

[SQUEAKING]

[RUSTLING]

[CLICKING]

**GIAN PAOLO
BERETTA:**

OK. Good afternoon. Here we are March 19, lecture 13. We are still in the process of studying mixtures, the wonderful world of mixtures. And so let me just briefly see where we are. You remember that we discussed the properties of isothermobaric mixing, which are measurable properties that allow us to infer the chemical potentials thanks to this expression here.

We discussed the intermolecular forces between molecules and, for example, the fact that in the gas or ideal gas, regime, molecules spend most of their time far away from each other in spite of having an enormous amount of collisions per second, during which they do get close. But the time they spend close is small with respect to the average. And so since the properties are determined by the average, the effect of the intermolecular interactions is relatively negligible and that's the ideal gas regime. We will see more about that today from another point of view, which we will see in a moment.

We introduced a first ideal behavior model based on this idea of particles not interacting, which is the ideal Gibbs-Dalton mixture model in which, as you remember, the idea is that we represent the properties of the mixture as equivalent to the properties of a composite system in which the various components are separate. They occupy the same volume. They are the same temperature. And they are at the partial pressure.

According to this model, if each of the substances behaves as an ideal gas, so obeys the ideal gas pV equals RT equation of state, we found that the partial pressures add up to the pressure of the mixture. That's Dalton law, this here. And the partial pressures are given by the pressure of the mixture times the mole fraction. Plus, the chemical potential of the components in the mixture of each component is equal to this important expression that is the sum of the chemical potential of the pure component at the same temperature and pressure of the mixture plus RT , the log of a mole fraction.

We then proved also the Amagat law of additive volumes whereby an ideal gas mixture is such that you can also view it as a composite of separate pure gases, all at the same temperature and pressure but each occupying the so-called proper volume, which is given at the volume that they would occupy when pure in the same amount at the same temperature and pressure.

Then we discussed spontaneous mixing. What happens when we have separate gases kept separate by partitions that are movable so they are all at the same temperature and pressure. And without any external effect-- you see the weight is at the same level-- you just let them mix. You take the partitions away and let them mix. We found that the mixing under these conditions yields the final temperature and the final pressure equal to the initial ones. And the entropy increases by the so-called entropy of mixing, which is given by this term, minus the number of particles, total number of particles, the gas constant, and the sum of $y \log y$, so the mole fractions times the logarithm of the mole fractions.

We also noticed that we could have done that mixing in a better way, better from the thermodynamic point of view, in other words reversibly, in a weight process. And we computed how much work we could-- so therefore how much weight lifting we could get if we did that, indeed, reversibly. And here's the formula.

And the last thing we discussed was a pictorial idea of how that extraction of the adiabatic availability could be done if we have semi-permeable membranes. So here we picture, first of all, the idea is that when we have a gas confined in a volume by a partition and on the other side of the partition there is volume-- I mean, there is vacuum-- the pressure of the molecules here on the piston will tend to push it so as to increase the volume. And that will lift the weight if properly connected. Similarly, the black particles.

Now the idea here is that in an ideal Gibbs-Dalton mixture behavior, the particles essentially, even in a mixture, ignore each other. They live in the entire volume even as if the others weren't there. So for the gray particles to expand, to here whether there is the black particles or not, isn't really a difference. So that is the idea that we can achieve with these semi-permeable membranes.

This one allows the black particles to go through but not the gray. So the gray will push it to the right and pull up this weight. This other membrane is permeable to the gray particles but not the black so the black will push it leftward and pull up this other weight. OK.

You asked, do we have membranes in reality? And yes, we do. We can manufacture them. We engineer materials. These are some YouTube videos that you may want to google and watch. Not doing it here. And so you can, essentially, have a wall of material with holes. And if the holes are the proper size, they can allow the passage of some particles and prevent the other particles to pass. So that is the way we achieve a semi-permeable membrane.

In nature, also in biology, this is very important on the surface of cells where you may have a-- you may have active and passive transport. Active transport is a different story because it requires some energy for the particles to move. But also there is passive transport and the fact that membranes can be changed by the cell depending on its needs.

So for example, if the cell needs to eat, so to speak, it could open little holes so that the, say, oxygen comes in. If the cell needs to discharge maybe CO₂, then it needs to open a little bigger holes so that these molecules can go out. I'll have another reference about biological structures but I don't want to talk about them because I'm not at all expert on this. But so the world is full of membranes. And we will see more in a few moments.

Now there is one historical topic that is maybe interesting, just fun to discuss. It's called the Gibbs paradox. And we have the tools to now understand what it means and essentially to resolve it. The idea is this. We just mentioned what happens when we have particles that mix.

