

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

PROFESSOR:

OK. Let's start. So last lecture, what we talked about was limitations of classical statistical mechanics, and what I will contrast with what I will talk about today, which is new version. The old version of quantum mechanics, which was based on the observation originally from Planck, and then expanded by Einstein, that for a harmonic oscillator, a frequency ω , the energies cannot take all values. But values that are multiples of the frequency of the oscillator and then some integer n .

What we did with this observation to extract thermal properties was to simply say that we will construct a partition function for the harmonic oscillator by summing over all of the states of e to the minus βe_n according to the formula given above. Just thinking that these are allowed states of this oscillator-- and this you can very easily do. It starts with the first term and then it's a algebraic series, which will give you that formula.

Now, if you are sitting at some temperature t , you say that the average energy that you have in your system, well, the formula that you have is minus the log z by the β , which if I apply to the z that I have about, essentially weighs each one of these by these Boltzmann weights by the corresponding energy and sums them.

What we do is we get, essentially, the contribution of the ground state. Actually, for all intents and purposes, we can ignore this, and, hence, this. But for completion, let's have them around. And then from what is in the denominator, if you take derivative with respect to β , you will get a factor of $\hbar \omega$. And then this factor of $1 - e^{-\beta \hbar \omega}$. Actually, we then additional $e^{-\beta \hbar \omega}$ in here.

Now, the thing that we really compare to was what happens if we were to take one

more derivative to see how much the heat capacity is that we have in the harmonic oscillator. So basically taking an average of the raw formula with respect to temperature, realizing that these betas are inverse temperatures. So derivatives with respect to t will be related to derivative with respect to β , except that I've will get an additional factor of 1 over $k_B t^2$.

So the whole thing I could write that as k_B and then I had the $\hbar \omega$ over $k_B t^2$ squared. And then from these factors, I had something like $e^{-\hbar \omega / k_B t}$ to the minus $e^{-\hbar \omega / k_B t}$ squared. E to the $\hbar \omega$ over $k_B t$ minus 1 squared.

So if we were to plug this function, the heat capacity in its natural unit that are this k_B , then as a function of temperature, we get behavior that we can actually express correctly in terms of the combination. You can see always we get temperature in units of k_B over $\hbar \omega$. So I can really plug this in the form of, say, $k_B t$ over $\hbar \omega$, which we call t to some characteristic temperature.

And the behavior that we have is that close to 0 temperatures, you go to 0 exponentially, because of essentially the ratio of these exponentials. We leave one exponential in the denominator. So the gaps that you have between n equals to 0 and n equals to 1 translates to behavior of that at low temperatures is exponentially decaying to leading order.

Then, eventually, at high temperatures, you get the classical result where you saturate to 1 . And so you will have a curve that has a shift from one behavior to another behavior. And the place where this transition occurs is when this combination is of the order of 1 . I'm not saying it's precisely 1 , but it's of the order of 1 . So, basically, you have this kind of behavior. OK?

So we use this curve to explain the heat capacity of diatomic gas, such as the gas in this room, and why at room temperature, we see a heat capacity in which it appears that the vibrational degrees of freedom are frozen, are not contributing anything. While at temperatures above the characteristic vibrational frequency, which for a gas is of the order of 10 to the 3 degrees K , you really get energy in the harmonic oscillator also in the vibrations. And the heat capacity jumps, because you have

another way of storing energy.

So the next thing that we asked was whether this describes also heat capacity of a solid. So basically, for the diatomic gas, you have two atoms that are bonded together into a molecule. And you consider the vibrations of that. You can regard the solid as a huge molecule with lots of atoms joined together. And they have vibrations.

And if you think about all of those vibrations giving you something that is similar to this, you would conclude that the heat capacity of a solid should also have this kind of behavior. Whereas we noted that in actuality, the heat capacity of a solid vanishes much more slowly at low temperatures. And the dependence at low temperatures is proportional to T^3 .

So at the end of last lecture, we gave an explanation for this, which I will repeat. Again, the picture is that the solid, like a huge molecule, has vibrational modes. But these vibrational modes cover a whole range of different frequencies.

And so if you ask, what are the frequencies ω of vibrations of a solid, the most natural way to characterize them is, in fact, in terms of a wave vector k that indicates a direction for the oscillatory wave that you set up in the material. And depending on k , you'll have different frequencies. And I said that, essentially, the longest wave length corresponding to $k = 0$ is taking the whole solid and translating it.

Again, thinking back about the oxygen molecule, the oxygen molecule, you have two coordinates. It's the relative coordinate that has the vibration. And you have a center of mass coordinate that has no energy. If you make a molecule more and more complicated, you will have more modes, but you will always have the 0 mode that corresponds to the translation. And that carries over all the way to the solid. So there is a mode that corresponds to translations-- and, in fact, rotations-- that would carry no energy, and corresponds, therefore, to 0 frequency.

