

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.

**MEHRAN
KARDAR:**

You decide at first look at the our simple system, which was the ideal gas. And imagine that we have this gas contained in a box of volume V that contains N particles. And it's completely isolated from the rest of the universe. So you can know the amount of energy that it has.

So the macroscopic description of the system consists of these three numbers, E , V , and N . And this was the characteristics of a microcanonical ensemble in which there was no exchange of heat or work. And therefore, energy was conserved.

And our task was somehow to characterize the probability to find the system in some microstate. Now if you have N particles in the system, there is at the microscopic level, some microstate that consists of a description of all of the positions of momenta. And there is Hamiltonian that governs how that microstate evolves as a function of time.

And for the case of ideal gas, the particles don't interact. So the Hamiltonian can be written as the sum of n terms that describe essentially the energy of the end particle composed of its kinetic energy. And the term that is really just confining the particle is in this box. And so the volume of the box is contained. Let's say that in this potential, that it's zero inside the box and infinity outside.

And we said, OK, so given that I know what the energy, volume, number of particles are, what's the chance that I will find the system in some particular microstate? And the answer was that, obviously, you will have to put zero if the particles are outside box. Or if the energy, which is really just the kinetic energy, does not much the energy that we know is in the system, we sum over i of P_i^2 over $2m$ is not equal to the energy of the system.

Otherwise, we say that if the microstate corresponds to exactly the right amount of energy, then I have no reason to exclude it. And just like saying that the dice can have six possible faces, you would assign all of those possible phases equal probability. I will give all of the microstates that don't conflict with the conditions that I have set out the same probability. I will call that probability 1 over some overall constant ω . And so this is one otherwise.

So then the next statement is well what is this number ω that you have put it? And how do we determine it? Well, we know that this P is a probability so that if I were to integrate over the entirety of the phase space of this probability, the answer should be 1 .

So that means this ω , which is a function of these parameters that I set out from the outside to describe the microstate, should be obtained by integrating over all q and p of this collection of 1 s and 0 s that I have out here. So I think this box of 1 s and 0 s, put it here, and I integrate.

So what do I get? Well, the integration over the q 's is easy. The places that I get 1 are when the q 's are inside the box. So each one of them will give me a factor of V . And there are N of them. So I would get V to the N .

The integrations over momenta essentially have to do with seeing whether or not the condition $\sum_i p_i^2 / 2m = E$ is satisfied or not. So this I can write also as $\sum_i p_i^2 = 2mE$, which I can write as R^2 . And essentially, in this momentum space, I have to make sure that the sum of the components of all of the momenta squared add up to this R^2 , which as we discussed last time, is the surface of hypersphere in $3N$ dimensions of radius R , which is square root of $2mE$.

So I have to integrate over all of these momenta. And most of the time I will get 0 , except when I hit the surface of this sphere. There's kind of a little bit of singularity here because you have a probability that there's 0 , except at the very sharp interval, and then 0 again. So it's kind of like a delta function, which is maybe a little bit hard to deal with.

So sometimes we will generalize this by adding a little bit be delta E here. So let's say that the energy does not have to be exactly E, but E minus plus a little bit, so that when we look at this surface in three n dimensional space-- let's say this was two dimensional space-- rather than having to deal with an exact boundary, we have kind of smoothed that out into an interval that has some kind of a thickness R, that presumably is related to this delta E that I put up there.

Turns out it doesn't really make any difference. The reason it doesn't make any difference I will tell you shortly.

But now when I'm integrating over all of these P's-- so there's P. There's another P. This could be P1. This could be P2. And there are different components. I then get 0, except when I hit this interval around the surface of this hypersphere.

So what do I get as a result of the integration over this 3N dimensional space? I will get the volume of this element, which is composed of the surface area, which has some kind of a solid angle in 3N dimensions. The radius raised to the power of dimension minus 1, because it's a surface.

And then if I want to really include a delta R to make it into a volume, this would be the appropriate volume of this interval in momentum space. Yes.

AUDIENCE: Just to clarify, you're asserting that there's no potential inside the [INAUDIBLE] that comes from the hard walls.

MEHRAN
KARDAR: Correct. We can elaborate on that later on. But for the description of the ideal gas without potential, like in the box, I have said that potential to be just 0 infinity. OK?

OK, so fine. So this is the description. There was one thing that I needed to tell you, which is the d, dimension, of solid angle, which is 2π to the d over 2 divided by d over 2 minus 1 factorial So again, in two dimensions, such as the picture that I drew over here, the circumference of a circle would be $2\pi r$.

So this s sub 2-- right there'd be 2π -- and you can show that it is 2π divided by 0

factorial, which is 1. In three dimensions it should give you $4\pi r^2$. Kind of looks strange because you get 2π to the $3/2$ divided by $1/2$ half factorial. But the $1/2$ factorial is in fact $\sqrt{2/\pi}$ -- $\sqrt{\pi}/2$. And so this will work out fine.

Again, the definition of this factorial in general is through the gamma function and an integral that we saw already. And factorial is the integral 0 to infinity, dx , x to the n into the minus x .

