

Reading

This week, you should read the first two chapters of K&K.

Entropy and the Number of States

As we discussed last time, in the statistical view, entropy is related to the number of “microstates” of a system. In particular, the entropy is the log of the number of states that are accessible to the system when it has specified macroscopic parameters (its “macrostate”).

The fact that entropy always increases is just a reflection of the fact that a system adjusts its macroscopic parameters, within the allowed constraints, so as to maximize the number of accessible states and hence the entropy.

So, a large part of statistical mechanics has to do with counting states and another large part has to do with deriving interesting results from these simple ideas.

Why is the Number of States Maximized?

Good question. We are going to take this as an axiom or postulate. We will not attempt to prove it. However, we can give some plausibility arguments.

First, remember that we are typically dealing with something like Avogadro’s number of particles, $N_0 = 6.02 \times 10^{23}$. As we discussed last time, this makes the probability distributions very sharp. Or put another way, improbable events are very improbable.

The other thing that happens with a large number of particles has to do with the randomness of the interactions. Molecules in a gas are in continual motion and collide with each other (we will see later in the term, how often). During these collisions, molecules exchange energy, momentum, angular momentum, etc. The situation in a liquid is similar, one of the differences between a liquid and gas has to do with the distance a molecule travels between collisions: in a gas, a molecule typically travels many molecular diameters; in a liquid, the distance between collisions is of the order of a molecular diameter. In a solid, molecules tend to be confined to specific locations, but they oscillate around these locations and exchange energy, momentum, etc. with their neighbors.

OK, molecules are undergoing collisions and interactions all the time. As a result, the distribution of molecular positions and speeds is randomized. If you pick a molecule and ask things like where is it located, how fast is it going, etc., the answers can only be given in terms of probabilities and these answers will be the same no matter which molecule you

pick. (Provided you pick the same kind of molecule - you'll probably get different answers for a N_2 molecule and an Ar atom, but you'll get the same answers for two N_2 molecules.)

Sticky point: suppose we assume that the world is described by classical mechanics. Also suppose we know the interactions between molecules in some isolated system. Suppose we also know all $\sim N_0$ positions \mathbf{r}_i and momenta \mathbf{p}_i (and whatever else we might need to know to specify the system, perhaps the angular momenta of the molecules, etc.). Then in principle, the equations of motion can be solved and the solution tells us the exact state of the system for all future times. That is, there is nothing random about it! How do we reconcile this with the probabilistic view espoused in the preceding paragraphs?

So far as I know, there are reasonable practical answers to this question, but there are no good philosophical answers. The practical answers have to do with the fact that one can't really write down and solve the equations of motion for $\sim N_0$ particles. But we can in principle! A somewhat better answer is that we can only know the initial conditions with some precision, not infinite precision. As we evolve the equations of motion forward, the initial uncertainties grow and eventually dominate the evolution. This is one of the basic concepts of chaos which has received a lot of attention in recent years: small changes in the initial conditions can lead to large changes in the final result. (Have you ever wished you could get a 10 day or 30 day weather forecast? Why do they stop with the 5 day forecast?)

Of course, the fact that we cannot measure infinitely precisely the initial conditions nor solve such a large number of equations does not mean (still assuming classical mechanics) that it couldn't be done in principle. (This is the philosophical side coming again!) So perhaps there is still nothing random going on. At this point one might notice that it's impossible to make a totally isolated system, so one expects (small) random perturbations from outside the system. These will disturb the evolution of the system and have essentially the same effect as uncertainties in the initial conditions. But, perhaps one just needs to include a larger system!

If we recognize that quantum mechanics is required, then we notice that quantum mechanics is an inherently probabilistic theory. Also, I'm sure you've seen or will see in your QM course that in general, uncertainties tend to grow with time (the spreading out of a wave packet is a typical example). On the other hand, the system must be described by a wave function (depending on $\sim N_0$ variables), whose evolution is determined by Schroedinger's equation

As you can see this kind of discussion can go on forever.

So, as said before, we are going to postulate that a system is equally likely to be in any state that is consistent with the constraints (macroscopic parameters) applied to the system.

As it happens, there is a recent *Physics Today* article on exactly this subject: trying to go from the reversibility of classical mechanics to the irreversibility of statistical mechanics. It's by G. M. Zaslavsky and is called, "Chaotic Dynamics and the Origin of Statistical Laws," 1999, vol. 52, no. 8, pt. 1, p. 39. I think you can read this article and get a feel for the problem even if some of it goes over your head (as some of it goes over my head).

Aside—Entropy and Information

In recent times, there has been considerable interest in the information content of data streams and what manipulating (computing with) those data streams does to the information content. It is found that concepts in information theory are very similar to concepts in thermodynamics. One way out of the "in principle" problems associated with classical entropy is to consider two sources of entropy: a physical entropy and an information or algorithmic entropy. This goes something like the following: if we had some gas and we had the knowledge of each molecule's position and momentum, then the physical entropy would be zero (there's nothing random about the positions and momenta), but the algorithmic entropy of our list of positions and momenta would be large (and equal to the physical entropy of a similar gas whose positions and momenta we hadn't determined). What is algorithmic entropy? Essentially, the logarithm of the number of steps in the algorithm required to reproduce the list.

