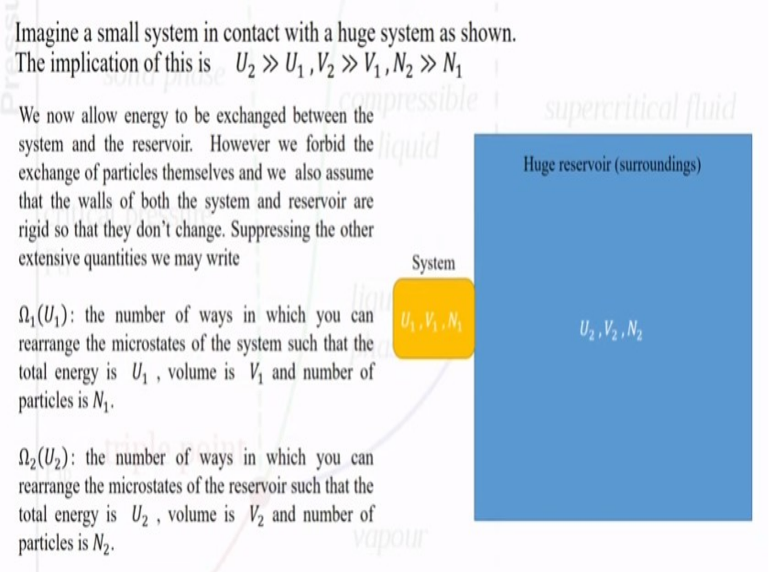
Statistical Mechanics
Prof. Girish S Setlur
Department of Physics
Indian Institute of Technology – Guwahati

Module No # 02
Lecture No # 07
Legendre transformation

(Refer Slide Time: 00:37)



Imagine a small system in contact with a huge system as shown.
The implication of this is $U_2 \gg U_1, V_2 \gg V_1, N_2 \gg N_1$

We now allow energy to be exchanged between the system and the reservoir. However we forbid the exchange of particles themselves and we also assume that the walls of both the system and reservoir are rigid so that they don't change. Suppressing the other extensive quantities we may write

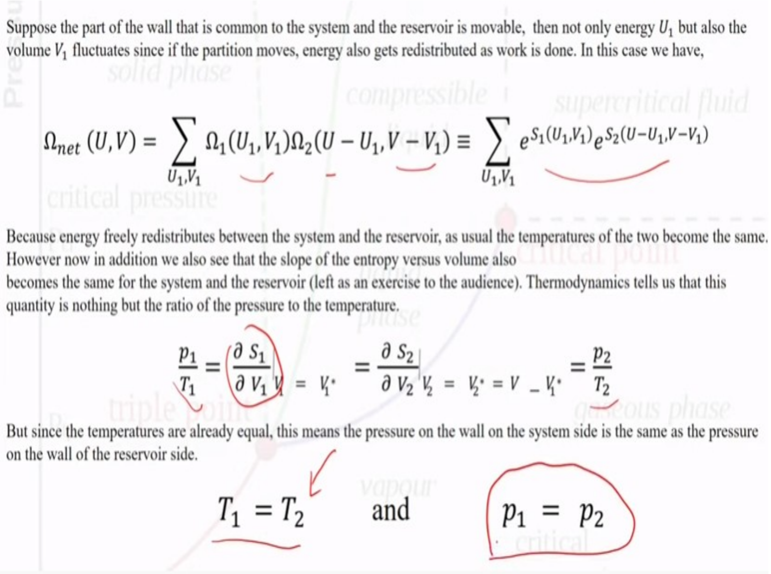
$\Omega_1(U_1)$: the number of ways in which you can rearrange the microstates of the system such that the total energy is U_1, volume is V_1 and number of particles is N_1.

$\Omega_2(U_2)$: the number of ways in which you can rearrange the microstates of the reservoir such that the total energy is U_2, volume is V_2 and number of particles is N_2.

Okay, let us continue so if you recall this is how we started, we imagined that we have a system that is in contact with a reservoir so there is a small system, well it is small compared to the reservoir but it is still you know huge in terms of the number of particle is much larger than 1. So we are going to assume that all the time, so this system is in contact with a reservoir and I am just going to you know fast forward because I have already told you that we have imagined a situation where the system is allowed to exchange energy with a reservoir or the surroundings but not anything else.

So in other words its volume and number of particles are fixed so when you do that you can count the number of ways in which you can rearrange the microstates of the system and the reservoir such that the total internal energy of the system is U_1 and because we do not really care about U_1 because that keeps fluctuating.

(Refer Slide Time: 01:50)



Suppose the part of the wall that is common to the system and the reservoir is movable, then not only energy U_1 but also the volume V_1 fluctuates since if the partition moves, energy also gets redistributed as work is done. In this case we have,

$$\Omega_{net}\,(U,V) = \sum_{U_1,V_1} \Omega_1(U_1,V_1)\Omega_2(U-U_1,V-V_1) \equiv \sum_{U_1,V_1} e^{S_1(U_1,V_1)}e^{S_2(U-U_1,V-V_1)}$$

Because energy freely redistributes between the system and the reservoir, as usual the temperatures of the two become the same. However now in addition we also see that the slope of the entropy versus volume also becomes the same for the system and the reservoir (left as an exercise to the audience). Thermodynamics tells us that this quantity is nothing but the ratio of the pressure to the temperature.

$$\frac{p_1}{T_1} = \frac{\partial S_1}{\partial V_1}\Big|_{V_1 = V_1^*} = \frac{\partial S_2}{\partial V_2}\Big|_{V_2 = V_2^* = V - V_1^*} = \frac{p_2}{T_2}$$

But since the temperatures are already equal, this means the pressure on the wall on the system side is the same as the pressure on the wall of the reservoir side.

$$T_1 = T_2 \quad \text{and} \quad p_1 = p_2$$

So what we have to do is find the total overall number of ways of doing it and that is the sum over all the U_1's. Because they are all independent ways of rearranging the microstates so now you can ask yourself like we did earlier what is the most probable value of U_1? and then that is found by maximizing this entropy, I mean rather this product which tells you the number of joint number of ways in which you can rearrange the microstates of both the system as well as the surroundings.

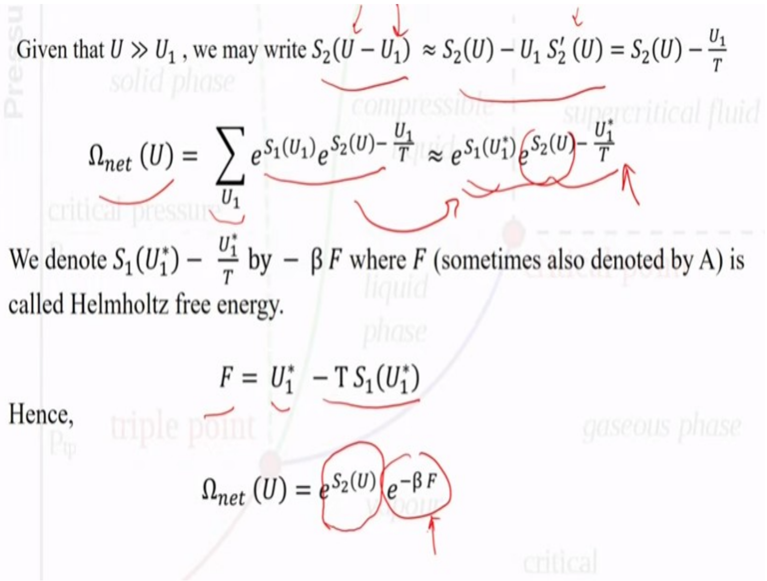
When you do that you are able to ascertain or you are able to convince yourself that there is a physical meaning to the slope of the entropy versus internal energy it is not merely formally defined as the reciprocal of the absolute temperature rather the physical meaning is that it is the temperature is what equalizes when you allow two systems in thermal contact to approach equilibrium.

