

The following content is provided under a Creative Commons license. Your support will help MIT OpenCourseWare continue to offer high quality educational resources for free. To make a donation or view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare at ocw.mit.edu.

PROFESSOR: So, last time we started in on a discussion of a new topic, which was statistical mechanics. So what we're trying to do now is revisit the thermodynamics that we've spent most of the term deriving and trying to use in a microscopic approach. So our hope is to be able to start with a microscopic model of matter, starting with atoms and molecules that we know are out there. And formulate thermodynamics starting from that microscopic point of view. In contrast to the way it was first formulated historically, and the way we've presented it also, which is as an entirely empirical subject based on macroscopic observation and deduction from that. So, we looked at the probabilities that states with different energies would be occupied and we inferred that there would be a simple way to describe the distribution of atoms or molecules in states of different levels.

First of all, although I didn't state it explicitly, essentially assumed there is that if I've got a bunch of molecules, and I've got states that they could be in of equal energy, then the probability that they would be in one or another of those states is the same. If the energies of the states are equal, the probabilities that those states will be occupied. That they'll be populated by molecules, will be equal. And then we deduced what's called the Boltzmann probability distribution. Which says that the probability that a molecular state, i , will be occupied is proportional to $e^{-E_i/kT}$, where E_i is the energy of that molecular state. And then realizing that the probability that some state out there has to be occupied, we saw that of course if we sum over all these probabilities that sum has to equal one, right? In other words, the sum over all of i of P_i is equal to one. So that means that we could write not just that this is proportional to this, but we could write that P_i is equal to $e^{-E_i/kT}$ divided by the sum of all such terms.

So now we know how to figure out how likely it is that a certain state is occupied. And what it means is, let's say we have a whole bunch of states whose energy is pretty low compared to kT . kT , the Boltzmann constant times temperature is energy units. So if it's pretty warm, maybe it's room temperature, maybe it's warmer. And there are a whole lot of states that are accessible because the energies are less than kT . What does that mean? That means that you could put a bunch of different i 's here whose energies are all quite low compared to kT . And they'll all have significant probabilities.

Now let's decrease the energy. Go to cold temperature. So kT becomes really small. So that $e^{-E_i/kT}$ starts to, if this gets to be bigger than kT by a lot, then this whole thing is a very small number. Suddenly, you get very few states accessible. So if we just plot this distribution, of course it's easy to do, it's just a

decaying exponential. So P_i , which is a function of E_i . It just looks like this. And here are a bunch of states. And if we have a classical mechanics picture of matter, then there would just be continuous states. And if we have a quantum mechanical picture of matter, there might be individual states with gaps in between the energies. Either way, what happens, what this is saying is, if we go out to higher and higher energies, then you have a smaller and smaller probability to be in a state like that.

And if you go to low energies, the probability gets bigger. And it depends on temperature. Because it's the ratio between the energy and kT that dictates the size of that term. So let's say this is moderate temperature. If we go to low temperature, it might look more like this. Hardly any states can be occupied because even states of rather moderate energies, suddenly now those energies are much bigger than kT . In other words, kT is measuring thermal energy. It's saying there's not enough thermal energy to populate, to knock things into states whose energy is much higher than that. So you have a very precipitous decay.

If you go to the other extreme, of very high temperature, then this will tend to flatten out. Eventually it'll decay, but it may take a long time. Because now kT is enormous. So the energy has to get to be very big. Before it's bigger than kT . And as long as it isn't, then this exponential term is not small. So there be a whole bunch of states that may be occupied. In other words, a whole bunch of states that are thermally accessible at equilibrium. Systems at thermal equilibrium, molecules are getting knocked around with whatever thermal energy is available. Crashing into each other or into the walls of a vessel. And there'll be a distribution of molecular energies. And that distribution is skewed either low or high depending on the temperature. So that's what that distribution, the Boltzmann distribution, is telling us. This is often called a Boltzmann factor, because it's telling what the population of some particular state is. OK.

