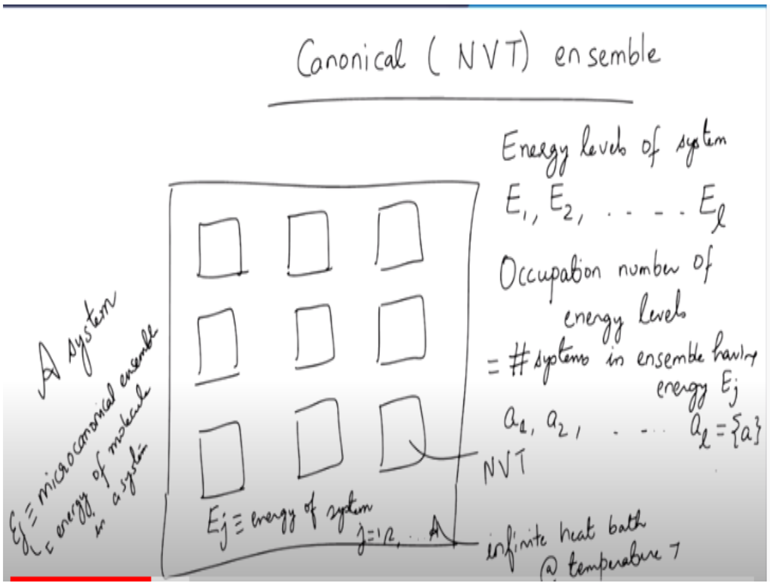
Advanced Thermodynamics and Molecular Simulations
Prof. Prateek Kumar Jha
Department of Chemical Engineering
Indian Institute of Technology-Roorkee

Lecture - 14
Canonical Ensemble: Most Probable Distribution, Partition Function

Hello, all of you. So in the last lecture, we have discussed the idea of thermodynamic ensembles and taken the example of micro canonical ensemble NVE. And I also discussed like what are the other types of ensemble. So today, I will do the discussion on the most common ensemble or most practical ensemble that is called the canonical ensemble or the NVT ensemble and I will show you that how can we compute thermodynamic properties using this idea of ensembles.

(Refer Slide Time: 00:59)



So as we discussed in the last class a canonical ensemble has the number of molecules the volume and the temperature as the control variables and therefore, we can imagine them, the ensemble as composed of systems which are having a constant number of molecules constant volume and constant temperature, right. So when I say I can control this that means, I can do the experiment at any possible value of number, volume and temperature. So all these systems in the ensemble are having the same number, volume and temperature in total there are some ‘A’ systems, right.

Now to visualize this one can imagine that you have this ‘A’ systems contained within some infinite heat bath this is simply a theoretical construct to imagine like how this can be realized this is not something that is practically done, right. You can imagine that you have infinite bath, so that the system is very small in comparison to the heat bath and temperature of the system is maintained by the surrounding, that is the heat bath and we also assume that there is no interaction among the systems. Systems are far separated from each other and the temperature of the system is maintained by using that heat path, right.

In reality, in a practical scenario, what we will do is like, we will take this systems as different states at different locations of the system or it can be imagined as a particular state, but at different time instants, right. The key idea is that the number, volume and temperature has to be constant but the configuration of the molecules, their positions and momentum and the distribution of energy among the molecules can vary between different systems in the ensemble.

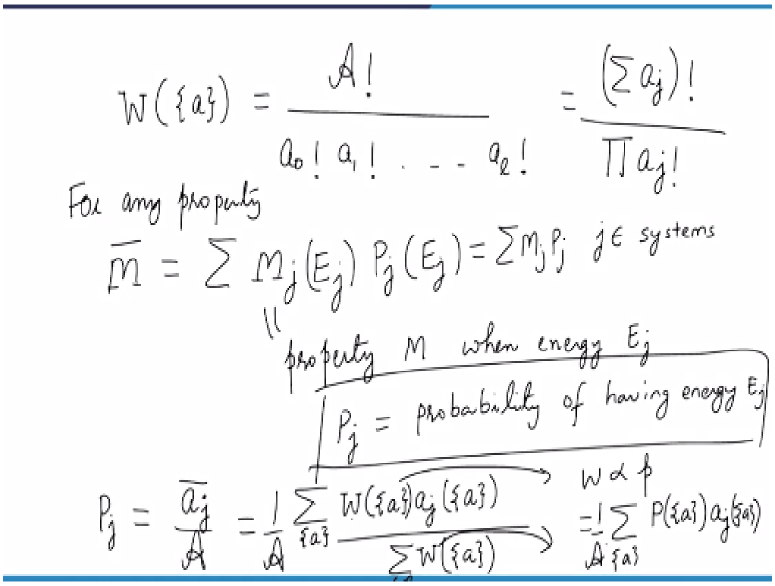
So since in this case NVT is being controlled, the other variables may change. In this case, unlike the micro canonical system, the energy of a system in the ensemble is not same for different systems of the ensemble but now we are referring to energy of the system. When I was doing the micro canonical ensemble I referred the energy as ϵj. But that energy was the energy of molecule within a system of the ensemble. Now I am referring to the energy of system as a whole right, energy of a system in the ensemble. So in this case, the index j is going from 1, 2 to A and in this case, the j went from the possible energy state within a system.

