oscillators with fixed total energy, if we take the macrostates to be characterized by the total energy, there is only one state, thus no comparisons with other macrostates can be made. There is nothing intrinsic to this picture that would preclude having all of the energy or all of the particles (say in a gas) concentrated in one small portion of the available volume.

Refining the picture

We could, however, compare regions in the system, to see whether they have comparable amounts of energy. Intuitively, we expect more even distributions to be more probable than ones with large place-to-place variations. This is one reason why we want to split our system into two (or more) parts and consider interactions between them.

In addition, one often encounters problems in both thermodynamics and statistical mechanics in which one needs to predict the equilibrium states of interacting systems. The simplest case (and the one to which a statistical treatment is most directly relevant) is that of two systems weakly coupled so that energy can flow between them (thermal contact), but volume and particle exchanges are precluded.

Two paramagnets in thermal contact

We'll address the statistical aspects of the process of equilibration of a pair of systems in thermal contact. Suppose we have a pair of two-state paramagnets in thermal contact, with the combined system isolated so that its total energy in an external magnetic field is constant. In this case, we'll characterize macrostates by the division of energy between the two subsystems, rather than by the combined total energy.

Let N_1 and N_2 be the numbers of magnetic moments in each subsystem. We'll define the zero of energy to be the ground state, with all magnetic moments aligned (up) with the external field. Then the energy is proportional to the number of antialigned (down) moments, let's call that n_i, i indexing the subsystem, with $n_1 + n_2 = n$ fixed by the total-energy constraint.

Two paramagnets in thermal contact

We might initialize the system with $n_1 = n$ and $n_2 = 0$ and ask what will be the equilibrium macrostate (set of values of n_1 and n_2) after a long time. If every microstate (characterized by the full distribution of down moments) is assumed equally likely, we can determine which values of n_1 and n_2 are most probable by simply finding the number of microstates corresponding to each.

We know that for each paramagnet, the number of microstates corresponding to the macrostate characterized by n_i is

$$\Omega(N_i, n_i) = \binom{N_i}{n_i} = \frac{N_i!}{n_i!(N_i - n_i)!} .$$

For each microstate of the first paramagnet, all the microstates of the second paramagnet with the same $n_2 = n - n_1$ are accessible, so the number of microstates corresponding to the macrostate n of the full system is just the product of the numbers $\Omega(N_i, n_i)$ of the subsystems.

Example

As an example, consider a pair of puny paramagnets with $N_1 = 5$ and $N_2 = 10$ moments, respectively, and suppose the energy is fixed by having a total of $n = 5$ antialigned moments. The statistics are summarized as follows:

n_1	$\Omega(5, n_1)$	n_2	$\Omega(10, n_2)$	$\Omega(5, n_1)\Omega(10, n_2)$	Probability
0	1	5	252	252	0.08392
1	5	4	210	1050	0.34965
2	10	3	120	1200	0.39960
3	10	2	45	450	0.14985
4	5	1	10	50	0.01665
5	1	0	1	1	0.00033

Clearly, the more even distributions (weighted by subsystem size) of the antialigned moments are statistically favored. Note that some of the states of paramagnet 2 are not accessible.

Example

So, if we started with all of the energy in paramagnet 1, a macrostate with only a single corresponding microstate, we can expaect that after sufficient time, almost no measurements of the moment distribution will find the combined system back in that state. Energy will "flow" from paramagnet 1 to paramagnet 2, simply because naturally occurring transitions among the microstates of the combined system will lead with high probability to some other state with a more even distribution of antialigned moments. No work has been done, and no particles have been exchanged between the subsystems.

This is the key concept in understanding equilibration through heat flow.

Homework

HW Problem

Schroeder problem 2.8, p. 59.

The macroscopic trend

A graph of the probabilities or the number of accessible microstates as a function of the parameter characterizing the macrostate (say n_1) has a peak around the most probable value of the parameter.

That peak becomes narrower as the number of constituents increases, becoming incredibly sharp for systems with macroscopic numbers of particles.

The consequence of that is that macrostates other than the most probable one will *essentially never* be observed.

Factorials are annoying

Many of the combinatorial expressions we've encountered contain multiple factorials, and these are often difficult to deal with analytically. There is a much simpler, but approximate, expression for the factorial, called Stirling's approximation (or Stirling's formula), that can simplify expressions substantially.

There are a few, slightly different, formulations, all of which rely on large values of the argument. That happens to be ideal for our purposes, since our factorials for macroscopic systems have arguments of magnitude typified by Avogadro's number.