Now, this is just dealing with individual molecule states. Then we said, OK, what about the whole system? Well, the same kind of relation holds there, too. In other words, if these are individual molecule energies, now if I look at the entire system, right well, still, there's no reason that same distribution doesn't hold. And it does. So in other words, P_i of E_i , that's the whole system energy, is $e^{-E_i/kT}$, over the sum over i , $e^{-E_i/kT}$. Now, this i doesn't refer to a single molecule state. We're talking about a whole system. It might be a mole of atoms or molecules in the gas phase, or what have you. It refers to a system state where the energy, the state of every one of those atoms or molecules is specified.

So this is a system microstate. Every molecular state is specified. So for a mole of stuff, that means there might be 10^{24} or so molecular states that this single subscript is indicating. But the point is, those states exist. They have a total energy. What's the probability that the whole system will be in such a state? Still going to be proportional to the total system energy.

It turns out that these summations end up taking on an enormous importance in statistical mechanics. And the reason is, as we'll see shortly, it turns out that every single macroscopic thermodynamic function can be derived by knowing just that. Just these sums, what are called partition functions. So of course they take on enormous importance. So, we call them partition functions because what they're doing is, they're indicating how the molecules are partitioned among the different available levels. The molecules or a whole system. So the molecular partition function is labeled little q . And the system partition function is labeled big Q . It's called the canonical partition function. And because they are going to take on such special importance, let's just look at some of their properties for different kinds of systems and in general.

OK, first of all, let's talk about units and values. They are unitless. This is an exponential function. Here are units of energy and energy. But this is a unitless was number. It's just some number, right? Could be 1. Could be 10. Could be 50. Whatever, right? Could be 10 to the 24. Its magnitude tells you about more or less how many states are thermally accessible. Because, look at it. And then go back to the example that I showed you if lots of these terms are big, are significant because this is really a big number, right? It's really hot. So lots of states have energies lower than kT , which means this is not too small, for lots and lots of values of i , then this just keeps adding up. And of course, same here. So the number might be very large.

But if we're in the low temperature limit, maybe we're in such low temperature that only the lowest possible state is occupied. And everything else, it's just too cold. There's not enough thermal energy to occupy anything but the very lowest state available. Well, in that case the lowest state, this would be one. Essentially we could label this zero. We could put the energy, the zero of energy there. Everything else is really big compared to kT , which means this exponential gets to be a really small number for every state except one. And this is just equal to one. In other words, the magnitude of these numbers tells us about how many states are accessible, thermally accessible, to molecules or to a whole system.

So let's just go through a couple of specific examples to try to make that a little more concrete. One is, let's start in the simplest case. Which it'll turn out I just alluded to. Let's consider a perfect atomic crystal at essentially zero degrees Kelvin. Zero Kelvin. So every atom. Or even if it's a molecular crystal, every molecule, they're all in the ground state. There's no excess thermal energy. Every molecule is in its proper place. Every atom, if it's an atomic lattice. In other words, it's in the ground state. That's it. So again, we can place the zero where we like. We'll place it there. That one, for that one state, this will be equal to one. And for everything else it'll be zero.

So Q is just the sum of $e^{-E_i/kT}$. It's equal to $e^{-0/kT}$ plus $e^{-E_1/kT}$ plus $e^{-E_2/kT}$. But remember, T is really tiny. It's almost zero degrees Kelvin. So all these things are much bigger than that. So this is vanishingly small. This is vanishingly small. The whole thing is approximately equal to one. That's it.

Also, if we say OK, now what's the probability of the system being in a particular state. Well, we have an expression for that. So let's look at P_0 . It's $e^{-E_0/kT}$ over the sum. Which we've just seen. So it's $e^{-E_0/kT}$ divided by $e^{-E_0/kT} + e^{-E_1/kT}$, and so on. This is the only term that's significant. So it's approximately equal to one.