So in the micro canonical ensemble the energy of all the systems were same, equal to E that is the total energy of system was same. In this case, the total energy or the internal energy is different between systems in the ensemble. So now just like we had different states of energy of a molecule in the micro canonical ensemble, we can have different possible energy of the system in the canonical ensemble. So just like we had energy levels for the molecules, now we can define energy levels for a system as a whole. The key point here is since the energy of a molecule is discretized so therefore the energy of the system is also a discretized although, as you increase the number of molecules, the discretization becomes very unimportant you can assume as a continuous energy spectra, but nonetheless, the energy is discretized that is the key point that I am borrowing here. So now we will talk about energy level spacing. So that refers to energy of the system containing molecules not the energy of molecules themselves that is true for that canonical example.

So now I can say you can have many possible energies, let us say E_1, E_2 to some E_l these are the energy levels of system and if you start with some A systems in the ensemble, we assume that A is a very large number. So for those systems, some of the systems will have energy E_1, some will have energy E_2 and so on, right. So keep in mind that now we are talking about number of systems having that energy instead of the molecules. So then we can write as the occupation number of energy levels, the occupation number means, how many states has that energy so number of systems in ensemble having energy let us say E_j and that I can define as a_1 for energy level E_1, a_2 to some al right. And as I enumerate the systems in the ensemble, then just like the energy levels, the distribution of energy levels can also change. So in total, we are talking about some particular distribution of energy level, let me call that {a}, so that a can change between systems in the ensemble. So I am not even coming to the most probable distribution yet. So we can pretty much have the ensemble comprised of all possible distribution the only criteria we are putting is the number of molecules, volume and temperature are the same. So in that case, there is a whole range of possibility of the occupation numbers or the number of system in a given energy level and that distribution I refer as {a}.

We can again think in terms of the example I was giving you of say three quanta among three particles the only difference here is that we are talking about three quanta among three systems, right. So it is a more microscopic view than compared to what we have used earlier. But then you will see some similarity in the math nonetheless. So just like we could distribute the energy levels in the molecule using certain number of ways, the same applies here, but for the systems.

(Refer Slide Time: 08:37)



$$W(\{a\}) = \frac{A!}{a_0!\, a_1! \, . \, . \, . \, a_l!} = \frac{(\sum a_j)!}{\prod a_j!}$$

For any property

$$\bar{M} = \sum M_j(E_j)\, P_j(E_j) = \sum M_j P_j \quad j \in \text{systems}$$

property M when energy E_j

P_j = probability of having energy E_j

$$P_j = \frac{\bar{a}_j}{A} = \frac{1}{A} \frac{\sum_{\{a\}} W(\{a\})\, a_j(\{a\})}{\sum W(\{a\})} \qquad W \propto P \qquad = \frac{1}{A} \sum_{\{a\}} P(\{a\})\, a_j(\{a\})$$

So we can now say that for a particular distribution {a}, the number of ways of doing that is something like-

$$W(\{a\}) = \frac{A!}{a_0!\, a_1! \ldots \ldots \ldots \ldots a_l!} = \frac{(\sum a_j)!}{\prod a_j!}$$

Where now the summation and product is over systems. So j is over systems in the ensemble and we are hoping that A is really so large that we are able to enumerate all possible ways again, it is a theoretical construct.

But it makes things easier to think of like if we enumerate all the possible state, then these are the number of ways this can possibly happen. So now if I am interested in asking the question, what is the average property of the system? So then the average property is the average over all these systems in the particular ensemble, right. So the way to think about it, let us say if I am interested in some property M and I want to find its average value that property can be for example pressure or energy. Whenever in thermodynamics, we talk about pressure or energy, it does not refer to a particular configuration, but an average over a large possible number of configurations and in this case, A is that large number of configuration. So basically, whenever we find an average property, here, it is average over all the possible states in the system,

So one way to define that is that I can find the energy, find the value of M for any property M, we can find the property M, when the energy of system is E_j. So for any property we have-

$$\bar{M} = \sum M_j(E_j)\, P_j(E_j) = \sum M_j P_j$$

So M_j is property M, when the energy is E_j and multiplied with the probability of that happening. So P_j is the probability of having energy E_j. . But keep in mind that even if I do not put E_j in the brackets, M_j refers to the property value when the energy is E_j. I am making an assumption here that it only depends on the energy of the system, right.