Now, the thing is that this is a quantity that for large values of dimension grows exponentially $\propto e^d$. So what I claim is that if I take the log of this surface area and take the limit that d is much larger than 1, the quantity that I will get -- well, let's take the log of this.

I will get \log of 2 . I will get $d/2 \log \pi$ and minus the log of this large factorial. And the log of the factorial I will use Sterling's formula. I will ignore in that large limit the difference between $d/2$ and $d/2 - 1$. Or actually I guess at the beginning I may even write it $d/2 - 1 \log$ of $d/2 - 1$ plus $d/2 - 1$.

Now if I'm in this limit of large d again, I can ignore the 1s. And I can ignore the $\log 2$ with respect to this $d/2$. And so the answer in this limit is in fact proportional to $d/2$. And I have the \log . I have π . I have this $d/2$ that will carry in the denominator. And then this $d/2$ times 1 I can write as $d/2 \log e$. And so this is the answer we get.

So you can see that the answer is exponentially large if I were to again write s of d . S of d grows like an exponential in d . OK, so what do I conclude from that?

I conclude that s over Kd -- and we said that the entropy, we can regard as the entropy of this probability distribution. So that's going to give me the log of this ω . And log off this ω , we'll get a factor from this v to the n . So I will get $N \log V$. I will get a factor from log of S of $3N$.

I figured out what that log was in the limit of large dimensions. So I essentially have $3N/2$ because my d is now roughly $3N$. It's in fact exactly $3N$, sorry.

I have the log of $2\pi e$. For d I have $3N$. And then I actually have also from here a $3N \log R$, which I can write as $3N$ over $2 \log$ of R squared. And my R squared is $2mE$. The figure I have here.

And then you say, OK, we added this ΔR . But now you can see that I can also ignore this ΔR , because everything else that I have in this expression is something that grows radially with N . What's the worse that I can do for ΔR ?

I could make ΔR even as big as the entirety of this volume. And then the typical volume would be of the order of the energy-- sorry, the typical value of R would be like the square root of energy. So here I would have to put this log of the square root of the energy. And log of a square roots of an extensive quantity is much less than the extensive quantity. I can ignore it.

And actually this reminds me, that some 35 years ago when I was taking this course, from Professor Felix Villiers, he said that he had gone to lunch. And he had gotten to this very beautiful, large orange. And he was excited. And he opened up the orange, and it was all skin. And there was just a little bit in the middle.

He was saying it is like this. It's all in the surface. So if Professor Villiers had an orange in $3N$ dimension, he would have exponentially hard time extracting an orange.

So this is our formula for the entropy of this gas. Essentially the extensive parts, $n \log v$ and something that depends on $n \log E$. And that's really all we need to figure out all of the thermodynamic properties, because we said that we can construct-- that's in thermodynamics-- dE is TdS minus PdV plus YdN in the case of a gas.

And so we can rearrange that to dS be dE over T plus P over T dV minus Y over T dN . And the first thing that we see is by taking the derivative of S with respect to the quantities that we have established, E , V , and N , we should be able to read off appropriate quantities.

And in particular, let's say 1 over T would be dS by dE of constant v and n . S will be

proportional to k_B . And then the dependence of this object on E only appears on this $\log E$. Except that there's a factor of $3N$ over 2 out front. And the derivative of $\log E$ with respect to E is 1 over E .

So I can certainly immediately rearrange this and to get that the energy is $\frac{3}{2} N k_B T$ in this system of ideal point particles in three dimensions. And then the pressure, P over T , is the S by dV at constant e and n . And it's again, k_B . The only dependence on V is through this $N \log V$. So I will get a factor of N over V , which I can rearrange to PV is $N k_B T$ by the ideal gas law.

And in principle, the next step would be to calculate the chemical potential. But we will leave that for the time being for reasons that will become apparent.

Now, one thing to note is that what you have postulated here, right at the beginning, is much, much, more information than what we extracted here about thermodynamic properties. It's a statement about a joint probability distribution in this $6N$ dimensional phase space. So it has huge amount of information.

Just to show you part of it, let's take a note at the following. What it is a probability as a function of all coordinates and momenta across your system. But let me ask a specific question.

I can ask what's the probability that some particular particle-- say particle number one-- has a momentum P_1 . It's the only question that I care to ask about this huge amount of degrees of freedom that are encoded in P of μ . And so what do I do if I don't really care about all of the other degrees of freedom is I will integrate them.

So I don't really care where particle number one is located. I didn't ask where it is in the box. I don't really care where part because numbers two through N are located or which momenta they have. So I integrate over all of those things of the full joint probability, which depends on the entirety of the phase space.

Fine, you say, OK. This joint probability actually has a very simple form. It is 1 over this $\Omega(E, V, N)$, multiplying either 1 or 0 . So I have to integrate over all of these q_i 's, all of these p_i . Of 1 over Ω or 0 over Ω , this delta like

function that we put in a box up there-- so this is this delta function that says that the particular should be inside the box. And the sum of the momenta should be on the surface of this hypersphere.