Now, while I've got this written here, let's just make sure we've got something clear. What if we hadn't arbitrarily set the zero of energy equal to zero? I mean, it's arbitrary, right? We can put the zero of energy anywhere. And you might think, well, gee that's going to have a big effect on everything. Well, it would have an effect on the actual number that we get for Q . But what you'll see is that any measurable quantity that we calculate won't be affected. It's only the zero of the energy scale that's going to be affected. So for example, let's look at this probability. It doesn't matter where we put the zero of energy. This term is still going to be enormously bigger than the next one. And the next one. And the next one. So in this sum only this term is going to matter. It's going to cancel with this term. So whether or not these individual terms are equal to one, which happens if we set this to zero, or whether they're equal to some other number. Still, the probability that the system is in the lowest state is one, right?

That doesn't depend on where we arbitrarily put the zero of the energy scale. The state is still going to be in the ground state, at essentially zero degrees Kelvin. And it'll turn out to be that way with any property that we can measure. Only the zero of this scale moves if we arbitrarily move it somewhere. But the observable, measurable quantities that we calculate, they won't change. Other than that scale. So that's one simple example.

Now let's look at another example. Let's consider a mole of atoms, roaming around in the gas phase at room temperature. OK, so now what I want to do is just have a simple model for their translational motion. And of course, we could do that either quantum mechanically or classically solve for that. We're going to use an even simpler model. And this model is going to be very, very useful for a lot of the things that we'll treat. It's called a lattice model. All it means is, we're going to divide up the available volume. This room, for example, into zillions of tiny little elements. Little volume elements. Each one about the size of an atom. The idea being, we're going to specify the state of the atom by saying where is it. Is it in this lattice sight, in this one, in this one, in this one.

So it's a lattice model. Might be an atom there. Might be one there. So we're going to divide the volume up. So let's call our atomic volume little v . Our total volume big V . And an atomic volume, it's going to be on the order of 10^{-30} meters cubed. Or 10^{-30} meters cubed. And our room, our volume macroscopic one, might be on the order of one meter cubed. Ordinary sort of size.

So now let's figure out our molecular partition function. Now, implicit in this, we're basically saying that the energy, the translational energy, is basically zero. In other words, all these states have the same energy. They're just

located in different positions. At any given instant of time. And what that means is all these terms, all these Boltzmann factors, we just set them equal to one. We'll set the zero of energy there, be done with it. It's a simple model. But it's going to give us the right order of magnitude that we're after. So, how many states are there that are accessible? Well, on the order of 10^{30} , right? So q , little q , we'll call it little q translational, it's just discussing where things are. Is on the order of 10^{30} . And if we do a more careful treatment, if we treat the translational energy of atoms, either classically or quantum mechanically and solve it. We'll still get, we still do get, about the same order of magnitude.

Now let's treat the whole system. So, what happens? What are our total possible states? Because we have to add this up for every possible state. Well, let's start with the first atom. It has to be somewhere. It has 10^{30} possibilities for where it can be. Let's start with the second atom. And put it somewhere. Well, it has 10^{30} minus one, which is still pretty close to 10^{30} . Let's let's go to the third atom. And the fourth. If we have about a mole of atoms, let's say 10^{24} atoms, that's still going to occupy a very small fraction of the site. Only one in a million. So we don't have to keep careful track. Every one of them, we can just say, look, there are 10^{30} available sites. Because we're not going to worry about a change in one in a millionth.

So what that means is, capital Q , trans for the system, is 10^{30} times 10^{30} . In other words, let's now take the whole state. Well, I could put atom one here. I have 10^{30} possibilities. Atom two, I can put anywhere else. So the joint probability of that particular state, for just the two atoms, is 10^{30} times 10^{30} . It's 10^{30} squared. So this is going to keep going. It's going to be 10^{30} to the N th power. Where N is the number of atoms, which is 10^{24} . Huge number, right? It is a huge number. And that, too, if we treat the whole thing classically or quantum mechanically and work it all out. We'll get that number. Or something on that order. Because you know there really is a simply astronomical number of states accessible to the whole bunch of atoms or molecules in this room. It really is that big. So in other words, capital Q is just an astronomical number. And it is the case.