So for any given energy, just like what we had in the micro canonical ensemble, even for the same energy you can have many possible distributions of the energy within the molecule but in the canonical ensemble, I am not considering that. In a canonical ensemble I am assuming that the property value is the same for a given energy E_j of the system irrespective of how the energy is being distributed that will be valid for microscopic properties that are typically used in thermodynamics.

So the properties I am trying to do this for are the properties which I compute on an average, not the properties of individual molecules. So with this kind of an idea, the question is how can we find the probability of having the energy P_j right? And that is clearly given as-

$$P_j = \frac{a_j}{A}$$

Here, a_j divided by the total number of states in the ensemble.

So this number is going to change for every possible distribution, right. So a_j is not going to be constant, because you can imagine different possible distributions and different possible distributions will have different number of systems having energy E_j. So therefore, I am interested actually in the average value of the occupation of the energy level, E_j, right and how can I find that is I can basically list all the distributions. And for all the possible distributions, I can compute what is the a_j value for that distribution and divide by the all the possible distributions.

$$\{a\} = \frac{1}{A}\frac{W(\{a\})a_j(\{a\})}{\sum_{\{a_j\}} W(\{a\})}$$

So what I am doing now is I am not even referring to only the most probable distribution. W {a} is the number of ways of having a particular distribution. If I sum this over all the possible distributions that will give me the total number of ways energies can be distributed. Now out of this for W of a particular a distribution, you have a_j ways in which we are getting energy level of E_j. So if I multiply W of a particular distribution with a_j, it gives me the number of ways in which a_j is happening. Now if I sum this over all possible distribution, I am going to get the total number of ways for all possible distribution. So if basically the probability is the number of ways in which an outcome is possible, divided by the total number of ways. If this confuses you, you can imagine that the number of ways is proportional to the probability.

$$W \propto p$$

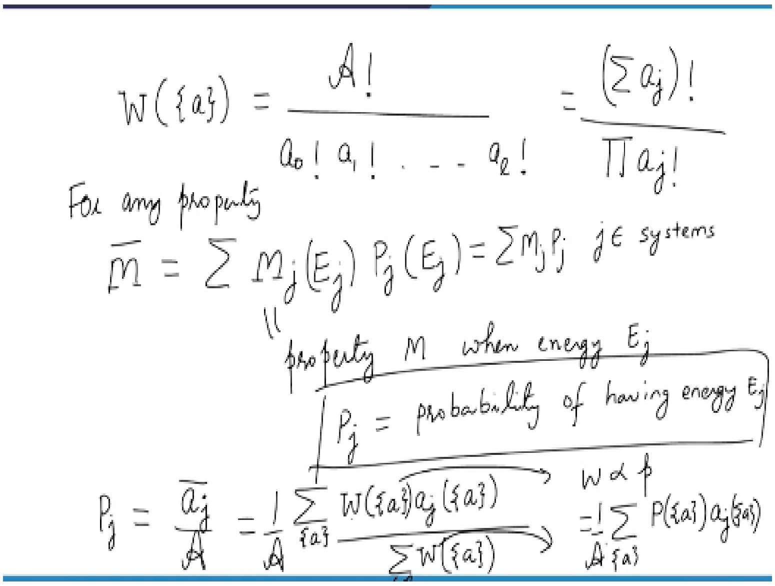
So whatever is appearing here, we can think of as the normalization of probability. So if I divide the W_a by sum over W_a, essentially what we are getting is the probability of getting a particular distribution {a} and I am multiplying that with the value of a_j for that particular distribution this is what we are doing.

$$P_j = \frac{1}{A}\sum_{\{a\}} P(\{a\}) a_j(\{a\})$$

So now clearly, if I start listing all the possible distribution, this becomes a very untraceable problem, right. It is very difficult to even think of all the possible ways of distributing something for a smaller number of quanta we can, but if we are thinking of a system as a whole, then the number of quanta are huge. And therefore, it is not quite easy or even practicable to think of all the possible distribution.

So we can simplify this problem by assuming that we are already in the most probable distribution and this is a great simplification that we can do, because in any case, we know that at equilibrium, the distribution should be close to the most probable distribution. we made that argument somewhat logically earlier, but we also demonstrated it for coin toss examples and so on. So even though other distributions are possible of the energy levels, the most probable distribution is the one that is only happening in most of the cases and therefore, instead of summing over all the possible distribution, we can sum over only the most probable distribution and that greatly simplifies the way we are doing it.