Now, let's do these integrations. Let's do it here. I may need space. The integration over Q_1 is very simple. It will give me a factor of V .

I have this ω , E , V , N , in the denominator. And I claim that the numerator is simply the following. $\omega E - P_1^2$ over $2m V N - 1$. Why? Because what I need to do over here in terms of integrations is pretty much what I would have to integrate over here that gave rise to that surface and all of those factors with one exception.

First of all, I integrated over one particle already, so the coordinate momenta here that I'm integrating pertains to the remaining $N - 1$. Hence, the ω pertains to $N - 1$. It's still in the same box of volume V . So V , the other argument, is the same.

But the energy is changed. Why? Because I told you how much momentum I want the first particle to carry. So given the knowledge that I'm looking at the probability of the first particle having momentum P_1 , then I know that the remainder of the energy should be shared among the momenta of all the remaining $N - 1$ particles.

So I have already calculated these ω 's up here. All I need to do is to substitute them over here. And I will get this probability.

So first of all, let's check that the volume part cancels. I have one factor a volume here. Each of my ω 's is in fact proportional to V to the N . So the denominator has V to the N . The numerator has a V to the $N - 1$. And all of the V 's would cancel out.

So the interesting thing really comes from these solid angle and radius parts. The solid angle is a ratio of-- let's write the denominator. It's easier. It is 2π to the $3N$ over 2 divided by $3N$ over 2 minus one factorial. The numerator would be 2π to the

$3N$ and minus 1 over 2 divided by $3N$ minus 1 over 2 minus 1 factorial.

And then I have these ratio of the radii. In the denominator I have $2mE$ to the power of $3N$ minus 1 over 2 minus 1. So minus $3N$ -- it is $3N$ minus 1 over 2. Same thing that we have been calculating so far.

And in the numerator it is $2mE$ minus P_1 squared over $2m$. So I will factor out the E . I have 1 minus P_1 squared over $2mE$. The whole thing raised to something that is $3N$ minus 1 minus 1 over 2.

Now, the most important part of this is the fact that the dependence on P_1 appears as follows. I have this factor of 1 minus P_1 squared over $2mE$. That's the one place that P_1 , the momentum of the particle that I'm interested appears. And it raised to the huge power, which is of the order of $3N$ over 2.

It is likely less. But it really doesn't make any difference whether I write $3N$ over 2, $3N$ over minus 1 over 2, et cetera. Really, ultimately, what I will have is 1 minus the very small number, because presumably the energy of one part they can is less than the energy of the entirety of the particle. So this is something that is order of 1 out of N raised to something that is order of N . So that's where an exponentiation will come into play.

And then there's a whole bunch of other factors that if I don't make any mistake I can try to write down. There is the $2s$ certainly cancel when I look at the factors of π . The denominator with respect to the numerator has an additional factor of π to the $3/2$. In fact, I will have a whole bunch of things that are raised to the powers of $3/2$. I also have this $2mE$ that compared to the $2mE$ that comes out front has an additional factor of $3/2$. So let's put all of them together. $2\pi mE$ raised to the power of $3/2$.

And then I have the ratio of these factorials. And again, the factorial that I have in the denominator has one and a half times more or $3/2$ times more than what is in the numerator. Roughly it is something like the ratio of $3N$ over 2 factorial divided by $3N$ minus 1 over 2 factorial.

And I claim that, say, N factorial compared to $N - 1$ factorial is larger by a factor of N . If I go between N factorial $N - 2$ factorial is a factor that is roughly N squared. Now this does not shift either by 1 or by 2, but by 1 and $1/2$. And if you go through Sterling formula, et cetera, you can convince yourself that this is roughly $3N$ over 2 to the power of 1 and $1/2$ -- $3/2$.

And so once you do all of your arrangements, what do you get? 1 minus a small quantity raised to a huge power, that's the definition of the exponential. So I get exponential of minus P^2 squared over $2m$. And the factor that multiplies it is E . And then I have $3N$ over 2.

And again, if I have not made any mistake and I'm careful with all of the other factors that remain, I have here $2\pi m E$. And this E also gets multiplied by the inverse of $3N$ over 2. So I will have this replaced by $2E$ over $3N$.

So statement number one, this assignment of probabilities according to just throwing the dice and saying that everything that has the same right energy is equally likely is equivalent to looking at one of the particles and stating that the momentum of that part again is Gaussian distributed. Secondly, you can check that this combination, $2E$ divided by $3N$ is the same thing as kT . So essentially this you could also if you want to replace 1 over kT . And you would get the more familiar kind of Maxwell type of distribution for the momentum of a single particle in an ideal gas.

And again, since everything that we did was consistent with the laws of probability, if we did not mix up the orders of N , et cetera, the answer should be properly normalized. And indeed, you can check that this is the three dimensional normalization that you would require for this gas. So the statement of saying that everything is allowed is equally likely is a huge statement in space of possible configurations.

On the one hand, it gives you macroscopic information. On the other hand, it retains a huge amount of microscopic information. The parts of it that are relevant, you can

try to extract this here. OK?