OK. There's an important sort of nuance that we need to introduce. And it's the following. Turns out, it is an astronomical number. But a tiny bit less astronomical than what I've treated so far would indicate. So let's look a little more carefully. What I've said is that Q translational is little q translational, that is, that 10^{30} number. To the N th power. But there's one failing here. Which is, when I got that, when I decided that if I have just the first two atoms I've got 10^{30} times 10^{30} possible states, the trouble with that is then if I keep counting all the possible atoms starting with each one, I'll double-count it, right? In other words, what if I interchange those two atoms? Well, those states are identical, right? Indistinguishable. And it's only distinguishable states that count. When you specify these things, these are indicating distinct states. In some way, at least in principle, measurably different. If the atom's are identical, in other words, if it's a mole of the same stuff,

I don't have a way of distinguishing between those two. I have to correct for that. And when I come to the third atom, I have to correct for all the possible interchanges. Of course, it's three factorial. And in general, it's N factorial. So this result needs to be modified. This is true for distinguishable particles. In other words, if I had all different atoms, so I could label them all, then I don't have that correction. Because then there's really a difference between that state and the state with the two atoms interchanged. But if it's a mole of identical atoms, that's no longer the case. So then, Q translational is little q translational to the N power divided by N factorial. It's still going to be an absolutely enormous number. But it's going to be a little less absolutely enormous than it was a minute ago.

Now finally, I just want to introduce a handy approximation to N factorial that's going to turn out to be very useful again and again. And that's called Stirling's approximation. For the log of a big number, $\ln N$ factorial is $N \log N$ minus N . If we take e to those, both sides, then we find that N factorial is equal to e to the minus N N to the N th power. So this, then, is approximately equal to q translational to the N th power. Over N to the N times e to the minus N . Now let's put the numbers back in. There's our 10 to the 30 th to the 10 to the 24 th power. And now we're going to diminish that just a bit. It's N to the N th power. So it's 10 to the 24 th to the 10 to the 24 th power. Times e to the minus 10 to the 24 th power. So, we can cancel something here. So we have 10 to the 6 th to the 10 to the 24 th power times e to the 10 to the 24 th power. e is about 10 to the 0.4 th power. So this whole thing is about 10 to the 6.4 th power to the 10 to the 24 th power. It's still a pretty respectable number. You could still say astronomical. But maybe astronomical, but not quite to the edge of the universe. Whereas the one before maybe hit the edge of the universe and then beyond.

So that's our second example. And again, part of what I'm trying to do here is just introduce some ideas. Things like this way of modeling positions and so forth. But also, again, orders of magnitude. How big are these numbers. For different sorts of systems. So let's do one more example.

Now, let's consider a polymer in a liquid. And it has different configurations. And they might be a little bit different in energy. For example, some configurations might bring neighboring regions of the polymer into proximity where they could hydrogen bond. And the point is that then the molecular energies involved will change a little bit. Because of some sorts of interactions that are possible in some configurations. And not in other configurations.

So what I'm trying to do is introduce a very simple framework through which we might be able to look at things like protein folding, or DNA hydrogen bonding. Things like this. And just in a simple way model how those work and what the forces are that drive them. So we'll think about polymer configurations. So let's look at a few configurations. Here is going to be one. I'm going to label this one here. And then here are a few others. So I've labeled them this way so that you can see how they are distinct from each other. This one would have a possible interaction. So we'll label its energy. So this is molecular state i , here's going to be our energy, E_i . And let's call

this one negative ϵ_{int} , for an interaction energy that's favorable. And these will be zero. Let's just put that there. Zero, zero, zero. And let's also indicate the degeneracy. How many states, different states are there with the same energy. And that's called g_i . And here it's one. And here are these three.