(Refer Slide Time: 08:37)



$$W(\{a\}) = \frac{A!}{a_0!\,a_1!\,\ldots\,a_l!} = \frac{(\sum a_j)!}{\prod a_j!}$$

For any property

$$\overline{M} = \sum M_j(E_j)\,P_j(E_j) = \sum M_j P_j \quad j \in \text{systems}$$

property M when energy E_j

P_j = probability of having energy E_j

$$P_j = \frac{\overline{a_j}}{A} = \frac{1}{A}\sum_{\{a\}} \frac{W(\{a\})a_j(\{a\})}{\sum W(\{a\})} = \frac{1}{A}\sum_{\{a\}} P(\{a\})a_j(\{a\})$$

$W \propto P$

So the assumption that we will make is we have the most probable distribution that is, I am looking at now a particular distribution, let me call that a*, which corresponds to some values a_1, a_2 to a_l of the occupation number.

$$\{a^*\} = (a_1^*, a_2^* \ldots\ldots\ldots\ldots\ldots\ldots\ldots..\, a_l^*)$$

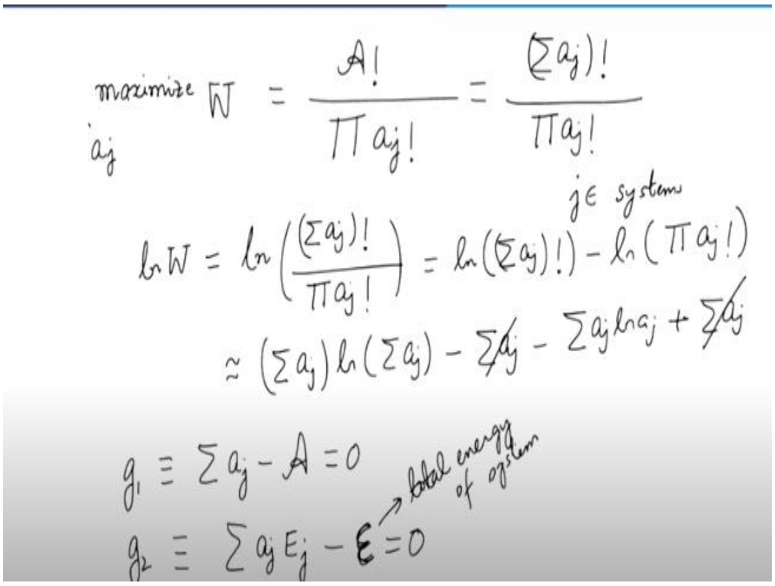
So now I am assuming that all the systems in the ensemble they can have different energies, but the distribution of the occupation numbers will follow the most probable distribution. So if I do that, then I can write-

$$P_j = \frac{\overline{a_j}}{A} = \frac{1}{A}\frac{W(\{a^*\})a_j(\{a_j^*\})}{W(\{a^*\})} = \frac{a_j(\{a^*\})}{A}$$

Now we are only doing it for the most probable distribution. So we can compute the number of ways for the most probable distribution and compute the a_j value for the most probable distribution.

So now if I do that, then the problem comes down to how do I find the a* for the most probable distribution and this we can do in the same way we did for the micro canonical ensemble with slight difference in the equations, because now I am talking about energy of systems instead of energy of the molecules.

(Refer Slide Time: 19:53)



$$\underset{a_j}{\text{maximize}}\ W = \frac{A!}{\prod a_j!} = \frac{(\sum a_j)!}{\prod a_j!}$$

$j \in$ system

$$\ln W = \ln\left(\frac{(\sum a_j)!}{\prod a_j!}\right) = \ln((\sum a_j)!) - \ln(\prod a_j!)$$

$$\approx (\sum a_j)\ln(\sum a_j) - \cancel{\sum a_j} - \sum a_j \ln a_j + \cancel{\sum a_j}$$

$$g_1 \equiv \sum a_j - A = 0$$

$$g_2 \equiv \sum a_j E_j - \mathcal{E} = 0$$

total energy of system

So the problem now is we need to maximize W and that is equal to-

$$maximise\ W = \frac{A!}{\prod a_j!}$$

Keep in mind that both the summation and the product is over different systems in the ensemble, which are assumed to be in the most probable distribution, that will come later after we do the minimization. So this we have to maximize for the a_j values for the most probable distribution. . So instead of maximizing this we can maximize ln of W and that is-

$$\ln W = \ln\left(\frac{(\sum a_j)!}{\prod a_j!}\right) = \ln((\sum a_j)\,!) - \ln(\prod a_j!)$$

$$\approx (\sum a_j)\ln(\sum a_j) - \sum a_j - \sum a_j \ln a_j + \sum a_j$$

Again we have the cancellations and we have this particular value of ln W i.e.-

$$\ln W = (\sum a_j)\ln(\sum a_j) - \sum a_j \ln a_j$$

And now just like what we had in the micro canonical ensemble, in this case also we have two constraints. The first constraint is-