So temperature has a very concrete physical meaning which is exemplified by this relation and this is something we rigorously derived by looking for the most probable configuration or the most probable energy which accomplishes this and so I told you that if you now start allowing for the walls of the container to be flexible so that it can compress and expand then naturally because work gets done.

So energy also gets exchanged so we are compelled to also imagine that not only the volume but also the energy fluctuates but then the total energy of the combined system of the reservoir and the system of interest remains fixed and so does the volume. So now you can ask the same question about the most probable volume of the system and what is the most probable energy of the system and so you get two different equilibrium conditions one is the usual equilibrium condition which forces us to equate the temperatures and the because the well the other condition forces us to equate this slope of the entropy versus volume on the reservoir side equal to the system side.

Now the slope of the entropy versus volume according to thermodynamics is the ratio of the pressure in the temperature. So recall we have just noticed that we just equated these two because energy gets exchanged and they reach thermal equilibrium so temperatures are anyway equal so because the ratio of p and T have to be equal and the temperatures are already equal that also implies that the pressures are equal.

(Refer Slide Time: 04:38)



Given that $U \gg U_1$, we may write $S_2(U - U_1) \approx S_2(U) - U_1 S_2'(U) = S_2(U) - \frac{U_1}{T}$

$$\Omega_{net}(U) = \sum_{U_1} e^{S_1(U_1)} e^{S_2(U) - \frac{U_1}{T}} \approx e^{S_1(U_1^*)} e^{S_2(U) - \frac{U_1^*}{T}}$$

We denote $S_1(U_1^*) - \frac{U_1^*}{T}$ by $-\beta F$ where F (sometimes also denoted by A) is called Helmholtz free energy.

$$F = U_1^* - T\, S_1(U_1^*)$$

Hence,

$$\Omega_{net}(U) = e^{S_2(U)} e^{-\beta F}$$

So this is where we left off last time so now I am going to go ahead and see if I can do this the similar thing to what I did well okay even this also we did earlier. So in order to study the overall number of ways in which you can rearrange the system what you do is you look at the entropy of the reservoir and then you realize that U is much larger than U_1 so you can do a Taylor series like this and then recall that this is nothing but the temperature the common temperature between the system and the reservoir and then you can write it like this.

So I am going to show you later that it is legitimate to completely ignore the fluctuations around the most probable state so what I have done here I have simply equated U_1 to its most probable value so the question is what gives me the right to do that after all there is a summation there over U_1. So if the implication is that the contributions from U_1 different from the most probable value is overwhelmingly suppressed compared to the most probable value and that is for systems which are large in size which we will establish this a little later.

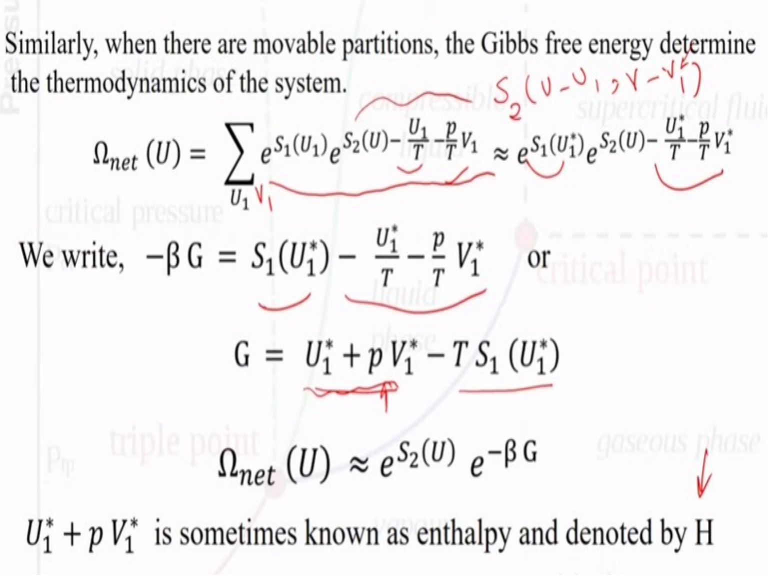
But for now let us assume that this is legitimate that I am going to legitimate on my part to simply replace the summation and drop the summation and replace U_1 by its most probable value. So in which case I end up getting an expression of this sort and so this e^{S2} goes outside because that is the reservoir which I do not care about. So if I look at only the system so I end up being able to write something like this.

So the net way of rearranging the microstates of course the net way of doing that for the reservoir times the system because it is in contact with the reservoir it is no longer e^{S} . So the influence of the reservoir has to be now taken into account. So it is taken into account through what is known as the free energy of the system. So the free energy is defined as the difference between the internal energy minus the temperature times the entropy at its most when evaluated at the most probable value of the energy.

So now of course the very term free energy implies that it has some physical meaning and indeed it does, so the answer I mean the physical meaning is that free energy is that energy which is available for you to utilize to do work. So it is something that is left over after you know all the entropy exchange I mean the energy exchanges are done and then there is some of that energy gets irretrievably lost in entropy and whatever remains is basically what you can use to do work.

So free energy therefore is the energy that remains for you utilize to do work you know useful work. So this is what is known as Helmholtz free energy so now I am going to repeat this calculation when I allow the walls of container to not be rigid. So in other words I allow of the walls of the container to move around.

(Refer Slide Time: 08:15)



Similarly, when there are movable partitions, the Gibbs free energy determine the thermodynamics of the system.

$S_2(U-U_1, V-V_1)$

$$\Omega_{net}(U) = \sum_{U_1 V_1} e^{S_1(U_1)} e^{S_2(U) - \frac{U_1}{T} - \frac{p}{T} V_1} \approx e^{S_1(U_1^*)} e^{S_2(U) - \frac{U_1^*}{T} - \frac{p}{T} V_1^*}$$

We write, $-\beta\, G = S_1(U_1^*) - \frac{U_1^*}{T} - \frac{p}{T} V_1^*$ or

$$G = U_1^* + p\, V_1^* - T\, S_1(U_1^*)$$

$$\Omega_{net}(U) \approx e^{S_2(U)}\, e^{-\beta\, G}$$

$U_1^* + p\, V_1^*$ is sometimes known as enthalpy and denoted by H

So as a result I told you already that not only the volume can fluctuate but also because if I allow the walls of the container to move around work gets done. As a result the internal energy also fluctuates so as a result I have to take into account both these processes and when I do that i can do something similar to what I was doing earlier namely this was this is the system so rather this is the reservoir and the reservoir entropy is now a function of U_2 and recall that

$$U_2 = U - U_1$$

So it also depends on the volume of the reservoir which is V_2 and V_2 is the difference between the total volume of the combined system and the reservoir which is fixed and the difference between that and the volume of the system V_1 which fluctuates. So now as usual as it was in the earlier case we may Taylor expand in powers of U_1 and V_1 and we end up with this expression because now the coefficient are nothing but the slopes of the entropy versus energy which is reciprocal temperature and slope of the entropy versus volume which is pressure by temperature.