So that's the framework of our model. Well, so what's our molecular partition function for this configurational degree of freedom? Well, we can label it little q configurational. So we're going to sum over these states. $e^{-\epsilon_i}$ to the minus ϵ_i for the different configurations over kT . So it's $e^{-\epsilon_{int}}$ over kT . Plus three times e^{-0} over kT , we'll get all those other three terms. So that's it. It's $e^{-\epsilon_{int}}$ over kT plus three. Remember, the interaction energy is negative ϵ_{int} . It's a favorable interaction. ϵ_{int} is a positive number.

So that's our result. Now we've described the probabilities in terms of the states that can be occupied. That is, we've added it up. But of course, even the way I wrote it just out of convenience, I didn't actually write out each term in the sum. In the notes I actually did that. But of course it's not necessary to write e^{-0} over kT plus e^{-0} over kT and write that three times for each of these three. Rather, it's convenient to group them together. And the point I'm illustrating here is that in our expression, for the partition function, we've written that in terms of the individual states. But we doing a sum over the individual molecular states. But we could also sum over energy levels. Including the degeneracy. So we could say, let's not do the sum over every state. After all, what if there are a hundred states that had the same energy. Rather than just three. Gets to be kind of painful, right? Instead let's just sum over energy levels. They're all going to be the same factor anyway. And then we'll have, we'll just explicitly write the degeneracy in there.

So, we can write the partition function as a sum over energy level. So, in other words q is the sum over i , $e^{-\epsilon_i}$ over kT . That's individual molecular states i . But we also could write it as the sum over i , where this now is molecular energy levels i . And then we need to incorporate the degeneracy $g_i e^{-\epsilon_i}$ over kT . Same thing, of course. But again, sometimes much more convenient to write things this way. And of course, we can write the probabilities of the occupancies in the same way, too. That is, we could talk about the P_i in terms of individual states. $e^{-\epsilon_i}$ over kT over q , the whole sum. Or P_i summing over energy levels, $g_i e^{-\epsilon_i}$ over kT , divided by q .

Here, of course, this is bigger if it's degenerate, right? It's saying that the probability of being in this energy level is three times the probability of it being in any individual one of the states. But sometimes it's useful to keep track of things in this way. What it shows you too is that, remember when we looked at, did I erase it? I guess it's gone. When we looked at this probability distribution for the occupancies of the levels, of course, what it says is at any temperature, the lowest level is the most probable. For intermediate temperatures. For low temperatures, for high temperatures. At high temperature, it might be only a little more probable than the next one over, and the next one, and so forth. But the very lowest state always has the highest probability.

But the lowest energy doesn't always have the highest probability because of degeneracies. There might be many states with an energy up here. And the probability of any one of them is only a little lower than the probability of the lowest energy. That could be the case here. Let's say, the interaction energy isn't enormously strong. So there is some energetic favoring of this state. Maybe there are 10% at room temperature. Maybe it turns out there are 10% more molecules like this than in any one of these states. But of course, these three altogether mean that there are many more molecules at this energy than at the lowest energy.

Now of course we could do the same thing for the canonical partition function. Not just the molecular one. So in other words, $Q = \sum_i e^{-\beta E_i}$. But we could also write it as a sum over energies. $Q = \sum_i g_i e^{-\beta E_i}$. And then we have to include the degeneracy. g_i here is the degeneracy of molecular energy E_i . Now, same form. Just an important difference, though. This number, this little g_i , is typically a small number. It could be one, it could be a few. This number is usually astronomical. That's basically, that is, what we calculated here. In other words, how many total system states are there with a particular energy. Well, in many, many, many cases the answer is some astronomical number.

So this number might often be between one and ten. This number might be 10^{24} . It might be 10^{24} to a large power. And, of course, that has a big effect on the way things end up working. In statistical mechanics and in thermodynamics. A lot of thermodynamics results are the way they are because you have so many possible states with a particular energy. That that energy can be strongly favored just by virtue of the number of states that there are. Just the way, in a very small way, this energy might be favored just because there are more states with it than there are states here. But again, with the system, it's not a factor of three to one. It might be a factor of 10^{24} to the power of something. It might be just enormously larger than other possibilities that'll tend to put the energy at a certain place.