So those were the successes. Question is why didn't I calculate for you this u over T ? It is because this expression as we wrote down has a glaring problem with it, which in order to make it explicit, we will look at mixing entropies.

So the idea is this is as follows. Let's imagine that we start with two gases. Initially, I have N_1 particles of one type in volume 1. And I have N_2 particles of another type in volume 2. And for simplicity I will assume that both of them are of the same temperature.

So this is my initial state. And then I remove the partition. And I come up with this situation where the particles are mixed. So the particles of type 1 could be either way. Particles of type 2 could be in either place.

And let's say I have a box of some toxic gas here. And I remove the lid. And it will get mixed in the room.

It's certainly an irreversible situation where is an increase of entropy i associated with that. And we can calculate that increase of entropy, because we know what the expression for entropy is. So what we have to do is to compare the entropy initially. So this is the initial entropy. And I calculate everything in units of k_B so I don't have to write k_B all over the place.

For particle number one, what do I have? I have $N_1 \log V_1$. And then I have a contribution, which is $\frac{3}{2} N_1$. But I notice that whatever appears here is really only a function of E over N . E over N is really only a function of temperature. So this is something that I can call a σ of T over here.

And the contribution of box 2 is $N_2 \log V$ plus $\frac{3}{2} N_2$. This-- huh-- let's say that they are-- we ignore the difference in masses. You could potentially have here σ_1 , σ_2 . It really doesn't make any difference.

The final state, what do we have? Essentially, the one thing that changed is that the N_1 particles now are occupying the box of volume V . So if call V to the V_1 plus V_2 ,

what we have is that we have $N_1 \log V$ plus $N_2 \log V$.

My claim is that all of these other factors really stay the same. Because essentially what is happening in these expressions are various ratios of E over N . And by stating that initially I had the things at the same temperature, what I had effectively stated was that E_1 over N_1 is the same thing as E_2 over N_2 .

I guess in the ideal gas case this E over N is the same thing as $3/2 kT$. But if I have a ratio such as this, that is also the same as E_1 plus E_2 divided by N_1 plus N_2 . This is a simple manipulation of fractions that I can make.

And E_1 plus E_2 over N_1 plus N_2 , by the same kinds of arguments, would give me the final temperature. So what I have to compute is that the final temperature is the same thing as the initial temperature. Essentially, in this mixing of the ideal gases, temperature does not change.

So basically, these factors of σ are the same before and after. And so when we calculate the increase in entropy, S_f minus S_i , really the contribution that you care about comes from these volume factors. And really the statement is that in one particle currently are occupying a volume of size V , whereas previously they were in V_1 .

And similarly for the N_2 particles. And if you have more of these particles, more of these boxes, you could see how the general expression for the mixing entropy goes. And so that's fine.

V is certainly greater than V_1 or V_2 . Each of these logs gives you a positive contribution. There's an increase in entropy as we expect.

Now, there is the following difficulty however. What if the gases are identical-- are the same? We definitely have to do this if I take a box of methane here and I open it, we all know that something has happened. There is an irreversible process that has occurred.

But if the box-- I have essentially taken the air in this room, put it in this box,

whether I open the lid or not open the lid, it doesn't make any difference. There is no additional work that I have to do in order to close or open the lid. Is there no increase of entropy one way or the other.

Whereas if I look at this expression, this expression only depends on the final volume and the initial volumes, and says that there should be an increase in entropy when we know that there shouldn't be. And of course, the resolution for that is something like this. That if I look at my two boxes-- and I said maybe one of them is a box that contains methane. Let's call it A. And the other is the room that contains the air.

Now this situation where all of the methane is in the box and the oxygen is freely floating in the room is certainly different from a configuration where I exchange these two and the methane is here and the oxygen went into the box. They're different configurations. You can definitely tell them apart.

Whereas if I do the same thing, but the box and outside contain the same entity, and the same entity is, let's say, oxygen, then how can you tell apart these two configurations? And so the meaning of-- yes.

AUDIENCE: Are you thinking quantum mechanically or classically. Classically we can tell them apart, right?

MEHRAN KARDAR: This is currently I am making a macroscopic statement. Now when I get to the distinction of microstates we have to-- so I was very careful in saying whether or not you could tell apart whether it is methane or oxygen. So this was a very macroscopic statement as to whether or not you can distinguish this circumstance versus that circumstance. So as far as our senses of this macroscopic process is concerned, these two cases have to be treated differently.

Now, what we have calculated here for these factors are some volume of phase space. And where in the evening you might say that following this procedure you counted these as two distinct cases. In this case, these were two distinct cases. But here, you can't really tell them apart.

So if you can't tell them apart, you shouldn't call them two distinct cases. You have over counted phase space by a factor of two here. And here, I just looked at two particles. If I have N particles, I have over counted the phase space of identical particles by all possible permutations of n objects, it is n factorial.

So there is an over counting of phase space or configurations of N identical particles by a factor of N factorial. I.e., when we said that particle number one can be anywhere in the box, particle number two can be anywhere in the box, all the way to particle number n , well, in fact, I can't tell which each is which. If I can't tell which particle is which, I have to divide by the number of permutations and factors.