$$g_1 = \sum a_j - A = 0$$

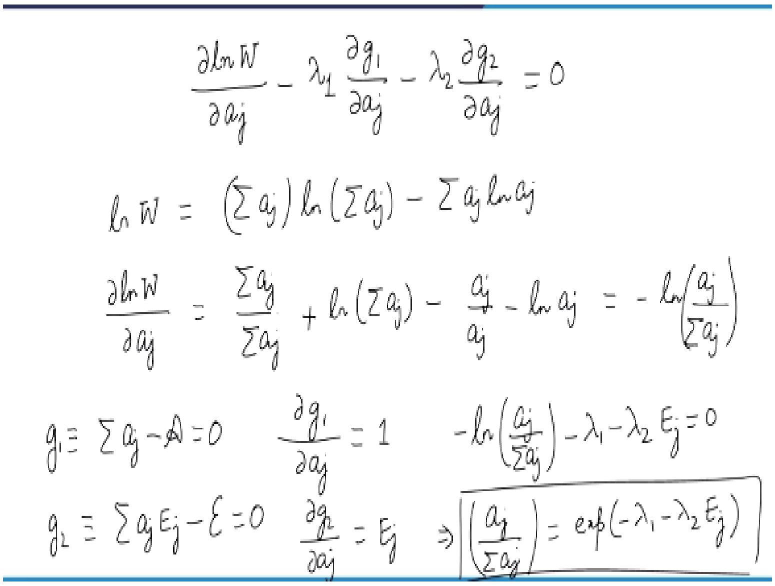
It is the total number of states in the ensemble, right. The meaning of the terms are different, but the equation looks identical to what we had for the micro canonical ensemble. Now we are talking about distribution within the systems as opposed to distribution within the molecules. So this is equal to zero.

And then again we have for the energy-

$$g_2 = \sum a_j E_j - \epsilon = 0$$

and the key point here is that, this ϵ is the total energy of a system, total energy of the ensemble to be precise. And the way to think about it is like although we have assumed that the systems have different energies that is placed in an infinite bath. So if I take the system along with the bath that becomes a system that is isolated so for the entire ensemble we may assume a total energy because there is nothing outside the ensemble right, because that systems plus the bath included must have some total energy and that we can assume to be constant. Other way to think about it is if I simply add over energy of all the systems in the ensemble that should be equal to constant and that I call as capital epsilon, not the same as small epsilon what I had earlier.

(Refer Slide Time: 24:03)



$$\frac{\partial \ln W}{\partial a_j} - \lambda_1 \frac{\partial g_1}{\partial a_j} - \lambda_2 \frac{\partial g_2}{\partial a_j} = 0$$

$$\ln W = \left(\sum a_j\right) \ln\left(\sum a_j\right) - \sum a_j \ln a_j$$

$$\frac{\partial \ln W}{\partial a_j} = \frac{\sum a_j}{\sum a_j} + \ln\left(\sum a_j\right) - \frac{a_j}{a_j} - \ln a_j = -\ln\left(\frac{a_j}{\sum a_j}\right)$$

$$g_1 = \sum a_j - A = 0 \qquad \frac{\partial g_1}{\partial a_j} = 1 \qquad -\ln\left(\frac{a_j}{\sum a_j}\right) - \lambda_1 - \lambda_2 E_j = 0$$

$$g_2 = \sum a_j E_j - \mathcal{E} = 0 \qquad \frac{\partial g_2}{\partial a_j} = E_j \qquad \Rightarrow \boxed{\left(\frac{a_j}{\sum a_j}\right) = \exp(-\lambda_1 - \lambda_2 E_j)}$$

So now the problem is we can write the maximization as-

$$\frac{\partial \ln W}{\partial a_j} - \lambda_1 \frac{\partial g_1}{\partial a_j} - \lambda_2 \frac{\partial g_2}{\partial a_j} = 0$$

And ln W we already have found is-

$$\ln W = (\sum a_j) \ln(\sum a_j) - \sum a_j \ln a_j$$

So if I do this derivative with respect to a_j, what I will have is using the same kind of math what we had for the canonical ensemble what we get is this,-

$$\frac{\partial \ln W}{\partial a_j} = \frac{\sum a_j}{\sum a_j} + \ln(\sum a_j) - \frac{a_j}{a_j} - \ln a_j = -\ln(\frac{a_j}{\sum a_j})$$