And then if you start to make long wavelength oscillations, the frequency is going to

be small. And, indeed, what we know is that we tap on the solid and you create sound waves, which means that the low-frequency long wavelength modes have a dispersion relation in which ω is proportional to k . We can write that as ω is v times k , where v is the velocity of the sound in the solid.

Now, of course, the shortest wave length that you can have is related to the separation between the atoms in the solid. And so, basically, there's a limit to the range of k 's that you can put in your system. And this linear behavior is going to get modified once you get towards the edge of the Brillouin zone and you have to solve a big dynamical matrix in order to extract what the frequencies are, if you want to have the complete spectrum.

There are three different possible polarizations. So in principle, you will have three of these curves in the system. And these curves could be very complicated when you get to the edge of the Brillouin zone and you have to solve a big dynamical matrix in order to extract what the frequencies are, if you want to have the complete spectrum.

So the solid is a collection of these harmonic oscillators that are, in principle, very complicated. But we have the following. So I say, OK, I have all of these. And I want to calculate at a given temperature how much energy I have put in the solid. So this energy that I have put in the vibrations at some temperature T , assuming that these vibrations are really a collection of these oscillators.

Well, what I have to do is to add up all of these terms. There's going to be adding up all of the $\hbar \omega$ over $2s$ for all of these. OK? That will give me something that I will simply call ϵ_0 , because it doesn't depend on temperature. Presumably will exist at 0 temperature. And I can even fold into that whatever the value of the potential energy of the interactions between the particles is at 0 temperature.

What I'm interested in really is the temperature dependence. So I basically take the formula that I have over there, and sum over all of these oscillators. These oscillators are characterized by polarization and by the wave vector k . And then I have, essentially, $\hbar \omega \alpha$ of k divided by $e^{-\beta \hbar \omega \alpha}$ of k minus 1. So I have to apply that formula to this potentially very

complicated set of frequencies.

The thing is, that according to the picture that I have over here, to a 0 order approximation, you would say that the heat capacity is 1 if you are on this side, 0 if you're on that side. What distinguishes those two sides is whether the frequency in combination with temperature is less than or larger than 1. Basically, low frequencies would end up being here. High frequencies would end up being here. And would not contribute.

So for a given temperature, there is some borderline. That borderline would correspond to kt over h bar. So let me draw where that borderline is. kt over h bar. For a particular temperature, all of these modes are not really contributing. All of these modes are contributing. If my temperature is high enough, everything is contributing.

And the total number of oscillators is $3n$. It's the number of atoms. So essentially, I will get $3n$ times whatever formula I have over there. As I come further and further down, there's some kind of complicated behavior as I go through this spaghetti of modes. But when I get to low enough structures, then, again, things become simple, because I will only be sensitive to the modes that are described by this ω being v_k . OK?

So if I'm interested in t going to 0, means less than some characteristic temperature that we have to define shortly. So let's say, replace this with t less than some θ_d that I have to get for you shortly, then I will replace this with $e^{-\sum_{\alpha} \frac{h \bar{v}_{\alpha} k}{k_B t}}$. OK?

Now, for simplicity, essentially I have to do three different sums. All of them are the same up to having to use different values of v . Let's just for simplicity assume that all of the v_{α} s are the same v , so that I really have only one velocity. There's really no difficulty in generalizing this. So let's do this for simplicity of algebra.

So if I do that, then the sum over α will simply give me a factor of 3. There are three possible polarizations, so I put a 3 there. And then I have to do the summation

over k . Well, what does the summation over k mean? When I have a small molecule for the, let's say, three or four atoms, then I can enumerate what the different vibrational states are.

As I go to a large solid, I essentially have modes that are at each value of k , but, in reality, they are discrete. They are very, very, very, very finely separated by a separation that is of the order of 2π over the size of the system. So to ensure that eventually when you count all of the modes that you have here, you, again, end up to have of the order of n states.

So if that's the case, this sum, really I can replace with an integral, because going from one point to the next point does not make much difference. So I will have an integral over k . But I have to know how densely these things are. And in one direction it is 2π over l . So the density would be l over 2π .

If I look at all three directions, I have to multiply all of them. So I will get v divided by 2π cubed. So this is the usual density of states. And you go to description in terms of wave numbers, or, later on, in terms of momentums. And what we have here is this integral $\hbar v k e^{-\beta \hbar v k}$ minus 1. OK?