So now as usual we are entitled or we take the liberty of replacing the energy of the energy and volume I should have ,sorry there is a volume summation as well, so the energy and volume of the system by it is most probable value and its most probable values are determined by equating the temperatures on either side of the system and reservoir and the pressures as well. So you have two equations and two unknowns which get fixed the two unknowns being the most probable energy and most probable volume of the system.

So as usual as before rather we will be able to identify this difference this plus this so in other words this plus this has a special name and that goes by the name of -βG. So the Gibbs free energy is therefore defined as

$$G = U + pV - TS$$

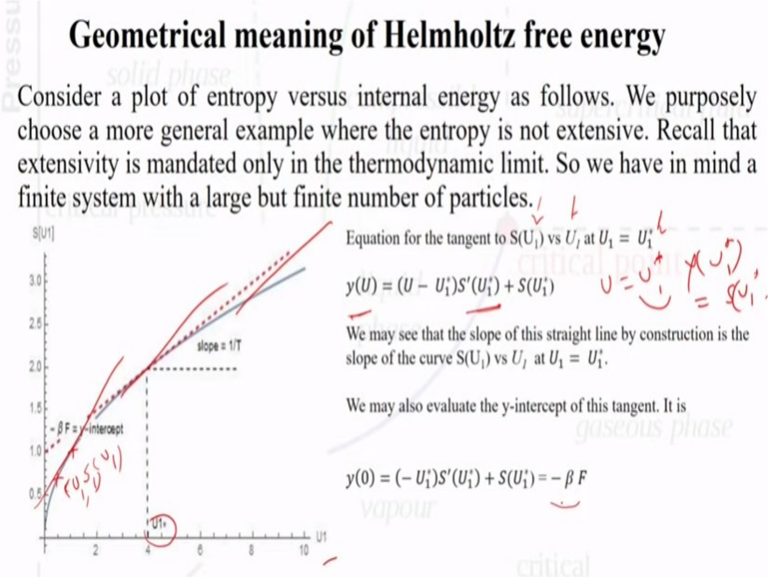
so remember that Helmholtz free energy where you only allow energy to get exchanged between system and reservoir was U – TS.

Now because you are also allowing volume to fluctuate you have an additional term called pV_1^* that is the most probable value. So this is called Gibbs free energy and this term here as a special name it is called enthalpy and it is sometimes denoted by H. So it is a so in other words Gibbs free energy is enthalpy minus TS just as Helmholtz free energy was internal energy minus TS, so here it is enthalpy minus TS.

So you can imagine that enthalpy is basically it takes on the role of internal energy when you allow volume to fluctuate. So it is the counter part of internal energy when you allow volume to fluctuate in that is called enthalpy. So now just as I told you that that Helmholtz free energy so it is called free energy precisely because you can use it to do work here too this is called Gibbs free energy precisely for the same reason.

But notice that the here when I say work I mean work other than the work that gets done because of the volume fluctuations so that is discounted because that is already taken into account that is not useful work that the system kind of tries to adjust itself. So you cannot do much about this so you cannot extract any further energy and do additional work. So that is taken away so when you take that away what remains is the energy that you can utilize actually to do work and that is called Gibbs free energy.

(Refer Slide Time: 12:40)



Geometrical meaning of Helmholtz free energy

Consider a plot of entropy versus internal energy as follows. We purposely choose a more general example where the entropy is not extensive. Recall that extensivity is mandated only in the thermodynamic limit. So we have in mind a finite system with a large but finite number of particles.

Equation for the tangent to $S(U_1)$ vs U_I at $U_1 = U_1^*$

$$y(U) = (U - U_1^*)S'(U_1^*) + S(U_1^*)$$

$U = U_1^*$, $y(U_1^*) = S(U_1^*)$

We may see that the slope of this straight line by construction is the slope of the curve $S(U_1)$ vs U_I at $U_1 = U_1^*$.

We may also evaluate the y-intercept of this tangent. It is

$$y(0) = (-U_1^*)S'(U_1^*) + S(U_1^*) = -\beta F$$

So now well that was a nice physical description of free energy and also temperature and pressure so I have told you what they mean physically in terms of you know quantites becoming equal on either side under various condition of the equilibrium so that is a physical description which is very useful and important. But now I am going to switch gears and also tell you about a geometrical or mathematical geometrical description of Helmholtz free energy.

So it is interesting to note that not only does the Helmholtz free energy and therefore also Gibbs free energy, I have a physical meaning which is easy to follow but it also has an important geometrical meaning which is also easy to follow. So in order to understand the geometrical significance of Helmholtz free energy so imagine there is a plot suppose I plot the entropy versus internal energy.

Of course I have purposely plotted it as not a straight line but in fact you should be wondering why it is not a straight line because that appears to violate extensivity. So we are actually purposely considering situations where extensivity is mildly violated because we are going to not operate in the thermodynamic limit. So I have told you extensivity is mandated only in the thermodynamic limit.

So if I relax that a little bit then it is the entropy function need not be completely extensive so it can only be approximately extensive. So if you remember that I told you that this I mean I claimed without proof that this goes through a maximum but that is actually a contradiction if you really assume that S_1 and S_2 are truly extensive. So if they are truly extensive then you see these are just linear functions of U_1 and V_1 and there is no chance that it goes through a maximum.

So linear function never goes through a maximum it just keeps increasing or decreasing so it never goes to maximum. So obviously already here when I say this what I have in mind already is you know I have something in mind namely that I am eventually going to accommodate deviations from thermodynamic limit which leads to deviations from extensivity, so that I already had in mind when I made these assertions because this would never reach a maximum or the most probable value would not even exist if that were not the case.

So if I strictly assume extensivity I would never be able to convince you that this product goes through a maximum. So I have to accommodate non extensive entropy and it is not scary I mean it sounds scary now but you will see that it is not that difficult to incorporate non extensivity in a systematic way around the basic result namely the extensive. So small fluctuation small deviations around extensivity can be accommodated easily.

So that is what we have in mind here so when we plot the entropy function versus internal energy and we purposely assume it is not extensive,so it is not a straight line in other words so there is a kind of curve here. So now you see you could either choose to describe this curve so you traditional way of describing any curve is to just mark out points on the curve and what are points? Points are just ordered collections of two number one number is corresponds to the x axis and the other is called ordinate and abscissa for those of you who are familiar to those terms.

But otherwise it is just a x axis and the y axis labels I mean the components rather so by specifying the you know the x coordinate of that point and the y coordinate of that point and whole bunch of such points and you should join them all by a smooth curve and that is how you describe a curve typically by listing all the points on the curve. But it so happens that there is a alternative way of describing a curve which does not involve listing all the points that lie on the curve.

So the alternative way of doing that is to instead list a whole bunch of straight lines such that a unique curve can be drawn which is tangent to all of them at the same time. So in other words so what you do is you draw a whole bunch of straight lines like this and then you ask yourselves which curve is tangent to all these lines means of course at different points.