Now, as somebody was asking the question, as you were asking the question, classically, if I write a computer program that looks at the trajectories of N particles in the gas in this room, classically, your computer would always know the particle that started over here after many collisions or whatever is the particle that ended up somewhere else. So if you ask the computer, the computer can certainly distinguish these classical trajectories. And then it is kind of strange to say that, well, I have to divide by N factorial because all of these are identical. Again, classically these particles are following specific trajectories. And you know where in phase space they are.

Whereas quantum mechanically, you can't tell that apart. So quantum mechanically, as we will describe later, rather than classical statistical mechanics-- when we do quantum statistical mechanics-- if you have identical particles, you have to write down of a wave function that is either symmetric or anti-symmetric under the exchange of particles.

And when we do eventually the calculations for these factors of 1 over N factorial will emerge very naturally. So I think different people have different perspectives. My own perspective is that this factor really is due to the quantum origin of identity. And classically, you have to sort of fudge it and put it over there.

But some people say that really it's a matter of measurements. And if you can't really tell A and B sufficiently apart, then you don't know. I always go back to the

computer. And say, well, the computer can tell.

But it's kind of immaterial at this stage. It's obvious that for all practical purposes for things that are identical you have to divide by this factor. So what happens if you divide by that factor?

So I have changed all of my calculations now. So when I do the log of-- previously I had V to the N . And it gave me $N \log V$. Now, I have \log of V to the N divided by N factorial. So I will get my Sterling's approximation additional factor of $-\ln N + N$, which I can sort of absorb here in this fashion.

Now you say, well, having done that, you have to first of all show me that you fixed the case of this change in entropy for identical particles, but also you should show me that the previous case where we know there has to be an increase in entropy just because of the gas being different that that is not changed because of this modification that you make. So let's check that.

So for distinct gases, what would be the generalization of this form S_f minus S_i divided by kV ? Well, what happens here? In the case of the final object, I have to divide $N_1 \log$ of V . But that V really becomes V divided by N_1 , because in the volume of size V , I have N_1 oxygen that I can't tell apart.

So I divide by the N_1 factorial for the oxygens. And then I have N_2 methanes that I can't tell apart in that volume, so I divide by essentially N_2 factorial that goes over there. The initial change is over here I would have $N_1 \log$ of V_1 over N_1 . And here I would have had $N_2 \log$ of V_2 over N_2 .

So every one of these expressions that was previously $\log V$, and I had four of them, gets changed. But they get change precisely in a manner that this $N_1 \log$ of N_1 here cancels this $N_1 \log$ of and N_1 here. This $N_2 \log$ of N_2 here cancels this $N_2 \log$ of N_2 here. So the ΔS that I get is precisely the same thing as I had before.

I will get $N_1 \log$ of V over V_1 plus $N_2 \log$ of V over V_2 . So this division, because the oxygens were identical to themselves and methanes were identical to themselves, does not change the mixing entropy of oxygen and nitrogen.

But let's say that both gases are the same. They're both oxygen. Then what happens?

Now, in the final state, I have a box. It has a N_1 plus N_2 particles that are all oxygen. I can't tell them apart. So the contribution from the phase space would be N_1 plus N_2 log of the volume divided by N_1 plus N_2 factorial. That ultimately will give me a factor of N_1 plus N_2 here.

The initial entropy is exactly the one that I calculated before. For the line above, I have N_1 log of V_1 over N_1 minus N_2 log of V_2 over N_2 .

Now certainly, I still expect to see some mixing entropy if I have a box of oxygen that is at very low pressure and is very dilute, and I open it into this room, which is at much higher pressure and is much more dense. So really, the case where I don't expect to see any change in entropy is when the two boxes have the same density. And hence, when I mix them, I would also have exactly the same density.

And you can see that, therefore, all of these factors that are in the log are of the inverse of the same density. And there's N_1 plus N_2 of them that's positive. And N_1 plus N_2 of them that is negative. So the answer in this case, as long as I try to mix identical particles of the same density, if I include this correction to the phase space of identical particles, the answer will be [? 0. ?] Yes?

AUDIENCE:

Question, [INAUDIBLE] in terms of the revolution of the [INAUDIBLE] there is no transition [INAUDIBLE] so that your temporary, and say like, oxygen and nitrogen can catch a molecule, put it in a [? aspertometer. ?] and have different isotopes. You can take like closed isotopes of oxygen and still tell them apart. But this is like their continuous way of choosing a pair of gases which would be arbitrarily closed in atomic mass.

MEHRAN

KARDAR:

So, as I said, there are alternative explanations that I've heard. And that's precisely one of them. And my counter is that what we are putting here is the volume of phase space. And to me that has a very specific meaning. That is there's a set of

coordinates and momenta that are moving according to Hamiltonian trajectories.

And in principle, there is a computer nature that is following these trajectories, or I can actually put them on the computer. And then no matter how long I run and they're identical oxygen molecules, I start with number one here, numbers two here. The computer will say that this is the trajectory of number one and this is the trajectory of numbers two.