So let's simplify this a little bit further. I have $e^{-\beta \hbar v k}$. I have $3v$. The integrand only depends on the magnitude of k , so I can take advantage of that spherical symmetry and write this as $4\pi k^2 v k$ divided by this 8π cubed. What I can do is I can also introduce a factor of β here, multiplied by k . βk is 1. And if I call this combination to be x , then what I have is k being $x \hbar v$. Of course, k is simply related to x by k being $x \hbar v$.

And so at the next level of approximation, this $k^2 v k$ I will write in terms of $x^2 v x$. And so what do I have? I have $e^{-\beta \hbar v k}$. I have $3v$ divided by 2π squared. Because of this factor of k that I will take outside I have a k . I have a $k^2 v k$ that I want to replace with $x^2 v x$. And that will give me an additional factor of $v x$ over $\hbar v$ cubed. And then I have an integral $\int_0^\infty x^2 e^{-x} dx$ minus 1.

Now, in principle, when I start with this integration, I have a finite range for k , which presumably would translate into a finite range for x . But in reality none of these modes is contributing, so I could extend the range of integration all the way to infinity, and make very small error at low temperatures. And the advantage of that is that then this becomes a definite integral. Something that you can look up in tables. And its value is in fact $\pi^4/15$.

So substituting that over there, what do we have? We have that the energy is $e_0 + 3$ divided by 15, will give me 5, which turns the 2 into a 10. I have π^4 divided by π^2 , so there's a π^2 that will survive out here. I have a kt . I have kt over $\hbar^3 v^3$. And then I have a factor of volume. But volume is proportional to the number of particles that I have in the system times the size of my unit cell. Let's call that a^3 . So this I can write this as $I a^3$.

Why do I do that is because when I then take the derivative, I'd like to write the heat capacity per particle. So, indeed, if I now take the derivative, which is de by dt , the answer will be proportional to n and kv . The number of particles and this kv , which is the function, the unit of heat capacities.

The overall dependence is t to the fourth. So when I take derivatives, I will get $4t^3$. That 4 will change the 1 over 10 to $2\pi^2/5$. And then I have the combination $kv^3 \hbar^3$, and that factor of a^3 raised to the third power. So the whole thing is proportional to t^3 . And the coefficient I will call θ_D for [INAUDIBLE]. And θ_D I have calculated to be $\hbar^3 v^3 / a^3 k t$. No, $\hbar^3 v^3 / a^3 k t$.

So the heat capacity of the solid is going to be proportional, of course, to $n k_B$. But most importantly, is proportional to t^3 . And t^3 just came from this argument that I need low ω 's. And how many things I have at the ω . How many frequencies do I have that are vibrating?

The number of those frequencies is essentially the size of a cube in k space. So it goes like this-- maximum k^3 in three dimensions. In two dimensions, it will be squared and all of that. So it's very easy to figure out from this dispersion relation

what the low temperature behavior of the heat capacity has to be.

And you will see that this is, in fact, predictive, in that later on in the course, we will come an example of where the heat capacity of a liquid, which was helium, was observed to have this T^3 behavior based on that Landau immediately postulated that there should be a phonon-like dispersion inside that superfluid.

OK. So that's the story of the heat capacity of the solid. So we started with a molecule. We went from a molecule into an entire solid. The next step that what I'm going to do is I'm going to remove the solid and just keep the box. So essentially, the calculation that I did, if you think about it, corresponded to having some kind of a box, and having vibrational modes inside the box.

But let's imagine that it is an empty box. But we know that even in empty space we have light. So within an empty box, we can still have modes of the electromagnetic field. Modes of electromagnetic field, just like the modes of the solid, we can characterize by the direction along which oscillations travel.

And whereas for the atoms in the solid, they have displacement and the corresponding momentum for the electromagnetic field, you have the electric field. And its conjugate is the magnetic field. And these things will be oscillating to create for you a wave.

Except that, whereas for the solid, for each atom we had three possible directions, and therefore we had three branches, for this, since e and b have to be orthogonal to k , you really have only two polarizations.

But apart from that, the frequency spectrum is exactly the same as we would have for the solids at low temperature replacing v that we have with the speed of light.

And so you would say, OK. If I were to calculate the energy content that is inside the box, what I have to do is to sum over all of the modes and polarizations. Regarding each one of these as a harmonic oscillator, going through the system of quantizing according to this old quantum mechanics, the harmonic oscillators, I have to add up the energy content of each oscillator. And so what I have is this $\hbar \omega$ of k .

And then I have $\frac{1}{2} + 1$ over e to the $\beta \hbar \omega$ of k minus 1.

And then I can do exactly the kinds of things that I had before, replacing the sum over k with a v times an integral. So the whole thing would be, first of all, proportional to v , going from the sum over k to the integration over k . I would have to add all of these $\hbar \omega$ over $2s$. Has no temperature dependence, so let me just, again, call it some ϵ_0 . Actually, let's call it ϵ_0 , because it's more like an energy density.