So they are all tangent to the curve at different point so the different straight lines are tangent to the same curve at different points but then the curve itself is unique and a whole bunch of such straight lines. So it is so happens that you can instead of listing all the points that lie on the curve so the alternative way of thinking about the curve is to list all the possible straight lines, the family of tangents to the curve.

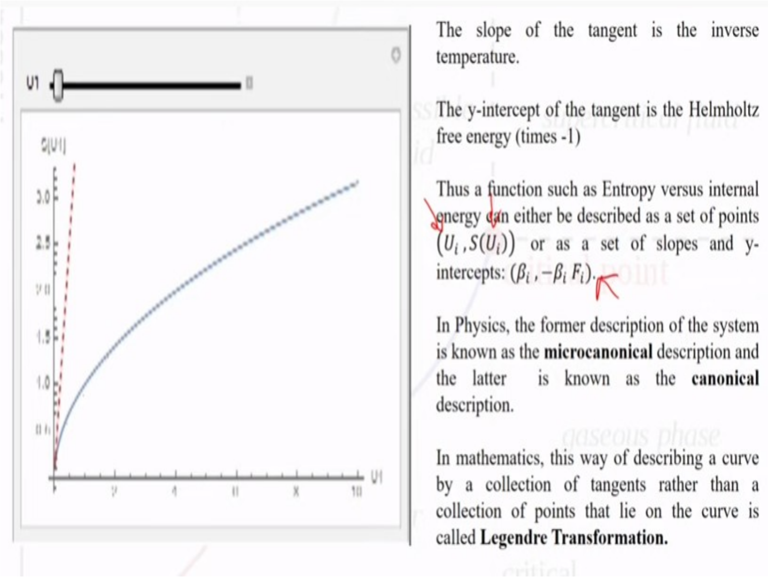
So that is the that is what we are going to do now so let us do that mathematically, so suppose S_1 vs U_1 is given now just focus on a particular value of U_1 called U_1^* and so remember that we have been calling the most probable value. So now suppose I draw a straight line tangent to this curve S(U) so what is the equation of that straight line it is going to have a slope which is corresponds to $S'(U_1^*)$ and it also has the value of $S(U_1^*)$ at U1. So in other words if $U = U_1^*$ then $U_1^* = S(U_1)$ because I mean they should touch I mean the straight line should touch the curve.

So the y coordinates of the curve and the straight line have to be the same at the place where it touches and the place where it touches is $U = U_1^*$. But then not only should it touch it should also touch in a tangential manner so in other words so how does one ensure that the straight line in the curve touch in a tangential manner, the way you ensure that it is by equating the slope of the straight line with this slope of the curve and which is why I have written it in this manner and that this way of doing it accomplishes that.

So now that I have got a straight line notice that this straight line is the unique the moment I pin down what is U_1^* and nature of the function S(U) this straight line becomes unique. Now because it is unique I can ask myself what is the value of the y intercepts so in other words if U is 0 so what is the value of y? So that is precisely what is known as Helmholtz free energy so that is the geometrical meaning of Helmholtz free energy, it is the y intercept so of course there is $-\beta$ there but other than that it is proportional to Helmholtz free energy.

So Helmholtz free energy is proportional to the y intercept of the tangent to the entropy versus the energy curve okay. So that is the geometrical meaning of Helmholtz free energy.

(Refer Slide Time: 20:56)



The slope of the tangent is the inverse temperature.

The y-intercept of the tangent is the Helmholtz free energy (times -1)

Thus a function such as Entropy versus internal energy can either be described as a set of points $(U_i, S(U_i))$ or as a set of slopes and y-intercepts: $(\beta_i, -\beta_i F_i)$.

In Physics, the former description of the system is known as the **microcanonical** description and the latter is known as the **canonical** description.

In mathematics, this way of describing a curve by a collection of tangents rather than a collection of points that lie on the curve is called **Legendre Transformation.**

You can see this happening for example by this animation that I just created, so this is the family of tangents. So you see that each point on the curve I can draw one tangent another tangent so the collection of all those tangents is specifying a collection of all those tangents is the same as specifying points on the curve. So notice that in order to specify a straight line which is basically what a tangent is you need two numbers, one is the slope and the other is the y intercept. So you have learnt that probably in your primary school already.

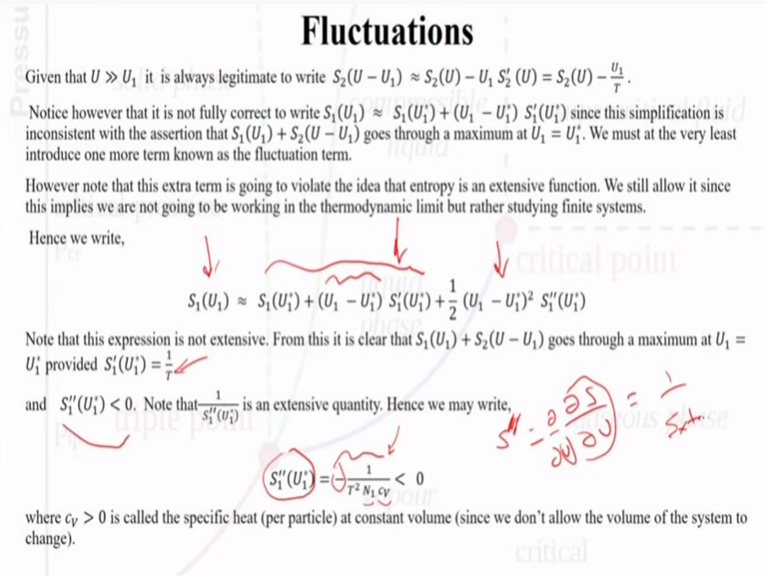
So by specifying the slope and the y intercept uniquely pin down a straight line but what is the physical meaning of the slope here. So this slope is just the inverse temperature and now the y intercept I have just given you a new interpretation and that is -βF. So what you are doing in a sense is specifying a straight line by a pair of numbers so the first entry in that pair is the inverse temperature of the second entry is Helmholtz free energy divided by –T which is temperature with a negative sign.

So by providing these two by providing these two numbers it is the equivalent of providing so there is alternative rather to providing U and S(U). So in other words you do not have to provide point that lie on the curve you can alternatively provide the details about the tangent to the curve. So this way of thinking about a curve so in physics the S vs U curve is called a micro-canonical description of the system.

Canonical I already told you what it means, canon means you know some rule book, so micro canonical because it kind of talks about combinotorics and counting and that sort of things it is very microscopic which is called micro-canonical. So the tangent way of doing things to describing that curve alternatively as a tangent rather than a collection of points is called a canonical description.

So that is the physics terminology, mathematics people call this transformation between you know thinking of the curve as a collection of points versus thinking of the curve as the collection of tangents so that is called in mathematic language it is called as Legendre transformation okay. So that is the name, so the free energy is sometimes called as Legendre transform of the entropy function.

(Refer Slide Time: 24:02)



Fluctuations

Given that $U \gg U_1$ it is always legitimate to write $S_2(U-U_1) \approx S_2(U) - U_1\, S_2'(U) = S_2(U) - \frac{U_1}{T}$.