And I can find the-

$$g_1 = \sum a_j - A = 0$$

$$g_2 = \sum a_j E_j - \epsilon = 0$$

$$\frac{\partial g_1}{\partial a_j} = 1$$

$$\frac{\partial g_2}{\partial a_j} = E_j$$

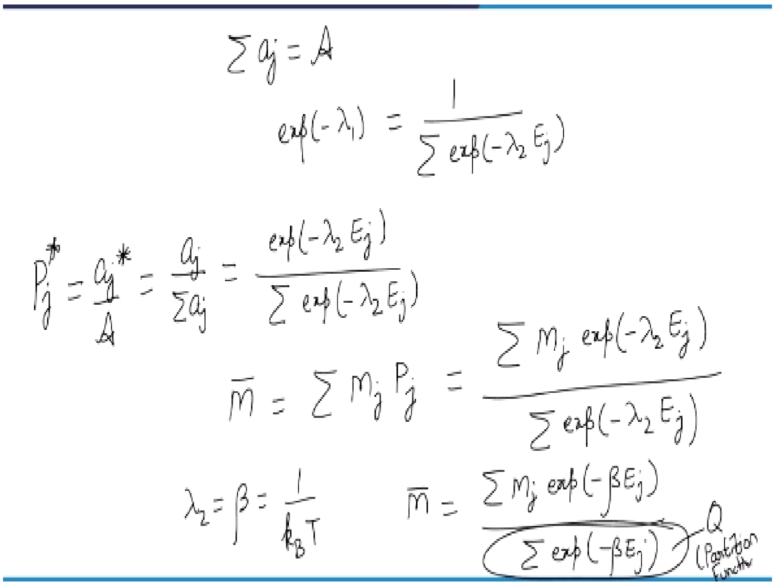
And therefore,

$$-\ln\left(\frac{a_j}{\sum a_j}\right) - \lambda_1 - \lambda_2 E_j = 0$$

And thus,

$$\left(\frac{a_j}{\sum a_j}\right) = \exp(-\lambda_1 - \lambda_2 E_j)$$

(Refer Slide Time: 26:34)



$$\sum a_j = A$$

$$exp(-\lambda_1) = \frac{1}{\sum exp(-\lambda_2 E_j)}$$

$$P_j^* = \frac{a_j^*}{A} = \frac{a_j}{\sum a_j} = \frac{exp(-\lambda_2 E_j)}{\sum exp(-\lambda_2 E_j)}$$

$$\bar{m} = \sum m_j P_j = \frac{\sum m_j\, exp(-\lambda_2 E_j)}{\sum exp(-\lambda_2 E_j)}$$

$$\lambda_2 = \beta = \frac{1}{k_B T} \qquad \bar{m} = \frac{\sum m_j\, exp(-\beta E_j)}{\boxed{\sum exp(-\beta E_j)}} \; Q \text{ (Partition function)}$$

If I again use the first constraint that is summation of a_j is equal to a. I will get pretty much the same result that we had earlier. And that gives me-

$$\sum a_j = A$$

$$\exp(-\lambda_1) = \frac{1}{\sum \exp(-\lambda_2 E_j)}$$

And therefore, what I have going back in the expression right here is-

$$\frac{a_j}{A} = \frac{a_j}{\sum a_j} = \frac{\exp(-\lambda_2 E_j)}{\sum \exp(-\lambda_2 (E_j)}$$

And now since we have found the most probable distribution, we can call it a_j* and this becomes the probability of the state j for the most probable distribution, right. So this again looks like the results of the Boltzmann distribution, but now it applies to system as a whole. So therefore, if I am now interested in any property M bar, I can write as-

$$\bar{M} = \sum M_j P_j = \frac{\sum M_j \exp\left(-\lambda_2 E_j\right)}{\sum \exp(-\lambda_2 E_j)}$$

We will show it later that the λ2 has the same meaning as earlier that is equal to β which is 1 over k_BT. We will establish that more clearly for the canonical ensemble. And actually, we will pretty much derive the zeroth law of thermodynamics in the coming classes. For the time being just take for granted that λ2 is equal to β. So we have-

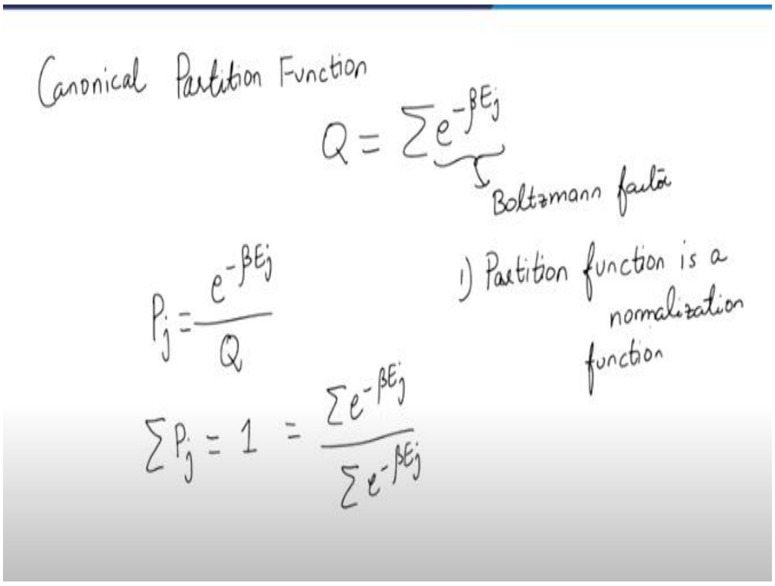
$$\lambda_2 = \beta = \frac{1}{k_B T}$$

And therefore,

$$\bar{M} = \frac{\sum M_j \exp(-\beta E_j)}{\sum \exp(-\beta E_j)}$$

Now there is something interesting about what you have in the denominator here. Let me give it a name and see what are the properties of this denominator that makes it very interesting. So let me call this thing Q, which is referred as the partition function.