And then I have the sum over all of the other modes. There's two polarisations. So as opposed to the three that I had before, I have two. I have, again, the integral over k of $4 \pi k^2 v k$ divided by $8 \pi^3$, which is part of this density of state calculation. I have, again, a factor of $\hbar \omega$. Now, I realize that my ω is ck . So I simply write it as $\hbar ck$. And then I have e to the $\beta \hbar ck$ minus 1.

So we will again allow this to go from 0 to infinity. And what do we get? We will get $v \epsilon_0$ plus, well, the 8 and 8 cancel. I have π over π^3 . So π over π^3 . So it will give me 1 over π^2 . I have one factor of kt . Again, when I introduce here a β and then multiply by kt , so that this dimension, this combination appears. Then I have, if I were to change variable and call this the new variable, I have factor of $k^2 dk$, which gives me, just as before over there, a factor of kt over $\hbar c^3$.

And then I have this integral left, which is the 0 to infinity $x^3 e^{-x}$ minus 1, which we stated is π^4 over 15.

So the part that is dependent on temperature, the energy content, just as in this case, scales as t to the fourth. There is one part that we have over here from all of the $0s$, which is, in fact, an infinity. And maybe there is some degree of worry about that. We didn't have to worry about that infinity in this case, because the number of modes that we had was, in reality, finite. So once we were to add up properly all of these 0 point energies for this, we would have gotten a finite number. It would have been large, but it would have been finite.

Whereas here, the difference is that there is really no upper cut-off. So this k here, for a solid, you have a minimum wavelength. You can't do things shorter than the separation of particles. But for light, you can have arbitrarily short wavelength, and that gives you this infinity over here. So typically, we ignore that. Maybe it is related to the cosmological constant, et cetera. But for our purposes, we are not going to focus on that at all. And the interesting part is this part, that proportional to t to the fourth.

There are two SOP calculations to this that I will just give part of the answer, because another part of the answer is something that you do in problem sets. One of them is that what we have here is an energy density. It's proportional to volume. And we have seen that energy densities are related to pressures.

So indeed, there is a corresponding pressure. That is, if you're at the temperature t , this collection of vibrating electromagnetic fields exerts a pressure on the walls of the container. This pressure is related to energy density. The factor of $1/3$ comes because of the dispersion relation. And you can show that in one of the problem sets. You know that already.

So that would say that you would have, essentially, something like some kind of p_0 . And then something that is proportional to t to the fourth. So I guess the correspondent coefficient here would be p squared divided by $45 k t$ over h bar c cubed. So there is radiation pressure that is proportional to temperature. The hotter you make this box, the more pressure it will get exerted from it.

There is, of course, again this infinity that you may worry about. But here the problem is less serious, because you would say that in reality, if I have the wall of the box, it is going to get pressure from both sides. And if there's an infinite pressure from both sides, they will cancel. So you don't have to worry about that.

But it turns out that, actually, you can measure the consequences of this pressure. And that occurs when rather than having one plate, you have two plates that there are some small separation apart. Then the modes of radiation that you can fit in

here because of the quantizations that you have, are different from the modes that you can have out here.

So that difference, even from the 0 point fluctuations-- the $\hbar \omega$ over $2s$ -- will give you a pressure that pushes these plates together. That's called a Casimir force, or Casimir pressure. And that was predicted by Casimir in 1950s, and was measured experimentally roughly 10 years ago to high precision, matching the formula that we had. So sometimes, these infinities have consequences that you have to worry about.

But that's also to indicate that there's kind of modern physics to this. But really it was the origin of quantum mechanics, because of the other aspect of the physics, which is imagine that again you have this box. I draw it now as an irregular box. And I open a hole of size a inside the box. And then the radiation that was inside at temperatures t will start to go out. So you have a hot box. You open a hole in it. And then the radiation starts to come out.

And so what you will have is a flux of radiation. Flux means that this is energy that is escaping per unit area and per unit time. So there's a flux, which is per area per time. It turns out that that flux-- and this is another factor, this factor of $1/3$ that I mentioned-- is related the energy density with a factor of $1/c$ over 4 . Essentially, clearly the velocity with which energy escaping is proportional to c . So you will get more radiation flux the larger c . The answer has to be proportional to c .

And it is what is inside that is escaping, so it has to be proportional to the energy density that you have inside, some kind of energy per unit volume. And the factor of $1/4$ is one of these geometric factors. Essentially, there's two factors of cosine of theta. And you have to do an average of cosine squared theta. And that will give you the additional $1/4$. OK?