Notice however that it is not fully correct to write $S_1(U_1) \approx S_1(U_1^*) + (U_1 - U_1^*)\, S_1'(U_1^*)$ since this simplification is inconsistent with the assertion that $S_1(U_1) + S_2(U-U_1)$ goes through a maximum at $U_1 = U_1^*$. We must at the very least introduce one more term known as the fluctuation term.

However note that this extra term is going to violate the idea that entropy is an extensive function. We still allow it since this implies we are not going to be working in the thermodynamic limit but rather studying finite systems.

Hence we write,

$$S_1(U_1) \approx S_1(U_1^*) + (U_1 - U_1^*)\, S_1'(U_1^*) + \frac{1}{2}(U_1 - U_1^*)^2\, S_1''(U_1^*)$$

Note that this expression is not extensive. From this it is clear that $S_1(U_1) + S_2(U-U_1)$ goes through a maximum at $U_1 = U_1^*$ provided $S_1'(U_1^*) = \frac{1}{T}$

and $S_1''(U_1^*) < 0$. Note that $\frac{1}{S_1''(U_1^*)}$ is an extensive quantity. Hence we may write,

$S'' = \frac{\partial}{\partial U}\left(\frac{\partial S}{\partial U}\right) = \frac{1}{S\ldots}$

$$\boxed{S_1''(U_1^*)} = -\frac{1}{T^2 N_1 c_V} < 0$$

where $c_V > 0$ is called the specific heat (per particle) at constant volume (since we don't allow the volume of the system to change).

So that is so much for the geometrical description of Helmholtz free energy so similarly also for Gibb's free energy. So let us get to the important thing which we have kind of glossed over which is something I told you already, so I told you that if you I kind of pretended that for instance this goes through a maximum as a function of U_1 and I did not actually prove it and that seems untenable if you assume extensivity.

So I already told you this is proportional to U_1, this proportional to $U - U_1$ which is U_2 there is no chance that this whole thing goes from through maximum because it is linear in U_1. So obviously there is something missing and what is missing is what is referred to as fluctuations and that is something we will have to get into right now okay. So what does that mean so the idea is that we of course do not stop here.

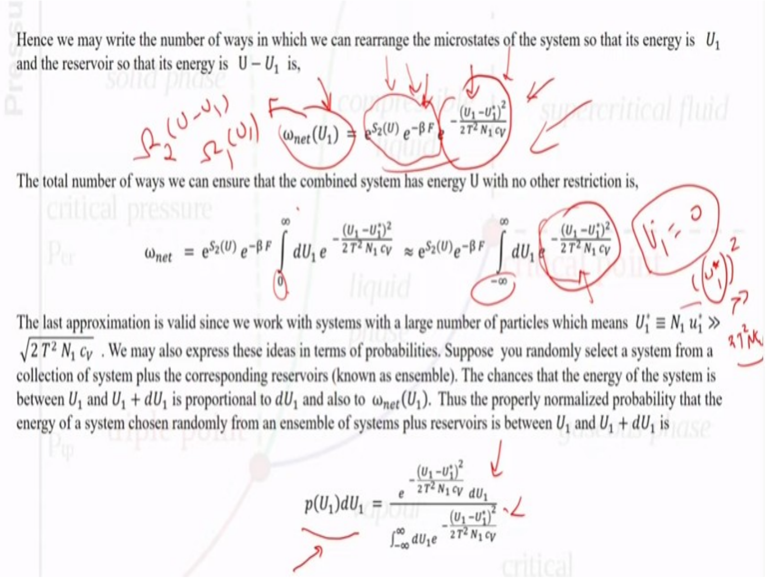
So what we do is we do a Taylor's series around U_1^* so remember that U_1^* is the most probable value of the internal energy of the system. So what I do is that I expand in powers of U1 – U_1^* so that is my small quantity so I am close to the most probable value of the internal energy. So I go ahead and expand but then instead of stopping here and I am mandated to stop here if my entropy function is truly extensive and then but then it is not because it is not I am allowed to or I take the liberty of going one step further.

So if I have a small deviation from extensivity I have to go one step further and I end up with this mixed term here. So recall that the slope is of course continues to be the reciprocal of temperature so in order for S_1 to reach a maximum it is imperative that this S, the second derivative being negative. So we going to assert that this is negative and we are going to give a name to second derivative of the entropy versus internal energy at the most probable value and that is in terms of what is called the specific heat per particle.

So C_V is V stands for constant volume so we will get to that later but right now is just a name so it is a whole bunch of things the important thing is there is a minus sign there and of course there is an implication that the second derivative of the entropy function with the internal energy is its reciprocal is extensive. So which is why I pulled out N_1 which is the number of particle so the implication is temperature which is already intensive and C_V means specific heat at constant volume per particle.

So that is the meaning of that and because it is per particle it is an intensive quantity so why do I have the right to claim that the reciprocal of this is extensive because if it is extensive and it is negative I have a right to say something like this, otherwise no. So how, what gives me the right to assume all this, negative I already told you because otherwise this would not go through a maximum but extensive is easy to see because after all what is S'' I mean two derivatives that means you take one derivative this becomes intensive and then you take one more derivative which is so this is intensive quantity one by temperature and then this is U in the denominator so it is 1 by extensive quantity so that is why I have written this as 1 / N_1.

(Refer Slide Time: 28:15)



Hence we may write the number of ways in which we can rearrange the microstates of the system so that its energy is U_1 and the reservoir so that its energy is $U - U_1$ is,

$$\omega_{net}(U_1) = e^{S_2(U)}\, e^{-\beta F} e^{-\frac{(U_1 - U_1^*)^2}{2T^2 N_1 c_V}}$$

The total number of ways we can ensure that the combined system has energy U with no other restriction is,

$$\omega_{net} = e^{S_2(U)}\, e^{-\beta F} \int_0^\infty dU_1\, e^{-\frac{(U_1 - U_1^*)^2}{2T^2 N_1 c_V}} \approx e^{S_2(U)} e^{-\beta F} \int_{-\infty}^\infty dU_1\, e^{-\frac{(U_1 - U_1^*)^2}{2T^2 N_1 c_V}}$$

The last approximation is valid since we work with systems with a large number of particles which means $U_1^* \equiv N_1\, u_1^* \gg \sqrt{2\,T^2\, N_1\, c_V}$. We may also express these ideas in terms of probabilities. Suppose you randomly select a system from a collection of system plus the corresponding reservoirs (known as ensemble). The chances that the energy of the system is between U_1 and $U_1 + dU_1$ is proportional to dU_1 and also to $\omega_{net}(U_1)$. Thus the properly normalized probability that the energy of a system chosen randomly from an ensemble of systems plus reservoirs is between U_1 and $U_1 + dU_1$ is

$$p(U_1)dU_1 = \frac{e^{-\frac{(U_1 - U_1^*)^2}{2T^2 N_1 c_V}}\, dU_1}{\int_{-\infty}^{\infty} dU_1 e^{-\frac{(U_1 - U_1^*)^2}{2T^2 N_1 c_V}}}$$

So now let us rewrite my you know the net way of doing things so what is this given as new symbol to this I do not want to confuse people so let me write down what this is this, this is nothing but

$$\omega_{net}(U_1) = \Omega_2(U-U_1)\Omega_1(U_1)$$

so that is what that is so I am just given it as a different name. So notice that is still a function of U_1 because I am taking into account fluctuations i have not replaced U_1 by U_1^*.