In 1998, Toby Marriage wrote a JP on this topic. You can find it at

http://physics.princeton.edu/www/jh/juniors_fall98.html .

One of our criteria for junior papers is that other juniors should be able to understand the paper; so I think you might get something out of this paper as well!

Macroscopic Parameters

We will be most concerned with systems in equilibrium. Such a system can usually be described by a small number of macroscopic parameters. For example, consider a gas. If the density of the gas is low enough, it can be described quite well by the ideal gas law when it's in equilibrium:

$$pV = NkT = nRT ,$$

where p is the pressure, V is the volume, N is the number of molecules, n is the number of moles, $k = 1.38 \times 10^{-16}$ erg K^{-1} is Boltzmann's constant or the gas constant per molecule, $R = 8.31 \times 10^7$ erg mole $^{-1}$ $\text{K}^{-1} = N_0 k$ is the gas constant per mole, and T is the absolute temperature.

Notice that some parameters depend on how much gas we have and some don't. For example, if we replicate our original system, so we have twice as much, then V , N , U (internal energy), and S (entropy) all double; p , and T stay the same. We are ignoring the contribution of any surface interactions which we expect to be very small. Can you think why? Parameters which depend on the size of the system are called *extensive* parameters. Parameters that are independent of the size of the system are called *intensive* parameters.

Note that the gas law is not the whole story. If more than one kind of molecule is in the gas, we need to specify the numbers of each kind: N_1, N_2, \dots . Also, the gas law does not say anything about the energy of the gas or its entropy. The gas law is an equation of state, but it needs to be supplemented by other relations in order that we know everything there is to know about the gas (macroscopically, that is!). For systems more complicated than a gas, other parameters may be needed.

Another thing to notice is that not all parameters may be specified independently. For example, having specified N , T , and V , the pressure is determined. Thus there is a certain minimum number of parameters which specify the system. Any property of the system must be a function of these parameters. Furthermore, we can often change variables and use a different set of parameters. For a single component ideal gas, we might have

$$p = p(N, V, T), \quad U = U(N, V, T), \quad S = S(N, V, T) .$$

We might imagine solving for T in terms of N , V , and U , and we can write

$$p = p(N, V, U), \quad T = T(N, V, U), \quad S = S(N, V, U) .$$

Anything that depends only on the equilibrium state of the system can be expressed as a function of the parameters chosen. Which parameters are to be used depends on the particular situation under discussion. For example, if the volume of a system is under our control, we would likely use that as one of the independent parameters. On the other hand, many processes occur at constant pressure (with the volume adjusting to what it needs to

be). In this case, using p rather than V as the independent parameter will probably be more convenient.

The Temperature

As we remarked, the entropy is the logarithm of the number of microstates accessible to a system. The number of states must be a function of the same macroscopic parameters that determine the macrostate of the system. Let's consider a system described by its internal energy U , its volume V , and the number of each kind of constituent particle N_a , N_b , \dots . For the moment, we ignore the possibility of reactions which can change particles of one kind into another kind. This means that our expressions will have the same form for N_a , N_b , etc., so we'll just assume a single kind of particle for the time being and assume we have N of them. Then the number of microstates is

$$g = g(U, V, N).$$

If we have two systems, that we prevent from interacting, then the number of microstates of the combined system is

$$g(U, V, N, U_1, V_1, N_1) = g_1(U_1, V_1, N_1)g_2(U_2, V_2, N_2),$$

with

$$U = U_1 + U_2, \quad V = V_1 + V_2, \quad N = N_1 + N_2.$$

This is straightforward. Any microstate in system 1 can be paired with any microstate in system 2, so the total number of microstates is just the product of the number for each system. Also, we have specified the macrostate in terms of extensive parameters, so we can write the parameters of the combined system as the sum of those for the individual systems as well as one set of the individual system parameters.

Following K&K, the *dimensionless* entropy is just

$$\begin{aligned} \sigma(U, V, N, U_1, V_1, N_1) &= \log g(U, V, N, U_1, V_1, N_1) = \log g_1 g_2 \\ &= \log g_1 + \log g_2 = \sigma_1(U_1, V_1, N_1) + \sigma_2(U_2, V_2, N_2). \end{aligned}$$

So far, we haven't really done anything. We've just written down some definitions twice. We have prevented the two systems from interacting, so nothing exciting can happen. Now let's suppose we allow the systems to exchange energy. In other words, we allow U_1 and U_2 to vary, but any change in U_1 has a compensating change in U_2 so that U is constant. In addition, we prevent changes in volume and numbers of particles, so that V_1 , V_2 , N_1 , and N_2 remain constant.

We're placing the systems in thermal contact, but preventing changes in volume or particle number. We know what will happen: energy flows from the hotter system to the cooler system until they come to thermal equilibrium at the same temperature. We know this from our intuitive understanding of the second law: heat flows from a hot object to a cold object.