But rather than looking-- so this would tell you that there is an energy that is streaming out. That is, the net value is proportional to t to the fourth. But more interestingly, we can ask what is the flux per wavelength? And so for that, I can just go back to the formula before I integrated over k , and ask what is the energy

density in each interval of k ?

And so what I have to do is to just go and look at the formula that I have prior to doing the integration over k . Multiply it by c over 4. What do I have? I have 8π divided by 8π cubed. I have a factor of k squared from the density of states. I have this factor of $\hbar c k$ divided by e to the beta $\hbar c k$ minus 1.

So there's no analogue of this, because I am not doing the integration over k . So we can simplify some of these factors up front. But really, the story is how does this quantity look as the function of wave number, which is the inverse of wave length, if you like. And what we see is that when k goes to 0, essentially, this factor into the beta $\hbar c k$ I have to expand to lowest order.

I will get beta $\hbar c k$, because the 1 disappears. $\hbar c k$ is cancelled, so the answer is going to be proportional to inverse beta. It's going to be proportional to kt and k squared. So, essentially, the low k behavior part of this is proportional to k squared c , of course, and kt .

However, when I go to the high k numbers, the exponential will kill things off. So the large k part of this is going to be exponentially small. And, actually, the curve will look something like this, therefore. It will have a maximum around the k , which presumably is of the order of kt over $\hbar c$.

So basically, the hotter you have, this will move to the right. The wavelengths will become shorter. And, essentially, that's the origin of the fact that can you heat some kind of material, it will start to emit radiation. And the radiation will be peaked at some frequency that is related to its temperature.

Now, if we didn't have this quantization effect, if \hbar went to 0, then what would happen is that this k squared kt would continue forever. OK? Essentially, you would have in each one of these modes of radiation, classically, you would put a kt of energy. And since you could have arbitrarily short wavelengths, you would have infinite energy at shorter and shorter wavelengths. And you would have this ultraviolet catastrophe.

Of course, the shape of this curve was experimentally known towards the end of the 19th century. And so that was the basis of thinking about it, and fitting an exponential to the end, and eventually deducing that this quantization of the oscillators would potentially give you the reason for this to happen.

Now, the way that I have described it, I focused on having a cavity and opening the cavity, and having the energy go out. Of course, the experiments for black body are not done on cavities. They're done on some piece of metal or some other thing that you heat up. And then you can look at the spectrum of the radiation. And so, again, there is some universality in this, that it is not so sensitive to the properties of the material, although there are some emissivity and other factors that multiply the final result.

So the final result, in fact, would say that if I were to integrate over frequencies, the total radiation flux, which would be c over 4 times the energy density total, is going to be proportional to temperature to the fourth power. And this constant in front is the Stefan-Boltzmann, which has some particular value that you can look up, units of watts per area per degrees Kelvin.

So this perspective is rather macroscopic. The radiated energy is proportional to the surface area. If you make things that are small, and the wavelengths that you're looking at over here become compatible to the size of the object, these formulas break down.

And again, go forward about 150 years or so, there is ongoing research-- I guess more 200 years-- ongoing research on-- no, 100 and something-- ongoing research on how these classical laws of radiation are modified when you're dealing with objects that are small compared to the wavelengths that are emitted, etc. Any questions?

So the next part of the story is why did you do all of this? It works, but what is the justification? In that I said there was the old quantum mechanics. But really, we want to have statements about quantum systems that are not harmonic oscillators.

And we want to be able to understand actually what the underlying basis is in the same way that they understand how we were doing things for classical statistical mechanics. And so really, we want to look at how to make the transition from classical to quantum statistical mechanics.

So for that, let's go and remind us. Actually, so basically the question is something like this-- what does this partition function mean? I'm calculating things as if I have these states that are the energy levels. And the probabilities are $e^{-\beta \epsilon_n}$. What does that mean? Classically, we knew the Boltzmann rates had something to do with the probability of finding a particle with a particular position and momentum. So what is the analogous thing here?

And you know that in new quantum mechanics, the interpretation of many things is probabilistic. And in statistical mechanics, even classically we had a probabilistic interpretation. So presumably, we want to build a probabilistic theory on top of another probabilistic theory. So how do we go about understanding precisely what is happening over here?

So let's kind of remind ourselves of what we were doing in the original classical statistical mechanics, and try to see how we can make the corresponding calculations when things are quantum mechanical.

So, essentially, the probabilistic sense that we had in classical statistical mechanics was to assign probabilities for micro states, given that we had some knowledge of the macro state. So the classical microstate μ was a point which was a collection of p 's and q 's in phase space.

So what is a quantum microstate? OK. So here, I'm just going to jump several decades ahead, and just write the answer. And I'm going to do it in somewhat of a more axiomatic way, because it's not up to me to introduce quantum mechanics. I assume that you know it already. Just a perspective that I'm going to take. So the quantum microstate is a complex unit vector in Hilbert space. OK?