So unless I change my definition of phase space and how I am calculating things, I run into this paradox. So what you're saying is forget about that. It's just can tell isotopes apart or something like that. And I'm saying that that's fine. That's perspective, but it has nothing to do with phase space counting. OK?

Fine, now, why didn't I calculate this? It was also for the same reason, because we expect to have quantities that are extensive and quantities that are intensive. And therefore, if I were to, for example, calculate this object, that it should be something that is intensive.

Now the problem is that if I take a derivative with respect N , I have $\log V$. And $\log V$ is clearly something that does not grow proportionately to size but grows proportionately to size logarithmically. So if I make volume twice as big, I will get an additional factor of $\log 2$ here contribution to the chemical potential. And that does not make sense.

But when I do this identity, then this V becomes V over N . And then everything becomes nicely intensive. So if I allowed now to replace this V over N , then I can calculate V over T as dS by dN at constant E and V . And so then essentially I will get to drop the factor of $\log N$ that comes in front, so I will get $kT \log$ of V over N . And then I would have $3/2 \log$ of something, which I can put together as $4 \pi N E$ over $3N$ raised to the $3/2$ power.

And you can see that there were these E 's from Sterling's approximation up there that got dropped here, because you can also take derivative with respect to the N 's that are inside. And you can check that the function of derivatives with respects to

the N's that are inside is precisely to get rid of those factors. OK?

Now, there is still one other thing that is not wrong, but kind of like jarring about the expressions that they've had so far in that right from the beginning, I said that you can certainly calculate entropies out of probabilities as minus log of P average if you like. But it makes sense only if you're dealing with discrete variables, because when you're dealing with continual variables and you have a probability density. And the probability density depends on the units of measurement.

And if you were to change measurement from meters to centimeters or something else, then there will be changes in the probability densities, which would then modify the various factors over here. And that's really also is reflected ultimately in the fact that these combinations of terms that I have written here have dimensions. And it is kind of, again, jarring to have expressions inside the logarithm or in the exponential that are not dimensionless.

So it would be good if we had some way of making all of these dimensionless. And you say, well, really the origin of it is all the way back here, when I was calculating volumes in phase space. And volumes in phase space have dimensions. And that dimensions of pq raised to the $3N$ power really survives all the way down here. So I can say, OK, I choose some quantity as a reference that has the right dimensions of the product of p and q , which is an action. And I divide all of my measurements by that reference unit, so that, for example, here I have $3N$ factors of this.

Or let's say each one of them is h . I divide by some quantity that has units of action. And then I will be set. So basically, the units of this h is the product of p and q .

Now, at this point we have no way of choosing some h as opposed to another h . And so by adding that factor, we can make things look nicer. But then things are undefined after this factor of h . When we do quantum mechanics, another thing that quantum mechanics does is to provide us with precisely [? age ?] of Planck's constant as a measure of these kinds of integrations.

So when we eventually you go to calculate, say, the ideal gas or any other

mechanic system that involves p and q in quantum mechanics, then the phase space becomes discretized. You would have-- The appropriate description would have energies that are discretized corresponding to various other discretization that are eventually the equivalent to dividing by this Planck's constant.

Ultimately, I will have additionally a factor of h squared appearing here. And it will make everything nicely that much [? less. ?] None of these other quantities that I mentioned calculated would be affected by this.

So essentially, what I'm saying is that you are going to use a measure for phase space of identical particles. Previously we had a product, $d^3 p_i$, $d^3 q_i$. This is what we were integrating and requiring that this integration will give us [INAUDIBLE].

Now, we will change this to divide by this N factorial, if the particles are identical. And we divide by h to the $3N$ because of the number of pairs of pq that appear in this. The justification to come when we ultimately do quantum study. Any questions?

So I said that this prescription when we look at a system at complete isolation, and therefore, specify fully its energy is the microcanonical ensemble, as opposed to the canonical ensemble, whereas the set of microscopic parameters that you identified with your system, you replace the energy with temperature. So in general, let's say there will be some bunch of displacements, x , that give you the work content to the system.

Just like we fixed over there the volume and the number of particles, let's say that all of the work parameters, such as x microscopically, we will fix in this canonical ensemble. So however, the ensemble is one in which the energy is not specified. And so how do I imagine that I can maintain a system at temperature T ?

Well, if this room is at some particular temperature, I assume that smaller objects that I put in this room will come to the same temperature. So the general prescription for beginning something other than temperature T is to put it in contact with something that is much bigger. So let's call this to be a reservoir. And we put

our system, which we assume to be smaller, in contact with it. And we allow it to exchange heat with the reservoir.

Now I did this way of managing the system to come to a temperature T , which is the characteristic of a big reservoir. Imagine that you have a lake. And you put your gas or something else inside the lake. And it will equilibrate to the temperature of the lake.

I will assume that the two of them, the system and the reservoir, just for the purpose of my being able to do some computation, are isolated from the rest of the universe, so that the system plus reservoir is microcanonical. And the sum total of their energies is sum E total.