So here suppose we have just a binary mixture, only two types of particles, the gray and the black. And the spontaneous mixing occurs this way. And if you compute the entropy of mixing, this summation, when you have only two types of particles, becomes a sum of just two terms. And of course, the mole fractions when you have two particles are not independent because they must add up to 1. So the mole fraction of second type must be equal to 1 minus the mole fraction of the first type of particles.

In the particular case in which you have 50% of 1 and 50% of the other, so the mole fractions are 50% or $1/2$, this entropy of mixing becomes $nR \log$ of 2. Now if we consider, for example, a case in which I have one particle, if we express one particle in terms of moles, that's 1 over the Avogadro's number mole because in one mole you have these many particles, 6.0×10^{23} . Right? So nR in that case is equal to R over the Avogadro number. And this is called the Boltzmann constant which is one of the physical constants.

So if we suppose here we have just one particle on one side and one particle on the other side so the total n is 2, it's actually 2 divided by Avogadro's number. So the entropy of mixing is twice, the Boltzmann constant $\log$ of 2.

So what is the Gibbs paradox? The Gibbs paradox runs as follows. There is a saying that I know it in Latin. *Natura non facit saltum*.

It means nature doesn't make jumps. And this I think it goes back to Aristotle or maybe Leibniz. I mean every philosopher has touched on this idea. So what Gibbs-- I think the Gibbs, they call it-- I really don't know. Maybe we should look it up about the history of these things is interesting as a chapter in itself. So I'm not an expert in that either. But suppose that it was Gibbs who realized that, all right, suppose I'm coloring these particles with a dye.

And so suppose the amount of distance in color is measured by this parameter λ . So I say 0 means the particles are the same color. λ greater than 0 means that the particles are different color. All right. So regardless of the color, here we see that the result, this result, the entropy of mixing, is the same regardless of how different the particles are. There is no detail about what the particles or the mass is. It's just how many there are.

So the entropy of mixing is always given by this amount, $nR \log$ of 2, except-- and so you would say, well, as λ goes to 0, I would expect that to also go continuously. Instead, as you go to 0, that jumps to 0 because, if we think in terms of the simple-system model, you remember that when we have a pure substance, putting a partition or removing it doesn't change the properties. So that's one quick way, since we have seen a lot before, to show that if the particles are identical, removing the partition will not change any of the properties, including entropy.

So there is no entropy of mixing in this case. If you mix the same substance, there is no entropy of mixing. As soon as they become different, there is entropy of mixing and it is discontinuous. So that's why the philosophers didn't like this idea. But in fact, it all has to do with the fact that we are doing modeling of reality. So Nature is the way we model it.

So if we have the possibility to distinguish the particles and therefore they are different for us, then there is mixing. And if there is entropy of mixing, there is also the possibility to do work. As we have seen, we can do this work, this mixing, properly and extract some work. But if we don't have the possibility to recognize, to identify, to distinguish the particles, that work is impossible for us. So it has to do with the way we model it or, if you really want to go Nature, if the particles are truly indistinguishable in the sense of quantum mechanics, distinguishable/indistinguishable, fermion, bosons, and that kind of discussion, then there is no way you can distinguish.

There is no way you can extract any work. If the particles are distinguishable, then it all depends on your model, namely on your kind of tool that you are considering in your model. So that's essentially the resolution of the paradox. There's no paradox.

It's true like this. And the discontinuity in going from here to there is essentially embedded in our modeling or, from the practical point of view, our ability to control the differences in particles because if the particles are different but we have no ways to control their differences and therefore build semi-permeable membranes that can select, then obtaining that work is out of our reach. OK. So that's the Gibbs paradox.

And since we are at it, I want to take a little detour today also on some other topics that are important in statistics, in information theory, and also in quantum theory. So again, look at this box with particles, different particles. And suppose you make an experiment by having a little window someplace like a control volume somewhere in that container where you can look.

If there is a particle in there, there can be only one at a time. It's a small control volume. You write it down in your experimental notebook and say, oh yeah, I saw a molecule of nitrogen, or I saw a molecule of oxygen, or I saw a molecule of argon, or I saw a molecule of CO₂. If that's air, this is the idea. And if you are in a stable equilibrium state and therefore the probabilities that you observe one of these molecules is proportional to the fraction of particles overall so it's the mole fraction. So the mole fraction, which by the way is a number like probability. It's a number between 0 and 1 and it adds up to 1. So the mole fraction is the probability to get a particle of that type.