So U_1 is still available so if I choose to replace U_1 by U_1^* I get back my old result which is this. So in other words this product is nothing but the number of ways of rearranging the microstates of the reservoir times the $e^{-\beta F}$ we did this already but what is the ingredient is this one and this is because of fluctuation so I am now taking into account the possibility that the entropy function may not be fully extensive.

So as a result I get a distribution of this sort and of course I have skipped some steps here and I encourage you to go ahead and fill in some of those steps yourself because it is very hard for, you know it does not look nice for me to include all the steps in a presentation of this sort and in any event you will be lost if you do not do it yourself. So this subject I already told you several times that physics is not learnt it is not like watching a movie on your you know Laptop screen or it is actually you know taking a piece of paper and pen and working out following along working out the details as a instructor explains things to you.

So we end up with this expression which tells me that the overall ways of doing things where the internal energy is U_1 is the most probable times value of fluctuation, so now I can go ahead and add up all the U_1's and unlike the other time when I you know took the easy way out instead of summing over all the U_1's. I just choose to replace it by its most probable value so now I am going to do something less lazy which is actually you know some over all the U_1's.

But notice that I am doing a classical physics type of analysis so the energies are not discrete so in other words summation here summing over all U_1's is the same as integrating over all U_1's. So even if U_1's are positive of course I mean I have at the back of my mind some kind of ideal gas where its energy is mostly kinetic energy and it is positive and well as a result the lower limit is U_1 but then you know I have the liberty of replacing the lower limit by $-\infty$ if I can convince myself that the contributions from U_1 all the way from $-\infty$ to 0 can be neglected so it is like I have just added a whole bunch of zero's.

Because i have included a new range I have to strictly start from 0 and go to ∞ but by making this approximation the claim is that you can go all the way from $-\infty$ to ∞ because $-\infty$ to 0 hardly contributes anything and that is going to be the case only if I can convince myself that you know if, I mean what is the maximum value of this integrand in that region of U_1 being $-\infty$ to 0.

So the maximum value is when U_1 is actually 0 so I mean when it is anything bigger than 0 is even when it is not bigger than so notice that U_1 is on the negative side. So it is a so you either have 0 or something slightly negative or even more negative. So if U_1 is very very negative notice that it is $U_1 – U_1^*$ so that is already a very negative and very negative makes this highly suppressed it of e raise to minus very negative squared.

So that is a large quantity and it suppressed but however the way to minimize this is to make U_1 as less as possible from the negative side. So in other words makes its magnitude as less as possible which is basically make it approach 0. So in other words all I have to make sure is that even if U_1 is 0 this is still negligible so if it is negligible when U_1 is 0 its certainly going to be negligible when U_1 is more negative than 0.

So how do I convince myself that it is negligible when U_1 is 0 you just have to realize that U_1 is extensive and it is just a number of particles times some intensive quantity so long as that is much larger than this quantity so in other words I am just convincing myself that U_1's is much larger than $2T^2$ NCV and that is certainly going to be the case because U_1 is extensive which is proportional to the number of particles.

And the square of that is going to be huge compared to this quantity which is proportional to the number of particles so the left hand side is proportional to the square of the number of particles the right hand side is proportional to the number of particles. So that is certainly justifying so given that is justified now we can also form a probabilistic interpretation of this expression that I just wrote down and the probabilistic description is like this.

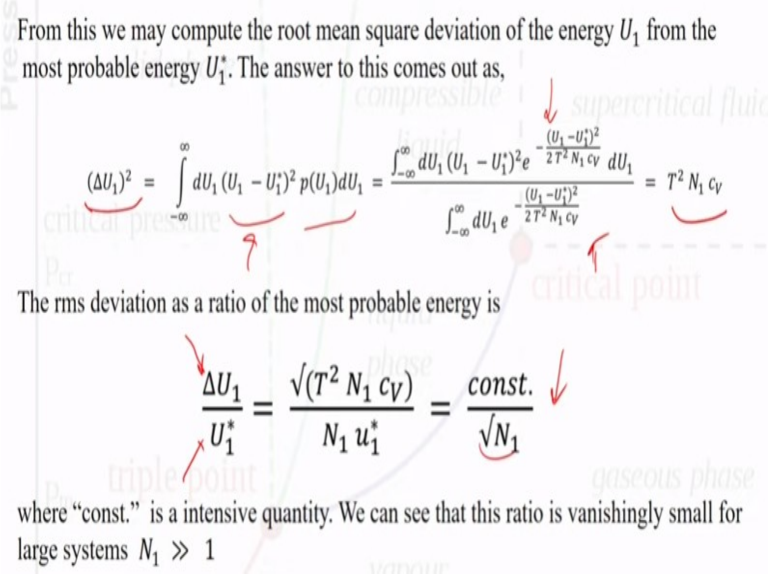
So suppose I ask myself suppose I have a collection of system plus a reservoir suppose I have many copies of that okay so I have copies of system plus reservoir I have a huge number of copies so I randomly you know put my finger on one of those systems and I ask myself what are the chances that the energy of that system is between U_1 and $U_1 + dU_1$.

So the answer is that is of course going to be proportional to dU_1 because it is so I mean so I am asking what are the chances that the energy of the system is between U_1 and $U_1 + dU_1$ is certainly proportional to dU_1 but then it is also proportional to this quantity because after all this is the number of ways in which you can rearrange the micro states of the system and of course the reservoir such that the system's energy is U_1.

So it is going to be proportional to this because this is the number of ways in which you can do that, so as a result I am entitled to make this claim that the probability. So if you have and this called an ensemble so this imaginary collection of system plus reservoir, so many Xerox copies of system plus reservoir so these are mental copies that you make and this is called an ensemble symbol. So this ensemble you can ask yourself what is the probability that in this ensemble the energy of the system is between U_1 and dU_1.

So it is like so the answer is that you count how many systems fall in that window that means with energy between U_1 and $U_1 + dU_1$ and you divide by the total number of systems in U ensemble so that is your number of system which fall in that window and that happens to be by definition P dU$_1$ and P is the now the probability distribution of finding that system in that window. So now from this expression it is clear that I can write like this and this is by construction properly normalized well if I integrate all over U$_1$ and I get 1.

(Refer Slide Time: 37:01)



From this we may compute the root mean square deviation of the energy U_1 from the most probable energy U_1^*. The answer to this comes out as,

$$(\Delta U_1)^2 = \int_{-\infty}^{\infty} dU_1\,(U_1 - U_1^*)^2\, p(U_1)dU_1 = \frac{\int_{-\infty}^{\infty} dU_1\,(U_1 - U_1^*)^2 e^{-\frac{(U_1-U_1^*)^2}{2T^2 N_1 c_V}}\, dU_1}{\int_{-\infty}^{\infty} dU_1\, e^{-\frac{(U_1-U_1^*)^2}{2T^2 N_1 c_V}}} = T^2 N_1 c_V$$

The rms deviation as a ratio of the most probable energy is

$$\frac{\Delta U_1}{U_1^*} = \frac{\sqrt{(T^2 N_1 c_V)}}{N_1 u_1^*} = \frac{const.}{\sqrt{N_1}}$$

where "const." is a intensive quantity. We can see that this ratio is vanishingly small for large systems $N_1 \gg 1$

So now I can go ahead and ask myself so remember that when I started off I kind of totally ignored the fluctuations altogether but that was conceptually wrong because then you know if it is extensive it does not go through a maximum so I would not have any right to replace summation over its most probable value. So obviously that is only valid if it goes most probable value and it goes through a most probable value because there are fluctuations.