But, what about our postulate that a system maximizes the number of accessible microstates? In this case, it means that the system adjusts U_1 and U_2 to maximize the entropy. So,

$$\begin{aligned} \left(\frac{\partial \sigma}{\partial U_1} \right)_{V,N} &= 0 && \text{since } \sigma \text{ is maximized} \\ &= \left(\frac{\partial \sigma_1}{\partial U_1} \right)_{V,N} + \left(\frac{\partial \sigma_2}{\partial U_1} \right)_{V,N} \\ &= \left(\frac{\partial \sigma_1}{\partial U_1} \right)_{V,N} + \left(\frac{\partial \sigma_2}{\partial U_2} \right)_{V,N} \frac{\partial U_2}{\partial U_1} \\ &= \left(\frac{\partial \sigma_1}{\partial U_1} \right)_{V,N} - \left(\frac{\partial \sigma_2}{\partial U_2} \right)_{V,N} && \text{since } \Delta U_1 = -\Delta U_2. \end{aligned}$$

This means

$$\left(\frac{\partial \sigma_1}{\partial U_1} \right)_{V,N} = \left(\frac{\partial \sigma_2}{\partial U_2} \right)_{V,N},$$

after equilibrium has been established.

So at equilibrium, the rate of change of entropy with respect to energy is the same for the two systems. If we started out with the two systems and we allowed them to exchange energy and nothing happened, then we know that $\partial\sigma/\partial U$ was already the same. If system 1 and system 2 are in equilibrium with respect to energy exchange and we allow system 1 to exchange energy with a third system and nothing happens, then $\partial\sigma_3/\partial U_3$ must also have the same value and nothing will happen if systems 2 and 3 are allowed to exchange energy. Thus, $\partial\sigma/\partial U$ has properties very similar to those we ascribe to temperature. In fact, we can define the temperature as:

$$\frac{1}{\tau} = \left(\frac{\partial \sigma}{\partial U} \right)_{V,N}.$$

This makes τ an intensive quantity (it's the ratio of two extensive quantities), and it makes the energy flow in the "correct" direction.

This can be seen as follows: if the two systems are not in equilibrium when we allow energy to flow, then the entropy of the combined systems must increase: The increase in

entropy after a very small amount of energy has been transferred is

$$\begin{aligned}\delta\sigma &> 0 \\ &= \delta\sigma_1 + \delta\sigma_2 \\ &= \frac{1}{\tau_1}\delta U_1 + \frac{1}{\tau_2}\delta U_2 \\ &= \left(\frac{1}{\tau_1} - \frac{1}{\tau_2}\right)\delta U_1.\end{aligned}$$

So if $\tau_1 < \tau_2$, $\delta U_1 > 0$, which means energy flows from the high τ system to the low τ system.

Finally, if you remember your elementary thermodynamics, recall that $dU = T dS - p dV$ which agrees with this definition of temperature.

Units: from our definitions σ is dimensionless and τ has the dimensions of energy. You recall that temperature T has the dimensions of Kelvins and entropy S has the dimensions of ergs per Kelvin. As it turns out,

$$\begin{aligned}S &= k\sigma, \\ \tau &= kT.\end{aligned}$$

Boltzmann's constant is really just a scale factor which converts conventional units to the fundamental units we've defined above.

It's often said that we measure temperature in Kelvins or degrees Celsius or Fahrenheit because the measurement of temperature was established before the development of thermodynamics which in turn took place before the connection to energy was fully appreciated. What would you think if you tuned in to the weather channel and found out that the high tomorrow was expected to be 4.14×10^{-14} erg or 0.0259 eV??? (If I did the arithmetic correctly, this is $\sim 80^\circ\text{F}$.)

Actually, to measure a temperature, we need a thermometer. Thermometers make use of physical properties which vary with temperature. (That's obvious I suppose!) The trick is to calibrate the thermometers so you get an accurate measure of the thermodynamic temperature, τ/k . A recent *Physics Today* article discusses some of the difficulties in defining a good practical scale for $\tau/k < 1$ Kelvin. (Soulen, Jr., R. J., and Fogle, W. E., 1997, *Physics Today*, vol. 50, no. 8, p. 36, "Temperature Scales Below 1 Kelvin.")

One other thing to point out here: You've no doubt noticed the V, N subscripts. When you read a thermodynamics text, you'll often find the statement that this a reminder that V and N are being held fixed in taking the indicated partial derivative. Well, this is true, but since we have a partial derivative, which already means hold everything else fixed, why do we need an extra reminder? Answer: since there are so many choices of independent variables, these subscripts are really a reminder of the set of independent variables in use.

Note that we can add energy to a gas keeping the volume and number of particles fixed. In this case the pressure and temperature rise. Alternatively, we can keep the pressure and number of particles fixed. In this case the volume and temperature increase. Furthermore,

$$\left(\frac{\partial\sigma}{\partial U}\right)_{V,N} \neq \left(\frac{\partial\sigma}{\partial U}\right)_{p,N}.$$

When it's obvious from the context which set of independent variables are in use, I will probably be lazy and omit the subscripts.