So now this system still is something like a gas. It's a has a huge number of potential degrees of freedom. And these potential number of degrees of freedom can be captured through the microstate of the system, u sub s .

And similarly, the water particle in the lake have their own state. An there's some microstate that describes the positions and the momenta of all of the particles that are in the lake. Yes?

AUDIENCE: When you're writing the set of particles used to describe it, why don't you write N ? Since it said the number of particles in the system is not fixed.

MEHRAN KARDAR: Yes, so I did want to [INAUDIBLE] but in principle I could add N . I wanted to be kind of general. If you like X , [$? it ?$] is allowed to include chemical work type of an X .

So what do I know? I know that there is also if I want to describe microstates and their evolution, I need to specify that there's a Hamiltonian that governs the evolution of these microstates. And presumably there's a Hamiltonian that describes the evolution of the reservoir microstate. And so presumably the allowed microstates are ones in which E total is made up of the energy of the system plus the energy of the reservoir.

So because the whole thing is the microcanonical, I can assign a probability, a joint

probability, to finding some particular μ_s, μ_r combination, just like we were doing over there. You would say that essentially this is a combination of these 1s and 0s.

So it is 0 if H of-- again, for simplicity I drop the s on the system-- H of μ_s plus H reservoir of μ reservoir is not equal to E_{total} . And it is 1 over some ω of reservoir in the system otherwise. So this is just again, throwing the dice, saying that it has so many possible configurations, given that I know what the total energy is. All the ones that are consistent with that are allowed.

Which is to say, I don't really care about the lake. All I care is about the states of my gas, and say, OK, no problem, if I have the joint probability distribution just like I did over here, I get rid of all of the degrees of freedom that I'm not interested in. So if I'm interested only in the states of the system, I sum over or integrate over-- so this would be a sum. This would be an integration, whatever-- of the joint probability distribution.

Now actually follow the steps that I had over here when we were looking at the momentum of a gas particle. I say that what I have over here, this probability see is this $1/\omega_{R,S}$. This is a function that is either 1 or 0. And then I have so sum over all configurations of the reservoir. But given that I said what the microstate of the system is, then I know that the reservoir has to take energy in total minus the amount the microstates has taken.

And I'm summing over all of the microstates that are consistent with the requirement that the energy in the reservoir is $E_{\text{total}} - H$ of microstate. So what that is that the ω that I have for the reservoir-- and I don't know what it is, but whatever it is, evaluated at the total energy minus the energy that is taken outside the microstate.

So again, exactly the reason why this became $E - \frac{p^2}{2m}$. This becomes $E_{\text{total}} - H$ of μ_s . Except that I don't know either what this is or what this is. Actually, I don't really even care about this because all of the H dependents on microstate dependents is in the numerator. So I write that as

proportional to exponential. And the log of omega is the entropy. So I have the entropy of the [INAUDIBLE] in units of k_B , evaluated at the argument that this E total minus H of μS .

So my statement is that when I look at the entropy of the reservoir as a function of E total minus the energy that is taken out by the system, my construction I assume that I'm putting a small volume of gas in contact with a huge lake. So this total energy is overwhelmingly larger than the amount of energy that the system can occupy. So I can make a Taylor expansion of this quantity and say that this is S_R of E total minus the derivative of S with respect to its energy. So the derivative of the S reservoir with respect to the energy of the reservoir times H of the microstate and presumably higher order terms that are negligible.

Now the next thing that is important about the reservoir is you have this huge lake. Let's say it's exactly at some temperature of 30 degrees. And you take some small amount of energy from it to put in the system. The temperature of the lake should not change.

So that's the definition of the reservoir. It's a system that is so big that for the range of energies that we are considering, this S by dE is 1 over the temperature that characterizes the reservoir. So just like here, but eventually the answer that we got was something like the energy of the particle divided by kT . Once I exponentiate, I find that the probability to find the system in some microstate is proportional to E to the minus of the energy of that microstate divided by kT . And of course, there's a bunch of other things that I have to eventually put into a normalization that will cause.

So in the canonical prescription you sort of replace this throwing of the dice and saying that everything is equivalent to saying that well, each microstates can have some particular energy. And the probabilities are partitioned according to the Boltzmann weights of these energies. And clearly this quantity, Z , the normalization is obtained by integrating over the entire space of microstates, or summing over them if they are discrete of this factor of E to the minus βH of μS . And we'll

use this notation $\beta = 1/kT$ sometimes for simplicity.

Now, the thing is that thermodynamically, we said that you can choose any set of parameters, as long as they are independent, to describe the macroscopic equilibrium state of the system. So what we did in the microcanonical ensemble is we specified a number of things, such as energy. And we derived the other things, such as temperature.

So here, in the canonical ensemble, we have stated what the temperature of the system is. Well, then what happened? On one hand, maybe we have to worry because energy is constantly being exchanged with the reservoir. And so the energy of the system does not have a specific value. There's a probability for it.