(Refer Slide Time: 29:34)



Canonical Partition Function

$$Q = \underbrace{\sum e^{-\beta E_j}}_{\text{Boltzmann factor}}$$

$$P_j = \frac{e^{-\beta E_j}}{Q}$$

1) Partition function is a normalization function

$$\sum P_j = 1 = \frac{\sum e^{-\beta E_j}}{\sum e^{-\beta E_j}}$$

So the partition function, it comes out to be a very useful concept is defined as-

$$Q = \sum e^{-\beta E_j}$$

This we can think of as some sort of a Boltzmann factor because this tells me the probability of having a state E_j with some pre-factors. So essentially, if I sum over the Boltzmann factor for all the possible states of the system, what we get is a partition function. Since we are doing it for the canonical ensemble, we can call this the canonical partition function.

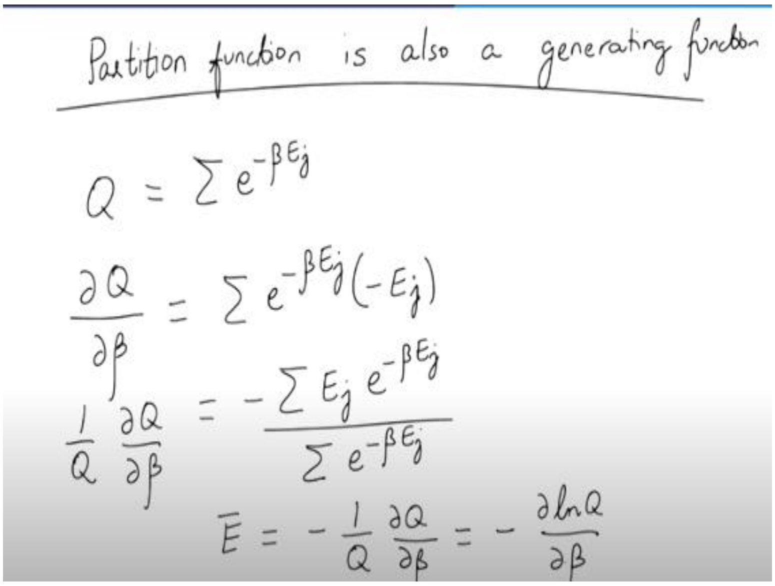
Now there is something very interesting about this partition function. First thing you may notice is that since it appears in the denominator right here, what essentially it is doing is somewhat normalizing the probabilities we know the probabilities are proportional to exponential of - βE_j.

So if I take a summation over that gives me a normalization of the probabilities. So basically, if I look at $e^{-\beta E_j}$ by Q, if the probability is defined as this, then the summation of the P_j for all the possible systems must be equal to 1, which has to be the case because essentially it is-

$$\sum P_j = 1 = \frac{\sum e^{-\beta E_j}}{\sum e^{-\beta E_j}}$$

And also because the probabilities must add to 1 so one use of the partition function is that partition function is a normalization function of probability.

(Refer Slide Time: 31:55)



Partition function is also a generating function

$$Q = \sum e^{-\beta E_j}$$

$$\frac{\partial Q}{\partial \beta} = \sum e^{-\beta E_j}(-E_j)$$

$$\frac{1}{Q}\frac{\partial Q}{\partial \beta} = -\frac{\sum E_j e^{-\beta E_j}}{\sum e^{-\beta E_j}}$$

$$\bar{E} = -\frac{1}{Q}\frac{\partial Q}{\partial \beta} = -\frac{\partial \ln Q}{\partial \beta}$$

The other even more interesting use is that the partition function also happens to be a generating function and what I mean by the generating function is if I know the partition function of the system, I can compute a whole host of properties. The other way of saying that is that if I know the partition function, we can find all the properties of the all thermodynamic property of the system and you can see by couple of examples.

So let us say if I start with Q as summation of $e^{-\beta Ej}$ now what happens is-