And just to continue, of course, the same thing goes for the system probabilities. P_i , right, which is $e^{-\beta E_i} / Q$. If this is a sum over states, or $\sum_i g_i e^{-\beta E_i} / Q$, let's write this over, if now we're calculating the probability of an energy level. And it's important to make the distinction. Because in many cases, that's what we care about. What's the energy of the system? And we often don't care about exactly where's that molecule and that one, and that one, and that one, right? The individual states that might be involved that would comprise that energy.

Now, what I want to do is start deriving thermodynamics. Like I promised, we're going to be able to derive every thermodynamic quantity if we just know the partition function. So now I just want to show that that might really be true.

So, the point is that from Q we're going to get all of our thermodynamics. And let's start with the energy. Remember u , right? That's our system energy. It's an average energy. It's an average of the energy that we would get by looking at the states that the system might be occupying. So we can write it as a sum over i . Of P_i times E_i . In other words, it's going to be determined by the energy of any system state times the probability that the system is in that state. Add them all up. Well, we know what that is. It's the sum i of $E_i e^{-\beta E_i}$ over Q . Now, I'm going to just make a simple substitution. I'm going to use the term β to mean one over kT . I'm really just using that so I don't need to use the chain rule a billion times and doing derivatives. So I'm going to write this is sum over i $E_i e^{-\beta E_i}$ over Q .

So Q , then, in this term is just a sum over i $e^{-\beta E_i}$. So. Now I do want to take some derivatives. If I take $dQ / d\beta$, keeping V and N constant, then I get derivative with respect to β . Sum over i $e^{-\beta E_i}$. So that's going to bring out E_i , right? So it's going to bring out minus E_i minus the sum over i , $E_i e^{-\beta E_i}$. That obviously looks like a handy thing. Because that looks like that term.

Then, our average energy, remember, it's one over Q . Sum of i , $E_i e^{-\beta E_i}$. So now, it's just minus one over Q , $dQ / d\beta$. So that's just the same as minus $d \log Q / d\beta$. And now I'm going to use the chain rule. So it's minus $d \log Q / dT$ times $dT / d\beta$. All at constant V and N . Now I can easily get $dT / d\beta$. Because $d\beta / dT$ is just minus one over kT squared. It's just the derivative of one over kT with respect to T . So finally, my average E , which is u , is just kT squared times $d \log Q / dT$ constant V, N . That's terrific. In other words, if I have an expression for Q , I know the partition function, and I can calculate it at any temperature. I just need to take log of it, take its derivative with respect to temperature. Multiply it by kT squared and I've got my energy. Not very complicated, right? So in other words, macroscopic thermodynamic properties come straight out of our microscopic model of statistical mechanics. Statistical thermodynamics.

Now I'm just going to state the next result, just because I want to get there and it'll be followed up more next time. But it's the following. You can see, of course, our Q is a function of V and N and T . It's a function of V because in principle the energies that are going into all this can be a function of volume. What thermodynamic function is naturally a function of N, V, T ? Who remembers? Gibbs free energy? Helmholtz free energy? Enthalpy? Which one? Nobody knows. What's a function of N, V , and T . Or, V and T , is what we really formulate it as. N was introduced later. The Helmholtz free energy. Thank you. What that suggests is that actually the simplest and most natural connection between Q and macroscopic thermodynamics is to the Helmholtz free energy. And the result that you'll see derived next time is A is just minus $kT \log$ of Q . What a simple result. And you'll see the derivations in your notes. You can see it's a couple of lines, right? But of course, if you know A , and you know E , you know everything, right? Because S is in there. And then other combinations can give S and G and μ . And p . And anything else you want. So that's the point, is that all of macroscopic thermodynamics follows. And you'll see that

elaborated more next time. And in addition, a very simple natural expression for the entropy in terms of the states available follows, that we've alluded to before. And now you'll see it played out. Right.