So there is a measure-- if you are facing this situation with probabilities, you are in an uncertainty situation. You are uncertain about what will be the outcome of the next observation. You cannot predict the next time that's what I will get. But you can predict, because the mole fractions, with what frequency you will get these results upon repeated measurement.

But when you are facing the next observation and you have a big question mark in your mind and ask, all right, what will be the next outcome, you are in a state of uncertainty. And the uncertainty can be measured by this measure of information, which is called the Shannon information entropy. Shannon, 1948, said, well, if I look at a situation described by probabilities like these events, and I want to find the most logical mathematical expression or function of these probabilities that expresses my state of uncertainty. He set up some postulates, some desiderata about this measurement function of probabilities. And the only function that satisfies those criteria is this one.

For example, see if you are-- that is equal to 0 if you have certainty. So if the probabilities are all zeros except one event that is equal to 1, so you have certainty. So the particles are only one type. The mole fraction is 1. OK. Log of 1 is 0. And $x \log x$ for those that are 0 is again 0 because you remember that these functions $p \log p$ look this way. Between 0 and 1 is the function $x \log x$. Right?

So in that case of complete certainty, the information entropy is 0, or on the opposite side, if all particles have equal probability so all the mole fractions are equal, and in that case, the entropy is equal to log of n of the number of-- log of r in this case. It's the number of different particles.

You can work it out. You substitute each probability is $1/r$ and that this summation gives $\log r$. That's the maximum, the most uncertain situation when they are all alike. If you have instead a situation like for the molecules of air, it's not even and therefore some are more frequent than others so you could guess some better than if everything was just 50%.

You can also view-- this is interesting also for us for the next step in the next definition of ideal solution. You can also view this as a weighted sum weighted on the mole fractions. So I can view this as a summation. The mole fraction times minus log of y . Now this term in parentheses, we could call it the surprise. The surprise.

The surprise is a measure. It's high for the least probable ones. So if I get a-- say that if I get-- if I see a molecule of CO₂ there and I know that there are only 500 PPMs, parts per million, so there is very low probability, so if I see that one, I'm more surprised than if I see a nitrogen or an oxygen.

Argon also is a bit of a surprise because argon is less than 1%. So I need to do 100 trials to have, on average, one argon. And here I need 2,000 trials to get one CO₂. So that's the surprise.

Now, in information theory, there is also this problem that is also relevant for measurements. For example, suppose that your device finds the particle and characterizes the particle by the mass. So when the particle is in your window, your device, maybe a spectrometer or whatever, it tells you, OK, it was a particle with this mass.

So in the end, when you do a measurement, like a repeated measurement procedure, in the end, you will find the average or the mean mass that shows up. And suppose that that's all that you can measure of your substance. So all you can measure is the average value of the mass.

If that's the only information I have about my system, could I make an educated guess of what the probability distribution, so therefore what the mole fractions are? And here comes information theory. And again, based on Shannon information entropy, you say, all right, in order to make the educated guess, I should avoid any bias that is not contained in the information I have about the system.

So I should use the information I have, the measured values, but not extra stuff. So I want to make the least biased guess. And the least biased guess is obtained by saying, all right, what is the probability distribution that maximizes the Shannon information? In other words, it leaves me with the highest possible uncertainty compatible with that mean value that I have observed. In this way, I have not built in the probability distribution, any extra information.

And so, mathematically, this means you have to maximize the Shannon entropy subject to the constraint that your mean value be the known or measured value. Of course, you also have to put the auxiliary constraint that probabilities must add up to 1. And if you solve this problem with the method of Lagrange multipliers-- I don't want to just go over that today, but maybe later on, since we will review it. In any case, I presume that you have seen it somewhere else.

So the method simply says, all right, if you have a constraint maximization, you must assign to each of the constraints a Lagrange multiplier, an undetermined multiplier, and then do an unconstrained maximization of the optimizing function minus the constraints multiplied by these multipliers. And then once you obtain the solution of the unconstrained maximum in terms of these multipliers, you substitute back into the constraints and get the values for the multipliers.

All right. That's a quick so I don't expect you to have learned it from these words. But in any case, if you do the job, here's the probability distribution that you get. And if you substitute this probability back into the constraint, this one that gives you the average M , it gives you a relation between the λ M , which is the Lagrange multiplier, and the mean value. So you get that the Lagrange multiplier depends on that mean value of the mass.

So this is the best or at least the least biased assignment of mole fractions that you can do upon a measurement of only the mass. We will see in a while that this probability distribution is very similar to what you have for the description of stable equilibrium states within statistical mechanics and quantum theory. It's called the canonical distribution if instead of mass you put energy. And that allows or allowed researchers like Jaynes-- it's a big name who made these very neat connections between information theory and statistical mechanics.