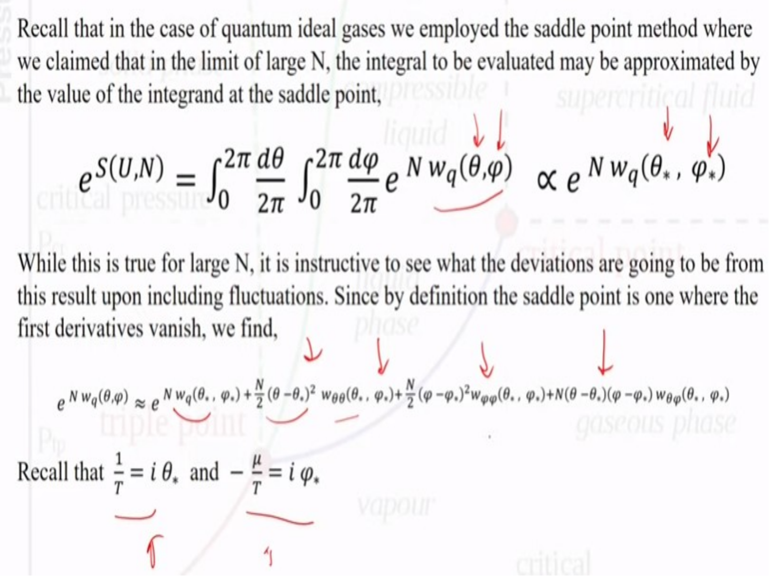
So it appears now that the it is a worrying situation because now it appears that I have to after all take into account fluctuations because they appear to play such a central role. But then happily we are not in such a difficult situation because we can go ahead and evaluate the precise magnitude of that fluctuation. So what we can do is estimate what is known as the root mean squared deviation of the energy of that system in that ensemble so using this probability distribution.

So when you do that so recall that the root mean squared is defined in this way if you look at the deviation you square it and you take the mean. So this is the taking the mean is basically averaging over the probability distribution which is what I have done here, so when I do this I end with this result but then notice this is the square of the deviation. So when I take the ratio so now I look at the deviation and I look at the ratio of the deviation of the fluctuation of the internal energy the estimate rather.

So in that ensemble and I look at the ratio of that fluctuation with the most probable energy and that comes to be inversely proportional to the square root of the size of the system. So this is the very important result central result in statistical mechanics which says that when the system sizes are large that even though conceptually fluctuations are important in order to ensure that you know the number of ways in which you can rearrange the microstates actually goes through a maximum.

So if you ignore fluctuations it does not go through a maximum it just keeps increasing so conceptually it is important for fluctuations to be there but fortunately the precise magnitude of the fluctuations are not that important it is just the fact that they exist is important because when you evaluate the fluctuations themselves they come out to be heavily suppressed by the number of particles in your system.

(Refer Slide Time: 39:55)



Recall that in the case of quantum ideal gases we employed the saddle point method where we claimed that in the limit of large N, the integral to be evaluated may be approximated by the value of the integrand at the saddle point,

$$e^{S(U,N)} = \int_0^{2\pi} \frac{d\theta}{2\pi} \int_0^{2\pi} \frac{d\varphi}{2\pi}\, e^{\,N\, \underbrace{w_q(\theta,\varphi)}} \propto e^{\,N\, w_q(\theta_*\,,\,\varphi_*)}$$

While this is true for large N, it is instructive to see what the deviations are going to be from this result upon including fluctuations. Since by definition the saddle point is one where the first derivatives vanish, we find,

$$e^{\,N\, w_q(\theta,\varphi)} \approx e^{\,N\, \underbrace{w_q(\theta_*\,,\,\varphi_*)} + \frac{N}{2}\underbrace{(\theta-\theta_*)^2}\, w_{\theta\theta}(\theta_*\,,\,\varphi_*) + \frac{N}{2}(\varphi-\varphi_*)^2 w_{\varphi\varphi}(\theta_*\,,\,\varphi_*) + N(\theta-\theta_*)(\varphi-\varphi_*)\, w_{\theta\varphi}(\theta_*\,,\,\varphi_*)}$$

Recall that $\underbrace{\frac{1}{T}} = i\,\theta_*$ and $-\underbrace{\frac{\mu}{T}} = i\,\varphi_*$

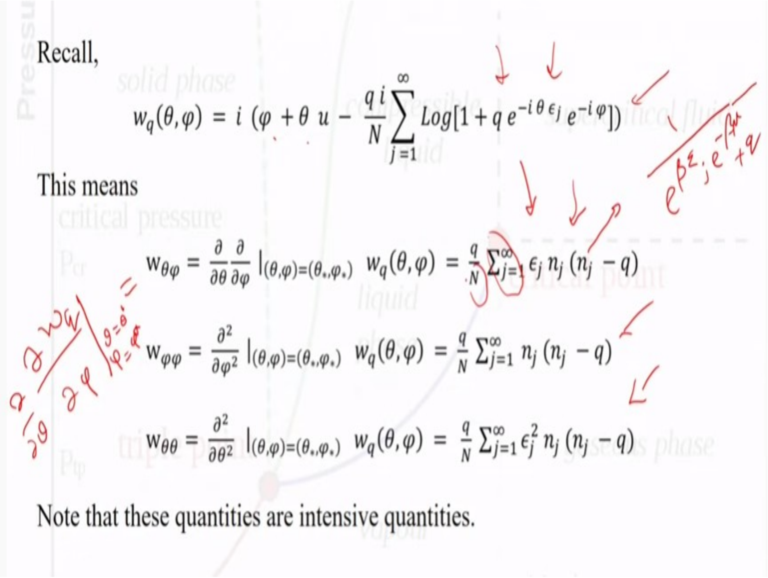
So that was mostly a classical description but then I can go ahead and also repeat this activity for a quantum system. So recall that in a quantum system I had done something very similar to what i did earlier for a classical system where I had you know energy, volume energy was continuous that sort of thing that is what we have been doing till now. But now for a quantum system if you recall there were this funny angles called θ and φ over which I had to integrate and then I invoked something called as saddle point method.

Of course without proving it but when I said that it is legitimate to replace this by some star you know some quantities called θ^* , φ^* these are the analogs for the most probable values that I was taking about till now. So the quantum analogs of the most probable value but here too this makes no sense at all if this w does not have a feature which is known as a saddle point. So it so happens that it is only when this w was goes through a saddle point then you can do this.

So I am not I am not going to go through the technical details we can kind of postpone the details to a later date or maybe you can look at up in any of the math's books. But from a physics point of view all it means is that just like in the earlier few slides I was compelled to take into account quadratic terms which correspond to fluctuations here too that is all I do the linear terms of course vanish identically precisely because that is the definition that yields by setting the linear terms in this Taylor series to 0 yields the definitions of.