So for any vector space, we can choose a set of unit vectors that form an

orthonormal basis. And I'm going to use this bra-ket notation. And so our ψ , which is a vector in this space, can be written in terms of its components by pointing to the different directions in this space, and components that I will indicate by this ψ_n .

And these are complex. And I will use the notation that ψ_n is the complex conjugate of ψ_n . And the norm of this I'm going to indicate by $\psi \psi$, which is obtained by summing over all $n \psi_n \psi_n^*$, which is essentially the magnitude of $n \psi$ squared. And these are unit vectors. So all of these states are normalized such that $\psi \psi$ is equal to 1.

Yes.

AUDIENCE: You're not allowing particle numbers to vary, are you?

PROFESSOR: At this stage, no. Later on, when we do the grand canonical, we will change our Hilbert space. OK?

So that's one concept. The other concept, classically, we measure things. So we have classical observable. And these are functions all of which depend on this p and q in phase space. So basically, there's the phase space. We can have some particular function, such as the kinetic energy-- $\sum_i p_i^2 / 2m$ -- that's an example of an observable.

Kinetic energy, potential energy, anything that we like, you can classically write a sum function that you want to evaluate in phase space, given that you are at some particular point in phase space, the state of your system, you can evaluate what that is.

Now in quantum mechanics, observables are operators, or matrices, if you like, in this vector space. OK? So among the various observables, certainly, are things like the position and the momentum of the particle. So there are presumably matrices that correspond to position and momentum.

And for that, we look at some other properties that this classical systems have. We had defined classically a Poisson bracket, which was a sum over all α d_α by d_α

$q_{\alpha} \delta_{\beta} - p_{\alpha} \delta_{\beta} - p_{\alpha} \delta_{\beta} - q_{\alpha} \delta_{\beta}$. OK?

And this is an operation that you would like to, and happens to, carry over in some sense into quantum mechanics. But one of the consequence of this is you can check if I pick a particular momentum, key, and a particular coordinate, q , and put it over here, most of the time I will get 0, unless the alphas match exactly the p q 's that I have up there. And if you go through this whole thing, I will get something like that is like a δ_{ij} .

OK. So this structure somehow continues in quantum mechanics, in this sense that the matrices that correspond to p and q satisfy the condition that p_i and q_j , thinking of two matrices, and this is the commutator, so this is $p_i q_j - q_j p_i$ is \hbar over i δ_{ij} .

So once you have the matrices that correspond to p and q , you can take any function of p and q that you had over here, and then replace the p 's and q 's that appear in, let's say, a series expansion, or an expansion of this o in powers of p and q , with corresponding matrices \hat{p} and \hat{q} . And that way, you will construct a corresponding operator.

There is one subtlety that you've probably encountered, in that there is some symmetrization that you have to do before you can make this replacement.

OK. So what does it mean? In classical theory, if something is observable the answer that you get is a number. Right? You can calculate what the kinetic energy is. In quantum mechanics, what does it mean that observable is a matrix?

The statement is that observables don't have definite values, but the expectation value of a particular observable o in some state ψ_i is given by $\psi_i o \psi_i$. Essentially, you take the vector that correspond to the state, multiply the matrix on it, and then sandwich it with the conjugate of the vector, and that will give you your state.

So in terms of elements of some particular basis, you would write this as a sum over n and m . $\psi_n o_m \psi_m$. And in that particular basis, your operator would have these matrix elements.

Now again, another property, if you're measuring something that is observable, is presumably you will get a number that is real. That is, you expect this to be the same thing as its complex conjugate. And if you follow this condition, you will see that that reality implies that n_{om} should be m_{on} complex conjugate, which is typically written as the matrix being its Hermitian conjugate, or being Hermitian. So all observables in quantum mechanics would correspond to Hermitian operators or matrices.

OK. There's one other piece, and then we can forget about axioms. We have a classical time evolution. We know that the particular point in the classical phase space changes as a function of time, such that \dot{q}_i is plus $d h$ by $d p_i$. \dot{p}_i is minus $d h$ by $d q_i$. By the way, both of these can be written as q_i and h Poisson bracket and p_i and h Poisson bracket.

But there is a particular function observable, h , the Hamiltonian, that derives the classical evolution. And when we go to quantum evolution, this vector that we have in Hilbert space evolves according to $i \hbar d$ by dt of the vector ψ is the matrix that we have acting on ψ .

OK. Fine. So these are the basics that we need. Now we can go and do statistical descriptions. So the main element that we had in constructing statistical descriptions was deal with a macrostate.