Or also, if you still go to the library here at MIT, there is a wonderful set of video lectures by Professor Myron Tribus back in, I think, the '60s or the '70s where he does the information theory point of view of thermodynamics. So he develops the entire thermodynamics just by building up on information theory concepts. And indeed there is a ground for doing it because the molecular view that we are discussing here and the tools developed by Shannon do provide that possibility.

So I said that I wanted to make a little digression today. So here is a digression in the digression. And it has to do with the treatment of a particle in a box. OK. So we take a box of sides l_1 , l_2 , and l_3 in the three directions. And we consider only one particle inside the box. That's the idea of this little survey with some equations that are borrowed from quantum mechanics, which you may be familiar or not.

But even if you are not, it's OK. This is just a quick view that tells you how all these things are related. So it might be interesting. I will not ask for these in the oral exam, of course. And actually, one day I will let you know which of the many things that we cover, the slides that are going to be part of the exam so that you can focus because otherwise it's too much, I perfectly realize.

All right. So the idea in quantum mechanics is that in order to describe a molecule like that, or even simpler, if you have a molecule on a one-dimensional box that we already mentioned-- so suppose we have just one-dimensional box. In classical mechanics, the particle is here and it goes back and forth. It's perfectly situated somewhere at every instant of time.

In quantum mechanics, the story is different. Particle cannot be-- is not here-- viewing the particle as localized is wrong. It gives you some results that are wrong. So the idea of quantum mechanics is that you start from-- you have to find a probability distribution which can be obtained by a function, let's say ψ . And this function is the solution of an eigenvalue problem.

And for example, first of all, this function can be also written as-- it factors. So it's a product of a function of x and a function of time. And the function of time, if I remember, it's minus $\hbar$ times the energy times time over something. It will come up somewhere-- over m probably. Not sure. So the idea is that this function here, the one that depends on x , is the solution of an eigenvalue problem.

And these functions are-- you can plot them as follows, as we know. Actually, I think we have the expression that's something like square root of 2 over l where l is the side-- so this is from 0 to l -- times sine of some j , some maybe x . And there are some other constants here that now slips my memory. It's maybe $\hbar$ over $2m$ or something like that or maybe also an l . There is also an l here.

So the idea is that this j is the integer numbers. So there are many possible contributions. Each one of these for each j corresponds to a solution and to that solution corresponds also an energy level. So for example, the lowest energy level corresponding to j is equal to 1 is a function like this. Then you may have the second energy level in which the function is something like this. And then j equals to 3 and so on.

Corresponding to each of these j 's, there are energies, or actually on the slide, I'm calling them ϵ_j . And these energies are $\frac{h^2}{8mL^2} j^2$. Maybe there's a π here because it needs to go to 0 at 0 and at L .

So OK. So that means that if this ϵ_1 is the lowest energy level, ϵ_2 is equal to-- that corresponds to j equals 2 so it equals-- j^2 is 4 times ϵ_1 . And ϵ_3 is 3 to the square so it's 9 times ϵ_1 and so on. So already this gives you the energy levels.

Now in an experiment that you do, for example, with spectroscopy on the particle, you can measure the energy of the particle. And you're facing a situation like the one that we've just done with the mass of the particles that you don't know which one.

OK. Here it's a similar situation. You do a measurement of energy. And you don't know which one of the possible energies you get because the particle is not definitely on this, not always. I mean, most of the times, it's in a superposition of energy levels. So each one of them has a probability. And the energy of the particle is the weighted sum with these probabilities of these energy levels.

So the state of the particles is represented by a probability distribution, which, at least in the stable equilibrium states, can be viewed as the probability of finding the energy levels among these possibilities. That is for the particle in one-dimensional box. If you have a three-dimensional box, each direction has its own contribution. And actually, these functions factor.

So the three orthogonal translational degrees of freedom of the particle are superposed in each other. So it's like having three kind of subparticles like in a perfect gas that ignore each other. Each degree of freedom lives by itself occupies its directional modes by itself without interacting with the other modes. That's why this is such a simple extension of that. Just multiply, so to speak, the various functions.

Not only, but from this one, from these results, you can also obtain the probability of finding-- the probability density of finding these particle if you make a position measurement, finding it between x and $x + dx$. That's the PDF, the probability density function. It's given by this ψ , the modulus of this one or also that one.

And I think there is an i here that I'm missing because this is supposed to be i is the square root of minus 1. OK. So that means that when I do the complex conjugate of this one, that i becomes minus i . And so the modulus of ψ is equal to-- this one cancels and you're left with this, which is the modulus of ψ squared. And that is the probability density function. So if you want to know how much is the probability in that interval, you multiply by dx .