So probability of system having energy ϵ -- it doesn't have a fixed energy. There is a probability that it should have energy. And this probability, let's say we indicate with $P(\epsilon)$ given that we know what temperature is. Well, on one hand we have this factor of $e^{-\epsilon/kT}$. That comes from the [INAUDIBLE].

But there isn't a single state that has that energy. There's a whole bunch of other states of the system that have that energy. So as I scan the microstates, there will be a huge number of them, $\Omega(\epsilon)$ in number, that have this right energy. And so that's the probability of the energy.

And I can write this as $e^{-\epsilon/kT}$ that I've called β . I have ϵ . And then the log of Ω that I take in the numerator is S divided by k_B . I can take that k_B here and write this as $T S$ of ϵ . And so this kind of should remind you of something like a free energy. But it tells you is that this probability to have some particular energy is some kind of [? a form. ?]

Now note that again for something like a gas or whatever, we expect typical values of both the energy and entropy to be quantities that are proportional to the size of the system. As the size of the system becomes exponentially large, we would expect that this probability would be one of those things that has portions that let's

say are exponentially larger than any other portion.

There will be a factor of E to the minus N something that will really peak up, let's say, the extremum and make the extremum overwhelmingly more likely than other places. Let's try to quantify that a little bit better. Once we have a probability, we can also start calculating averages.

So let's define what the average energy of the system is. The average energy of the system is obtained by summing over all microstates. The energy of that microstate, the probability of that microstate, which is E to the minus βH microstate divided by the partition function, which is the sum-- OK, the normalization, which we will call the partition function, which is the sum over all of these microstates.

Now this is something that we've already seen. If I look at this expression in the denominator that we call Z and has a name, which is the partition function, then it's certainly a function of β . If I take a derivative of Z with respect to β , what happens I'll bring down a factor of H over here.

So the numerator up to a sine is the derivative of Z with respect to β . And the denominator is 1 over Z . And so this is none other than minus the log Z with respect to β . So OK, fine, so the mean value of this probability is given by some expression such as this.

Well, you can see that if I were to repeat this process and rather than taking one derivative, I will take n derivatives and then divide by 1 over Z . Each time I do that, I will bring down a factor of H . So this is going to give me the average of H to the N . The end moment of this probability distribution of energy is obtainable by this procedure.

So now you recognize, oh, I've seen things like such as this. So clearly this partition function is something that generates the moments by taking subsequent derivatives. I can generate different moments of this distribution.

But then there was something else that maybe this should remind you, which is that if there's a quantity that generates moments, then its log generates cumulants. So

you would say, OK, the n th cumulant should be obtainable up to this factor of minus 1 to the n , as the n th derivative with respect to the beta of logs. And it's very easy to check that indeed if I were to take two derivatives, I will get the expectation value of H squared minus the average of H squared, et cetera.

But the point is that clearly this to $\log Z$ is, again, something that is extensive. Another way of getting the normalization-- I guess I forgot to put this 1 over Z here. So now it is a perfectly normalized object.

So another way to get z would be to look at the normalization of the probability. I could integrate over ϵ this factor of E to the minus $\beta \epsilon$ minus $T S$ of ϵ . And that would give me Z . Now, again the quantities that appear in the exponent, energy-- entropy, their difference, free energy-- are quantities that are extensive.

So this Z is going to be dominated again by where this peak is. And therefore, \log of Z will be proportional to \log of what we have over here. And it be an extensive quantity. So ultimately, my statement is that this \log of Z is something that is order of N .

So we are, again, kind of reminiscent of the central limit theorem. In a situation where we have a probability distribution, at large N , in which all of the cumulants are proportional to N . The mean is proportional to N . The variance is proportional to N . All of the cumulants are proportional to N , which means that essentially the extent of the fluctuations that you have over here are going to go the order of the square root of N .

So the bridge, the thing that again allows us, while we have in principle in the expression that we have said, a variable energy for the system. In fact, in the limit of things becoming extensive, I know where that energy is, up to fluctuations or up to uncertainty that is only of the order of square root of N . And so the relative uncertainty will vanish as the N goes to infinity, limit is approached.

So although again we have something that is in principle probabilistic, again, in the

thermodynamic sense we can identify uniquely an energy for our system as, let's say, the mean value or the most likely value. They're all the same thing of the order of $1/N$. And again, to be more precise, the variance is clearly the second derivative of $\log Z$. 1 derivative of a $\log Z$ is going to give me the energy.

So this is going to be $d/d\beta$ up to a minus sign of the energy or the expectation value of the Hamiltonian, which we identified as the energy of the system. The derivative with respect to β , I can write as $k_B T^2$. The derivative of energy with respect to T , everything here we are doing that conditions of no work.

So the variance is in fact $k_B T^2$, the heat capacity of the system. So the extent that these fluctuations squared is $k_B T^2$ times the heat capacity the system.

OK, so next time what we will do is we will calculate the results for the ideal gas. First thing, the canonical ensemble to show that we get exactly the same macroscopic and microscopic descriptions. And then we look at other ensembles. And that will conclude the segment that we have on statistical mechanics of non-interacting systems.