$$Q = \sum e^{-\beta E_j}$$

$$\frac{\partial Q}{\partial \beta} = \sum e^{-\beta E_j}\,(-E_j)$$

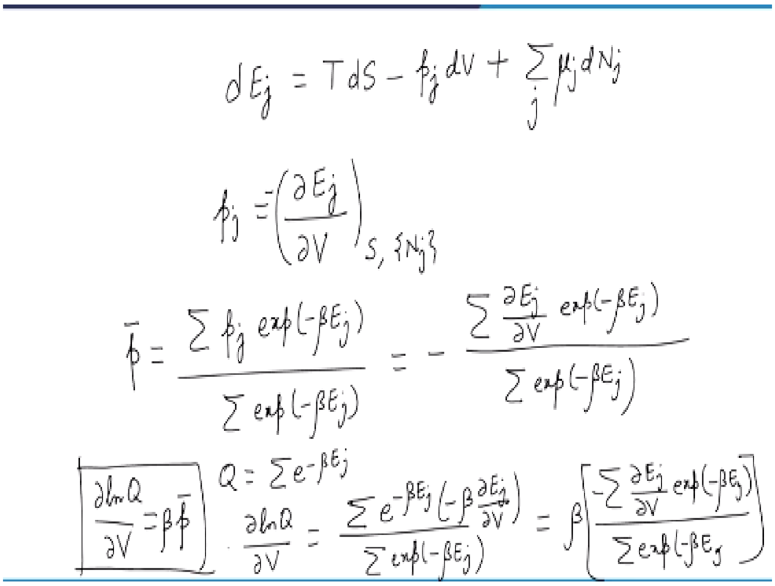
$$\frac{1}{Q}\,\frac{\partial Q}{\partial \beta} = \frac{-\sum E_j e^{-\beta E_j}}{\sum e^{-\beta E_j}}$$

If you go back the way we have defined our average properties, you may notice that if I simply replace M with E, what we are actually getting right there is the average value of the internal energy, right. So the E bar can therefore be defined as-

$$\bar{E} = -\frac{1}{Q}\,\frac{\partial Q}{\partial \beta} = -\frac{\partial \ln Q}{\partial \beta}$$

You can see I do not need to know anything else apart from Q, because if I know the function form of Q, I can find doh ln Q by doh β and I can basically recover the average energy of the system and therefore I call partition function as the generating function.

(Refer Slide Time: 34:36)



$$dE_j = TdS - p_j dV + \sum_j \mu_j dN_j$$

$$p_j = \left(\frac{\partial E_j}{\partial V}\right)_{S,\{N_j\}}$$

$$\bar{p} = \frac{\sum p_j \exp(-\beta E_j)}{\sum \exp(-\beta E_j)} = -\frac{\sum \frac{\partial E_j}{\partial V} \exp(-\beta E_j)}{\sum \exp(-\beta E_j)}$$

$$\boxed{\frac{\partial \ln Q}{\partial V} = \beta \bar{p}} \quad Q = \sum e^{-\beta E_j} \quad \frac{\partial \ln Q}{\partial V} = \frac{\sum e^{-\beta E_j}\left(-\beta \frac{\partial E_j}{\partial V}\right)}{\sum \exp(-\beta E_j)} = \beta\left[\frac{-\sum \frac{\partial E_j}{\partial V} \exp(-\beta E_j)}{\sum \exp(-\beta E_j)}\right]$$

Let us think of something else. So we know that for a particular state with energy E_j, we can use the thermodynamic expression for that particular state. And what we have is-

$$dE_j = TdS - p_j dV + \sum_j \mu_j dN_j$$

So if I am interested in the pressure of the state having energy E_j, the pressure can be defined as the partial derivative of E_j with respect to the volume of course, for constant S and constant number of molecules of all the species.

$$p_j = \left(\frac{\partial E_j}{\partial V}\right)_{S,\{N_j\}}$$

And if I want to find the average of that, it is going to be-

$$\bar{p} = \frac{\sum p_j \exp(-\beta E_j)}{\sum \exp(-\beta E_j)} = \frac{\sum \frac{\partial E_j}{\partial V} \exp(-\beta E_j)}{\sum \exp(-\beta E_j)}$$

And the same expression I can get if I simply do a derivative of the ln Q with respect to V-

$$\frac{\partial \ln Q}{\partial V} = \frac{\sum e^{-\beta E_j}\left(-\beta \frac{\partial E_j}{\partial V}\right)}{\sum \exp(-\beta E_j)} = \beta\left[\frac{-\sum \frac{\partial E_j}{\partial V} \exp(-\beta E_j)}{\sum \exp(-\beta E_j)}\right]$$

Therefore,

$$\frac{\partial \ln Q}{\partial V} = \beta \bar{p}$$

So therefore, if I know Q, not only I can find the e bar the average energy, I can also find the average pressure. So therefore, we say that the partition function is a generating function. You can extend the idea and take derivative with respect to n_j and find that even the chemical potential can be obtained using the partition function and what we obtain is the average chemical potential. So once we know the partition function of a system in the ensemble, then we can pretty much get the properties that we are interested in, in this case energy, pressure and the chemical potential. So this is the idea that I wanted to present for the canonical ensemble today.

In the next lecture, what you will see is, first of all, how can we also define fluctuations apart from the average using the idea of partition function. And then we also will have a closer connection with the part where we said that Lagrange multiplier should be equal to β. Actually, we will re-derive the third law of thermodynamics.

So with that I conclude here, thank you.