And if you want to know the probability to go from, say, from here to there, you just do the integral from whatever x , say x_1 to x_2 , to get the overall distribution. So these functions are these ones that I've plotted. I wish I had more space. OK. Do it this way.

So the probabilities are the square of these functions. The wave functions can go negative. But the probability is essentially the square. So take the square of this sine and it becomes something like this. The square of this other sine, it becomes something like this and so on. And so on.

When the particle is in a given state, it is not typically in just one of these energy levels. It is a probability distribution. So each one of these distributions in space is weighted by a probability distribution. So there will be a p_1 . There will be a p_2 , p_3 , and so on. So you have a weighted sum of these probabilities with which you can compute, for example, the probability that your particle is between x and $x + dx$, it's this area times p_1 plus this area times p_2 plus that area times p_3 up to infinity because you have to go up. This j goes up to infinity. And that's how you compute the actual stable equilibrium state probability distribution.

We already mentioned the fact that when you do this job, you can compute therefore the probability density. For every x , you will find that it will always go to 0 at the edges because at the edges, no matter how you weight it, it's always 0. So the sum of 0 is 0 whereas, in the center, you will have some wiggles, especially if the low energy levels are important. As you go upwards, this becomes so you may have, if the quantum number goes-- as it goes higher, these wiggles become more intense and they add up. And so the one cancels the others. And so the 0, you'll get only 0 at the walls but not in the middle, not in the bulk.

OK. So these are the energy levels. Here is the name for the probabilities that we have simplified here. In three dimensions, you have three labels because you have one quantum number per direction. And the typical problem that you face if you do energy measurements is, well, how do I-- what's my best guess for this probability distribution? That's the information theory point of view.

The thermodynamics point of view is, all right, the probabilities if we accept that the probabilities represent the state, and if we accept that entropy is measured by a function like this, then we can use the maximum entropy principle which says, for a given energy, my system has an energy and entropy diagram.

And for a given energy, I may have many different states. But if I'm looking for the stable equilibrium state, I know that it has to have the maximum entropy for that given energy. So that's precisely the same problem that we have seen before. Maximize entropy provided you accept that formula for the entropy of the quantum particle. Then the mathematics of maximization is the same using Lagrange multipliers.

Now since we know already the answer, instead of calling the Lagrange multiplier λ , which since later on we recognize that λ is the inverse of the temperature divided by the Boltzmann constant, we already call it $1/kT$. And so you find this distribution. So the probabilities are given by $e^{-\text{energy}}$, these energy values, divided by-- or multiplied by the Lagrange multiplier divided by this Q .

You can view it as a normalization factor. It's essentially the summation of all these things at the numerator, this Q . I also see that since the energy eigenvalues add up. The exponential of a sum is equal to the product of the exponentials. So these probability distributions factor into separate probability distributions for each of the directions of the box. So you can view the problem also just as-- you can solve it in one direction per time and then simply multiply the results.

These denominators or denominator for the overall is called the partition function. Yes. It plays the role of the normalization factor so that when you sum the probabilities, they add up to 1. But it also plays an interesting role because, essentially, you can extract everything if you know this partition function. If you know this summation, by playing with it in terms of taking some derivatives, you can get everything of the distribution, including the probabilities themselves. So if you take the partial derivative of the logarithm of the partition function as written here and take the derivative with respect to the energy level, you get the probability and multiply by kT and minus, you get the probability of observing that energy level.

Or you take this other derivative. You take the derivative with respect to temperature because, yes, in a moment we'll recognize that this T is also-- this Lagrange multiplier, satisfies our definition of temperature-- is the partial of E with respect to S . And so you identify it with the temperature.

So if you take the log of-- the partial of log of Q with respect to T , you get the energy, the mean energy. The mean energy is the average value of the energy. So it's the weighted sum of these energy levels weighted on the probability that represents that state. And also, the entropy. Also the entropy because the probabilities factor, the entropies add, because the logarithm-- remember, the logarithm of the product is equal to the sum of the logarithms.

So you can write the overall entropy and split it into the entropies of the three directions of the box. And each one of them can be obtained as a proper derivative of Q or if you've already taken the derivative that gives you the energy, you can write S as E over T plus Boltzmann constant times the logarithm of the partition function. This is somehow as close as you can get to the Euler free energy or the free energy. In fact, some people call it-- you could also rewrite that as E_i minus TS_i is equal to $k_B T \log$ of Q , probably with a minus sign. Yeah. OK.