We said that if I'm interested in thinking about the properties of one cubic meter of gas at standard temperature and pressure, I'm not thinking about a particular point in phase space, because different gases that have exactly the same macroscopic properties would correspond to many, many different possible points in this phase space that are changing as a function of time.

So rather than thinking about a single microstate, we talked about an ensemble. And this ensemble had a whole bunch of possible microstates. In the simplest prescription, maybe we said they were all equally likely. But, actually, we could even assign some kind of probability to them.

And we want to know what to do with this, because then what happened was from this description, we then constructed a density which was, again, some kind of a probability in phase space. And we looked at its time evolution. We looked at the averages and all kinds of things in terms of this density.

So the question is, what happens to all of this when we go to quantum descriptions? OK. You can follow a lot of that. We can, again, take the example of the one cubic meter of gas at standard temperature and pressure. But the rather than describing the state of the system classically, I can try to describe it quantum mechanically. Presumably the quantum mechanical description at some limit becomes equivalent to the classical description.

So I will have an ensemble of states. I don't know which one of them I am in. I have lots of boxes. They would correspond to different microstates, presumably. And this has, actually, a word that is used more in the quantum context. I guess one could use it in the classical context.

It's called a mixed state. A pure state is one you know exactly. Mixed state is, well, like the gas I tell you. I tell you only the macroscopic information, you don't know much about microscopically what it is. If these are possibilities, and not knowing those possibilities, you can say that it's a mixture of all these states.

OK. Now, what would I use, classically, a density for? What I could do is I could calculate the average of some observable, classically, in this ensemble. And what I would do is I would integrate over the entirety of the six n-dimensional phase space the ρ at some particular point in phase space and the density at that point in phase space. And this average I will indicate by a bar. So my bars stand for ensemble average, to make them distinct from these quantum averages that I will indicate with the Bra-Kets. OK?

So let's try to do the analogue of that in quantum states. I would say that, OK, for a particular one of the members of this ensemble, I can calculate what the expectation value is. This is the expectation value that corresponds to this observable o , if I was in a pure state ψ_α . But I don't know that I am there. I

have a probability, so I do a summation over all of these states. And I will call that expectation value ensemble average. So that's how things are defined.

Let's look at this in some particular basis. I would write this as a sum over α $\langle \alpha | \rho | \alpha \rangle$. So, essentially, writing all of these ρ 's in terms of their components, just as I had done above.

OK. Now what I want to do is to reorder this. Do the summation over n and m first, the summation over α last. So what do I have? I have $\langle n | \rho | m \rangle$. And then I have a sum over α of $\langle \alpha | n \rangle \langle m | \alpha \rangle$. So what I will do, this quantity that α is summed over-- so it depends on the two indices n and m . I can give it a name. I can call it ρ_{nm} .

If I do that, then this \bar{o} average becomes simply-- let's see. This summation over n gives me the matrix product ρo . And then summation over m gives me the trace of the product. So this is the trace of ρo . OK?

So I constructed something that is kind of analogous to the classical use of the density in phase space. So you would multiply the density and the thing that you wanted to calculate the observable. And the ensemble average is obtained as a kind of summing over all possible what values of that product in phase space.

So here, I'm doing something similar. I'm multiplying this ρ by some matrix row. So, again, this I can think of as having introduced a new matrix for an operator. And this is the density matrix.

And if I basically ignore, or write it in basis-independent form, it is obtained by summing over all α 's the $\langle \alpha | \rho | \alpha \rangle$, and, essentially, cutting off the n and m . I have the matrix that I would form out of state α by, essentially, taking the vector and its conjugate and multiplying rows and columns together to make a matrix. And then multiplying or averaging that matrix over all possible values of the ensemble-- elements of the ensemble-- would give me this density.

So in the same way that any observable in classical mechanics goes over to an

operator in quantum mechanics, we find that we have another function in phase space-- this density. This density goes over to the matrix or an operator that is given by this formula here.

It is useful to enumerate some properties of the density matrix. First of all, the density matrix is positive definite. What does that mean? It means that if you take the density matrix, multiply it by any state on the right and the left to construct a number, this number will be positive, because if I apply it to the formula that I have over there, this is simply sum over α p_α .

Then I have $\langle \psi | \rho | \psi \rangle$, and then $\langle \psi | \rho | \psi \rangle^*$, which is its complex conjugate. So I get the norm of that product, which is positive. All of the p_α are positive probabilities. So this is certainly something that is positive.

We said that anything that makes sense in quantum mechanics should be Hermitian. And it is easy to check. That if I take this operator ρ and do the complex conjugate, essentially what happens is that I have to take sum over α . Complex conjugate of p_α is p_α itself. Probabilities are real numbers. If I take $\langle \psi | \rho | \psi \rangle$ and conjugate it, essentially I take this and put it here. And I take that and put it over there. And I get the same thing. So it's the same thing as ρ .