Gamma functions and factorials

The derivation makes use of the Gamma function to represent the factorial function, so let's start by recalling the connection between them.
We'll define the Gamma function by

$$\Gamma(x+1) = \int_0^\infty t^x e^{-t}\, dt\,.$$

To see how it relates to the factorial, we'll begin by integrating by parts once. Let

$$\begin{aligned} u &= t^x & dv &= e^{-t}\, dt \\ du &= xt^{x-1} & v &= -e^{-t}\,. \end{aligned}$$

Gamma functions and factorials

Then

$$\Gamma(x+1) = \underbrace{-t^x e^{-t}\big|_0^\infty}_{0} + x \underbrace{\int_0^\infty t^{x-1} e^{-t}\, dt}_{\Gamma(x)} = x\Gamma(x)\,.$$

This recursion relation allows us to write the Gamma function of an integer as the product

$$\Gamma(n+1) = n(n-1)(n-2)\cdots 1\Gamma(1)\,.$$

But $\Gamma(1)$ can be evaluated easily:

$$\Gamma(1) = \int_0^\infty e^{-t}\, dt = 1\,.$$

Thus, the Gamma function and the factorial function are closely related:

$$\Gamma(n+1) = n!$$

Toward Stirling's approximation

We'll use the integral form of the Gamma function to derive Stirling's approximation for the factorial. Begin by rewriting the integrand in a form that proves to be more convenient:

$$\begin{aligned} N! &= \int_0^\infty t^N e^{-t}\, dt \\ &= \int_0^\infty e^{N \ln t - t} dt\,. \end{aligned}$$

Next we'll use a somewhat magic-looking substitution that we'll find does useful things for us. Let

$$t = N + y\sqrt{N}\,.$$

Then

$$dt = \sqrt{N}\, dy\,, \quad t = 0 \Leftrightarrow y = -\sqrt{N}\,, \quad \text{and} \quad t = \infty \Leftrightarrow y = \infty\,.$$

Toward Stirling's approximation

This changes the expression for $N!$ to the much messier looking

$$N! = \sqrt{N} \int_{-\sqrt{N}}^{\infty} e^{N \ln(N+y\sqrt{N}) - N - y\sqrt{N}} dy \,.$$

This form is still exact, but this is the point where we need to begin thinking about ways to simplify the mess through approximation. The logarithm in the exponent is a prime candidate, since there's a well-known expansion for $\ln(1+x)$. To get something like that, we'll factor out the N in the argument of the logarithm:

$$\ln(N + y\sqrt{N}) = \ln\left[N\left(1 + \frac{y}{\sqrt{N}}\right)\right] = \ln N + \ln\left(1 + \frac{y}{\sqrt{N}}\right) \,.$$

For large N, the fraction $y/\sqrt{N}$ is small compared to one, so that the expansion of the logarithm proves useful.

Toward Stirling's approximation

The expansion, expressed in terms of ϵ to emphasize its utility for small values of ϵ is

$$\ln(1+\epsilon) = \epsilon - \frac{\epsilon}{2} + \frac{\epsilon^2}{3} - \ldots .$$

We'll truncate this to two terms, in which case the factorial becomes

$$\begin{aligned} N! &\approx \sqrt{N} \int_{-\sqrt{N}}^{\infty} e^{N \ln N + y\sqrt{N} - \frac{y^2}{2} - N - y\sqrt{N}} dy \\ &= \sqrt{N} e^{N \ln N - N} \int_{-\sqrt{N}}^{\infty} e^{-\frac{y^2}{2}} dy \,. \end{aligned}$$

Now it's starting to look much simpler! The integral could be evaluated easily if the lower limit were 0 or ∞, but it's the rather inconvenient-looking $-\sqrt{N}$. However, the integrand, a Gaussian function, falls off quite rapidly as y deviates from 0.

Stirling's approximation at last

This means we can replace the lower limit of the integral by $-\infty$ with little damage to the result when N is large. Then the integral becomes the well-known

$$\int_{-\infty}^{\infty} e^{-y^2/2} dy = \sqrt{2\pi}\,,$$

and the approximation to the factorial becomes

$$N! \sim \sqrt{2\pi N} e^{N \ln N - N} = \sqrt{2\pi N} \left(\frac{N}{e}\right)^N \qquad \text{(for large } N\text{)},$$

which is Stirling's approximation to $N!$.

Taking the logarithm

It turns out that we'll usually be more interested in the logarithm of $N!$, so let's see what Stirling's approximation gives for that:

$$\begin{aligned}
\ln N! &\approx \underbrace{\frac{1}{2}\ln 2\pi}_{\approx 0.92} + \frac{1}{2}\ln N + N\ln N - N \\
&\approx \left(N + \frac{1}{2}\right)\ln N - N \\
&\approx N\ln N - N\,,
\end{aligned}$$

where we've neglected quantities of order 1 compared to N. This very simple approximation, valid for large N, is also often called Stirling's approximation, and it's the form we'll use most frequently.

Looking back at what we've done

It's interesting to note a simple connection between this result and our starting point. We started with the exact expression

$$N! = \int_0^\infty e^{N \ln t - t} dt \,.$$

Now the function comprising the integrand, let's call it $f(t)$, goes to $-\infty$ as $t \to 0$ and as $t \to \infty$, and it has a single maximum in between. To find it, just differentiate:

$$\frac{df}{dt} = \frac{N}{t} - 1 = 0 \,,$$

which locates the maximum at

$$t = N \,.$$

Looking back at what we've done

The value at that maximum is then

$$f(N) = N \ln N - N \,,$$

which is the same as the approximation we found for $\ln N!$. So that approximation is equivalent to replacing the value of the integral by the value of the integrand at its maximum:

$$N! = \int_0^\infty e^{N \ln t - t} dt \sim e^{N \ln N - N} \quad \text{(for large N.)}$$

Homework

HW Problem

Schroeder problem 2.16, p. 63.