Now going back to these probabilities, this probability distribution is the one that characterizes the stable equilibrium state. It characterizes this state. How about these other ones? So this canonical distribution, the maximum entropy one, looks like this. Here, for simplicity, I have drawn a system which doesn't have this eigenvalue structure but it has energy levels that are equally spaced and there is only seven of them. The lowest one, 1, 2, 3, 4, 5, 6, 7. OK. Equally spaced.

We have already mentioned that for a system with an upper bound to the energy spectrum-- this needs an r here-- that should be spectrum-- the E versus S diagram doesn't go up indefinitely but it bends and returns. So that's the case.

So we have only seven energy levels. And we can picture these probabilities by a histogram. So this is the maximum entropy, the canonical distribution. For example, this is a zero entropy. A zero entropy state is one in which you are certain-- remember, the entropy in terms of probabilities represents your state of uncertainty. Well, if I know that every time I measure, I always find the particle in that energy level, I have no uncertainty. The entropy is zero. And that is when I have zeros for all probabilities except 1, which is equal to 1.

A point like here is one that has an intermediate probability between 0 and the maximum. You can concoct an infinite number of possibilities of different probabilities, probability distributions, that have the same entropy and, of course, the same energy. Here are the ones. OK.

Why is that important? When we are interested in describing non-equilibrium states, what is important is to be able to understand or model how these states or these probability distributions evolve towards equilibrium because Nature non facit saltum, but Nature usually likes to go towards equilibrium. And so there is a spontaneous tendency.

So there is a natural law that pushes a system to even out or redistribute the energy load. You can interpret these histograms-- the probabilities-- you see as written here-- measure the degree of involvement of the level e_j in sharing the mean energy load. You give some energy to the particle. And the particle has to decide where to put it. It can put it in translational mode. But as we've seen here, if the particle is confined, the translational modes have these kind of energy levels.

So how much energy do I put in this mode and how much in that and how much in that? If the particle also has an internal structure, so for example, it could vibrate, or if it has a structure so that it can rotate, it can put also, you get additional modes, rotational modes, and then you have electronic modes. So the particle has to decide how to distribute and distribute evenly.

So if it is unconstrained, the most even distribution is the maximum entropy, the equilibrium. At equilibrium, it will have distributed things so as to maximize the entropy. So entropy also measures the degree of sharing. The highest degree of sharing is achieved here. And the natural phenomena have in common the idea that there is a spontaneous tendency to share the energy as evenly as possible among the available energy levels.

So the question about dynamics is how do these things evolve in time. And here I have a slide with some animation. Hopefully, it should work. Let's see. Yeah. So for example, I started from very close to zero entropy. And this is a smooth transition in which the particle, while keeping the energy constant, redistributes the probabilities until it eventually achieves the stable equilibrium canonical distribution.

And also, this one is another example. Started from an intermediate state with two initially full energy levels and the other ones get filled until, in the end, you get to the same state. I rescued these viewgraphs from work that I had done for my PhD thesis here, 1981, to devise an equation of motion that would take any probability distribution of this sort and evolves it smoothly towards a stable equilibrium. And that's developed in the subject that I now call it steepest entropy ascent but also has to do with what I like to call the fourth law of thermodynamics because, essentially, this kind of dynamical description of non-equilibrium states is common of a large number of theories of non-equilibrium.

There are some features and there are some aspects, some general aspects of this modeling ideas, that are so common in spite of the different fields in which you find them, that you can say that perhaps they represent another law of-- shouldn't say of Nature, but another law of modeling Nature correctly, like the second law we have described as something that your model should not violate.

Otherwise, it is a bad model. Also, the fourth law will be something that is similarly to be satisfied if your model has to provide a compatible description of how states evolve towards equilibrium-- compatible with all we know about evidence of physical reality. Well, a little more because we haven't finished yet with the particle in the box.

And let me take it-- start again from the partition function Q of the one dimensional case, which is therefore the summation of $e^{-\beta e_j}$ to the energy levels over kT . And here are the energy levels substituted directly.

So you see now in this expression, what are the fixed parameters. I mean, some are physical constants-- Planck constant h , Boltzmann constant k_B . m is the mass of your particle. j is just the dummy variable in this summation. There are only two parameters in this formula-- T , which is the temperature, and l , which is the size of the box or the side of the box.

So Q is a function of two variables. That means that, as we like to do when we have functions, we take the differentials. So if I take the differential of the logarithm of this summation, logarithm of Q , and it takes a little while, but you can do it. You can check. It can be written this way. So of course, it has a dT and a dl because the two variables are T and l .