So in other words these are obtained by setting the linear terms to 0 so you have 0th term then immediately I jump to the quadratic term so the linear terms are 0 and which leads to the expression. So this is how I am going to stop so if I mean stop means stop the Taylor series. So I have stopped the Taylor series at the quadratic term so let me evaluate these coefficients so these coefficients are nothing but the second derivative of w with respect to θ and φ and then there is mixed derivative with θ and φ.

(Refer Slide Time: 42:31)



Recall,

$$w_q(\theta,\varphi) = i\,(\varphi + \theta\, u - \frac{q\, i}{N}\sum_{j=1}^{\infty} Log[1 + q\, e^{-i\,\theta\,\epsilon_j}\, e^{-i\,\varphi}])$$

This means

$$w_{\theta\varphi} = \frac{\partial}{\partial\theta}\frac{\partial}{\partial\varphi}\,|_{(\theta,\varphi)=(\theta_*,\varphi_*)}\; w_q(\theta,\varphi) = \frac{q}{N}\sum_{j=1}^{\infty}\epsilon_j\, n_j\,(n_j - q)$$

$$w_{\varphi\varphi} = \frac{\partial^2}{\partial\varphi^2}\,|_{(\theta,\varphi)=(\theta_*,\varphi_*)}\; w_q(\theta,\varphi) = \frac{q}{N}\sum_{j=1}^{\infty} n_j\,(n_j - q)$$

$$w_{\theta\theta} = \frac{\partial^2}{\partial\theta^2}\,|_{(\theta,\varphi)=(\theta_*,\varphi_*)}\; w_q(\theta,\varphi) = \frac{q}{N}\sum_{j=1}^{\infty}\epsilon_j^2\, n_j\,(n_j - q)$$

Note that these quantities are intensive quantities.

$\frac{\partial^2 w_q}{\partial\theta\,\partial\varphi}\Big|_{\theta=\theta^*,\ \varphi=\varphi^*} =$

$e^{\beta\epsilon_j} e^{-\beta\mu} + q$

So now w is defined in this fashion and then if I go ahead and evaluate this I get these nice expressions. So for $w_{\theta\varphi}$ what does this mean? This means

$$w_{\theta\varphi} = \partial_\theta\partial_\varphi|_{(\theta,\varphi)=\theta^*,\ \varphi^*}\ w_q(\theta,\varphi)$$

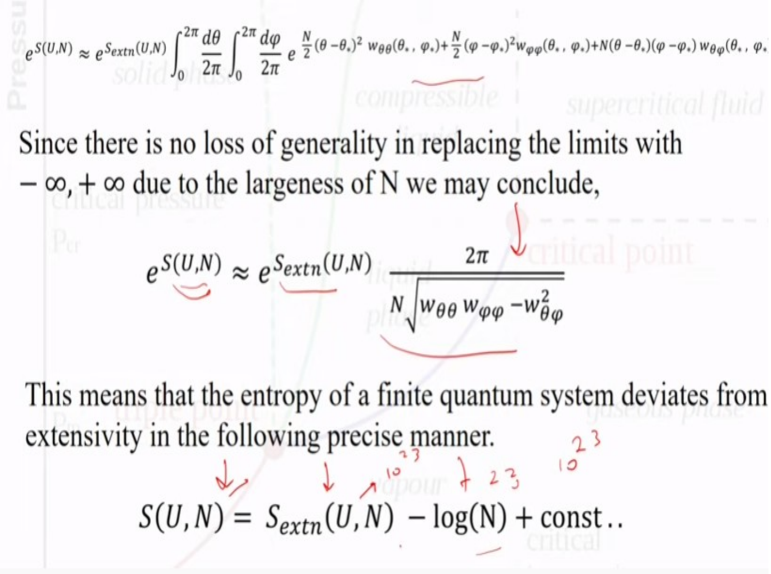
so that is what that means. So similarly here it means you differentiate with respect to φ twice and you put $\varphi = \varphi^*$ and $\theta = \theta^*$.

So when you do all that it is little bit tedious but it can be done notice if I differentiate twice these thing go away. So only this survives and then I rearrange my terms and what is n_j so remember what n_j was this is the number of the average number of ways in which I mean the average number of Bosons or Fermions depending upon the value of q that occupies energy level ϵ_j.

So n_j is the average number of quantum particles in energy level ϵ_j and that happens to have a mathematical expression of this sort so that is what that is. I am just recollecting, so bottom line is that this $w_{\theta\varphi}$ has this expression $w_{\varphi\varphi}$ has this expression and $w_{\theta\theta}$ has this expression but what is common to all this because they are very simple looking but what is also common is the fact that they are intensive quantities because recall that this is a this is kind of a of the size of the system.

So summing over all the j’s effectively counts the number of energy levels and each levels are populated by particles so in some sense that becomes extensive so it becomes scales with the size of the system. So that divided by n is basically an intensive quantity. So these coefficients here they are intensive quantities so we have to keep that in mind as we proceed.

(Refer Slide Time: 44:50)



$$e^{S(U,N)} \approx e^{S_{extn}(U,N)} \int_0^{2\pi} \frac{d\theta}{2\pi} \int_0^{2\pi} \frac{d\varphi}{2\pi}\, e^{\frac{N}{2}(\theta-\theta_*)^2\, w_{\theta\theta}(\theta_*,\, \varphi_*) + \frac{N}{2}(\varphi-\varphi_*)^2 w_{\varphi\varphi}(\theta_*,\, \varphi_*) + N(\theta-\theta_*)(\varphi-\varphi_*)\, w_{\theta\varphi}(\theta_*,\, \varphi_*}$$

Since there is no loss of generality in replacing the limits with $-\infty, +\infty$ due to the largeness of N we may conclude,

$$e^{S(U,N)} \approx e^{S_{extn}(U,N)} \frac{2\pi}{N\sqrt{w_{\theta\theta}\, w_{\varphi\varphi} - w_{\theta\varphi}^2}}$$

This means that the entropy of a finite quantum system deviates from extensivity in the following precise manner.

$$S(U,N) = S_{extn}(U,N) - \log(N) + \text{const}..$$

So now we go ahead and proceed and we see that we can evaluate this integral and we end up with this so when you take into account fluctuations the bottom line is the answer for the entropy including fluctuations is related to the answer for the entropy by ignoring fluctuation which is why I have called it with the subscript extn meaning extensive entropy. So this is extensive entropy and then there is a contribution due to fluctuation which leads to deviations from extensivity.

If you take the log on both sides you will see that it tells you the precise way in which entropy deviates from exensivity. So of course you know this is not at all important if N is huge so keep in mind that if n is the order of 10^{23} this is going to be proportional to 10^{23} this is so this is going to be proportional to 23 because it is log of that. So 1 10^{23} is huge compared to 23 so it makes no difference whether I subtract this or not.

So for huge systems fluctuations just like conceptual role in the analysis they do not play practical role so that is something to keep in mind okay. So I am going to stop here in the next hour I am going to tell you what is an important notion known as specific heat so we are going to study the specific heat of ideal gas, classical ideal gas, classical Fermi gas and Classical Bose gas alright so hopefully you will join me for the next hour. Thank you.