And, finally, there's a normalization. If, for my $\langle \psi | \rho | \psi \rangle$ over here in the last formula, I choose 1, then I get the expectation value of 1 has to be the trace of ρ . And we can check that the trace of ρ , essentially, is obtained by summing over all α p_α , and the dot product of the two ψ_α . Since any state in quantum mechanics corresponds to a unit vector, this is 1. So I get a sum over α of p_α . And these are probabilities assigned to the members of the ensemble. They have to add up to 1. And so this is like this.

So the quantity that we were looking at, and built, essentially, all of our later classical statistical mechanics, on is this density. Density was a probability in phase space. Now, when you go to quantum mechanics, we don't have phase space. We have Hilbert space. We already have a probabilistic theory.

Turns out that this function, which was the probability in phase space classically, gets promoted to this matrix, the density matrix, that has, once you take traces and do all kinds of things, the kinds of properties that you would expect the probability to have classically. But it's not really probability in the usual sense. It's a matrix.

OK. There is one other thing element of this to go through, which is that classically, we said that, OK, I pick a set of states. They correspond to some density. But the microstates are changing as a function of time. So the density was changing as a function of time. And we had Liouville's theorem, which stated that $d\rho/dt$ was the Poisson bracket of the Hamiltonian with ρ .

So we can quantum mechanically ask, what happens to our density matrix? So we have a matrix ρ . I can ask, what is the time derivative of that matrix? And, actually, I will insert the $i\hbar$ here, because I anticipate that, essentially, ρ having that form, what I will have is sum over α .

And then I have $i\hbar d/dt$ acting on these $p_\alpha \psi_\alpha \psi_\alpha^*$. So there ρ is sum over $\alpha p_\alpha \psi_\alpha \psi_\alpha^*$. Sum over αp_α I can take outside. $i\hbar d/dt$ acts on these two ψ s that are appearing a complex conjugates.

So it can either, d/dt , act on one or the other. So I can write this as sum over $\alpha p_\alpha i\hbar d/dt$ acting on $\psi_\alpha \psi_\alpha^*$, or i , or ψ_α^* , and then $i\hbar d/dt$ acting on this ψ_α .

Now, $i\hbar d/dt \psi_\alpha$, we said that, essentially, the quantum rule for time evolution is $i\hbar d/dt$ of the state will give you is governed by \hat{H} times acting on ψ_α . If I were to take the complex conjugate of this expression, what I would get is minus $i\hbar d/dt$ acting on ψ_α^* that is pointing the other way of our complex conjugation is \hat{H} acting on the ψ in the opposite way. So this thing is minus ψ_α^* with \hat{H} acting on it.

OK. So then I can write the whole thing as \hat{H} -- for the first term take the \hat{H} out front. I have a sum over $\alpha p_\alpha \psi_\alpha \psi_\alpha^*$, minus, from this complex

conjugation-- here, h is completely to the right-- I have a sum over α p α ψ α ψ α . And then we have h .

Now, these are again getting ρ back. So what I have established is that $i \hbar$, the time derivative of this density matrix, is simply the commutator of the operators h and ρ .

So what we had up here was the classical Liouville theorem, relating the time derivative of the density in phase space to the Poisson bracket with h . What we have here is the quantum version, where the time derivative of this density matrix is the commutator of ρ with h .

Now we are done, because what did we use this Liouville for? We used it to deduce that if I have things that are not changing as a function of time, I have equilibrium systems, where the density is invariant. It's the same. Then ρ of equilibrium not changing as a function of time can be achieved by simply making it a function of h . And, more precisely, h and conserved quantities that have 0 Poisson bracket with h .

How can I make the quantum density matrix to be invariant of time? All I need to do is to ensure that the Poisson bracket of that density with the Hamiltonian is 0. Not the Poisson bracket, the commutator. Clearly, the commutator of h with itself is 0. h minus h is 0. So this I can make a function of h , and any other kind of quantity also that has 0 commutator with h .

So, essentially, the quantum version also applies. That is, the quantum version of ρ equilibrium, I can make it by constructing something that depends on the Hilbert space through the dependence of the Hamiltonian and on the Hilbert space and any other conserved quantities that have 0 commutator also with the Hamiltonian.

So now, what we will do next time is we can pick and choose whatever ρ equilibrium we had before. Canonical $e^{-\beta h}$. We make this matrix to be $e^{-\beta h}$. Uniform anything, we can just carry over whatever functional dependence we had here to here. And we are ensure to have something that is quantum mechanically invariant. And we will then interpret what the various

quantities calculated through that density matrix and the formulas that we described actually. OK?