So once you have done it, you may say that the coefficient that multiplies dT is equal to the derivative of the log of Q with respect to T . And so you obtain this relation. And this other one is equal to the derivative of log of Q with respect to l . Actually, notice that since it's dl/l , this is also the derivative of the log of l . So you obtain that derivative of log of Q over log of l is equal to $E_{\text{sub } i} \text{ over } kT$.

Now, $E_{\text{sub } i}$, I don't find it in this slide, but it is $E_{\text{sub } i}$. If you go a couple of slides earlier, you find it. And you should correct me if I'm writing it wrong. $E_{\text{sub } i}$ should be the weighted sum of the probabilities of the energy levels of the single direction.

So essentially, it's the partial energy. You see, the overall energy is equal to the sum of the energies that the particle has in the three directions. It's like additive. They are independent modes, independent lives that the particle has in the three directions. So here we are watching it in just one direction.

Then, another formula that we had a couple of slides ago is this one, the one that I also wrote here for the free energy. So if I take the differential of this one, here is what I get.

Again, it takes a little bit of doing because I have to take the differential of log of Q . But the differential of log of Q , we just take it. We just took it. And so here it is. We substitute it there. This one cancels with the minus $1/T$ squared that comes from the derivative of this.

And so here is what I'm left with-- $dS \frac{1}{T} dE$ plus something dl . This is an entropy form of the fundamental relation for the stable equilibrium state. So this is the Gibbs relation. You like to put it in energy form. You just solve for dE . And here is it. It's dE is minus TdS minus something dl .

Now, if you remember, when we have a substance in the simple-system model limit, the Gibbs relation is that dE is TdS minus pdV . pdV , OK? Volume is the parameter, and pressure is the coefficient in the differential that measures the partial of the energy with respect to volume at constant S .

All right, so that prompts us by just looking at this. Here, we don't have volume because there is no volume for this particle. It is only the length of the box. The size of the box is the equivalent of the volume.

So we can define the equivalent of a pressure. I will call it-- I don't know what name to give, but I will call it directional pressure. It has dimensions of pressure. And I say-- and I define it as the negative of the partial of the energy with respect to volume but done while changing only the-- one direction.

So you see, since I have a box now this size, this shape, there are three different ways in which I can change the volume. One is to keep l_2 and l_3 fixed and make l_1 longer. And then there are the other two, OK? So I'm changing the volume just by changing l_1 . That's the pressure assigned to direction one.

How much the energy changes in the stable equilibrium at fixed S , fixed l_2 and l_3 when I change l_1 . So you can also rewrite it this way. And if you rewrite it this way, because the volume-- I mean, yeah, I'll let you think about this. This is very simple.

All right, so the partial of E with respect to the logarithm of l_1 is obtained here. It goes from this differential, this coefficient minus $2E$ is exactly the partial of E with respect to the logarithm of the dl/l is the logarithm of l . So here, what we find is that pressure is $2E/V$. That is the closest we can get for the moment without additional assumption. So we found that p_i -- let's say-- 1 times the volume-- is that the volume-- is equal to $2E_i$. This is p_i .

In a moment, if you add the information that eventually we will have that $E_{\text{sub } i}$ is equal to $1/2$ of kT , you substitute up there. You find that $p_i V$ equals-- the 2 cancels, equals kT . k , remember, is the gas constant $n R$. So you find that $p_i V$ equals nRT . That's the ideal gas equation of state.

It requires, though, this step, which we don't get yet. And it's not always. But at this stage, without assumptions about how large the temperature is, which is something that we require to discuss to get to this result, we have obtained essentially the equation of state. And of course, you have one such equation of state for each direction. You may have energy in the x , y , and z direction, and so you have these directional pressures. And so you have three such relations here.

You also find that since the energy sums-- I mean is additive, the energy of the three directions adds up to the overall energy of your particle, this relation between energy and pressure. So you can also be written as the sum of these three pressures times $V/2$. And from the fact that the energy sums, I can also take the-- if I take the differential of this, so dE is equal to the sum of the dE_i 's but the expressions for the dE_i 's is here.

So the Gibbs relations for the various directions are here. I can sum them up to get the overall Gibbs relation for my particle in the box. And so I have dE equal TdS . And instead of $p dV$, we have three contributions, one for each direction. So this is, in general, the Gibbs relation for the particle in the box. And it requires three pressures.

Now, we can make an observation that the partition function that we now recognize the role is a summation. So is a summation of various terms. You can think of a summation as the various terms. Consider them as the terms in a histogram. And then the summation is the area. So this is like the first term, say, from 0 to 1 and from here to 2 and 3 and so on.

And so the summation gives you essentially the area. And this is the value of the element in the sum. The idea here is that you sometimes can approximate this summation as the integral of a curve. And that approximation is good when the error that you make, these differences here become small. So this is the idea.

It requires that these things be small somehow. And that condition is written here. It means that this combination of terms must be smaller than 1 . This is a dimensionless number, which contains temperature and the square of the side. And it has to be small.

If this is small, you can approximate the summation with an integral. And then this is a kind of integral that can be done because it's an integral from 0 to infinity. And here is the result.

And this is the partition function for one direction. If you multiply the partition functions for the various directions, so you might apply the three. You get the partition function for the overall particle in the box. And you see here, I have an l_i . It's squared, but actually, it's to the $1/2$. So you can take it outside also.

And when you multiply the three of them, it becomes l_1 times l_2 times l_3 , which is the volume of the box. And so the partition function also here is Q to the $2/3$, but it's to the $3/2$. So you can take the V outside.

So essentially, this says that the overall partition function depends on the volume only. Essentially in this limit, you are losing the effect of the details of the shape of the box. What counts is just the product of the three, which is the volume.

In fact, in that limit, if you take the formulas we have up here for, say, for the energy $E_{sub i}$, you have to take the partial of the logarithm with respect to T . Well, here is Q . Take the logarithm with respect to T . It's just $1/2$ times kT . That's the result.

And therefore-- and it's equal in all directions. This is so-called equipartition theorem. The particle has decided in this limit to distribute equally the energy in the three translational modes. And each one gets a half of kT . Meaning that since the energy of the particle is the sum of the energies of the three directions, E is equal to $3/2$ of kT . This is a structureless particle. And so the specific heat, which is the partial of E with respect to T is $3/2$ of the Boltzmann constant.

Also, these pressures become all equal to one another. So if I take the Gibbs relation and these p_i 's are equal to one another, I can write it this way. And so I get-- so these three terms become kT , which is here, times the sum of the $d l_i / l_i$. This is the sum of the logarithm. So it's the derivative of the volume over the volume. It's d of $\log$ of V .

And so we have the Gibbs relation. But now, if we take the Gibbs relation, and remember that what multiplies dV is what we call pressure, here we find that, therefore, the pressure of the particle in the box is kT/V . So we get pV is equal to kT , which is the ideal gas law.

We also get, of course, the expression for the entropy, which can be written as a function of temperature and volume only. So we lost the information about-- the need to know the details of the box. So that means that in that limit-- and this limit, this condition that this is small can also be written this way, that the temperature, if you fix the volume, it means that the temperature must be big enough so that these constants divided by the volume to the $2/3$ -- so that the temperature is much greater than this number.

Or you could also put it differently. If you fix the temperature, you need the volume to be big enough. The system must be diluted enough. So in these limits, we find the ideal gas law. And we find that a box of the same volume, it could be very-- many different shapes can have the same volume. They all have the same properties.

That's why in most practical situations, since this is satisfied except when you go to very low temperatures, that's why we typically use just the volume as parameter because otherwise you would say, all right, you said at the beginning you need to put the parameters in the fundamental relation. But then you end up always putting only volume. Why? Well, here is the reason because for most practical situations, but not for small systems, volume is all that counts.

OK, so just a couple of minutes we are left. I want to introduce this idea, which is going to be our next step. Remember that we talked about the gas state as one in which the particles spend most of their time distant from one another, in spite of their colliding, OK?

So in the ideal gas-- Gibbs-Dalton mixture of ideal gases, we found that essentially the particles live in the space without seeing each other except when they collide. And that results to the fact that the entropy of mixing is the only effect that results upon mixing. You mix gases at a given spontaneous process. The temperature doesn't change. The pressure doesn't change. And the only thing that changes is the entropy. It's the mixing.

Now, we need to define also an ideal behavior for the liquid phase. Now, in liquid phase, the particles do see each other. Actually, they live most of the times close to one another, so close to where this intermolecular potential has a minimum. So for them, that is important.

However, we are going to define an ideal solution behavior, one in which the curves for the interaction between equal particles or different particles are all very similar. So if they are all very similar, the particle doesn't care what the neighbor is. In this limit, again, when you mix stuff, the only effect-- you don't get any energy of mixing or enthalpy of mixing. You don't get any change of volume of mixing because the particle doesn't want to change their preferred distances, which would imply a enthalpy of mixing or volume of mixing, which would imply changing the distance. Because essentially they like each other in the same way.

The only effect upon mixing is the randomness, so the entropy of mixing. And therefore, in this model that we define, we will assume that the entropy of mixing of the single particle is related to the surprise. And that is the only effect of this ideal situation that we're going to define. OK?

So with that, we've set the ground for the next lecture.