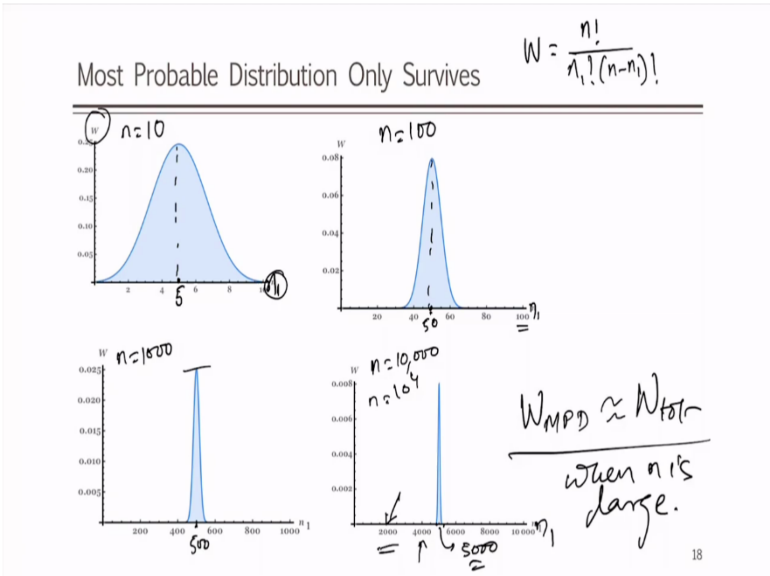
Chemical Principles II

Professor Dr Arnab Mukherjee
Department of Chemistry
Indian Institutes of Science Education and Research, Pune

Most Probable Distribution

Okay, so now as I said most probable distribution. Now you see we give two examples, 210 particle and we gave you an example of 2 particle. Now we said that somehow we went to 210 it was fluctuating around 105, 105 but not really going to 50 or 75, why is that the case? Why all those possible distributions that could be possible 1-209, 2-208, 3-207 like that all possible, how many will be there? n possible, so 210 possibilities, possible distributions are there, why? They were not happening, why only one distribution which is just half or most probable distributions are surviving you can see that from here.

(Refer Slide Time: 1:10)



So, here we are plotting the W with respect to the number of you know W is a number of micro state for possible distributions. Let's say you can call that as n1 so n1 in one box versus other. So we're just plotting this function where W is its factorial n by n1 factorial and n minus n1 factorial, we are plotting this particular quantity as W and plotting against n1.

Now what you see is that, it's pretty flat distribution so our n is equal to 10 particles, yeah, its 10 particles. And the peak as you can see is at five. Now, let us go to the next one. Now I'm plotting 100 particle and you see the peak is at 50. Again its peak is in the middle, but somehow the distribution has become narrow one, isn't it? Now, with respect to what it is become narrow one with respect to the percentage deviation. So if you call that as a half, then with respect to its half it has narrowed down.

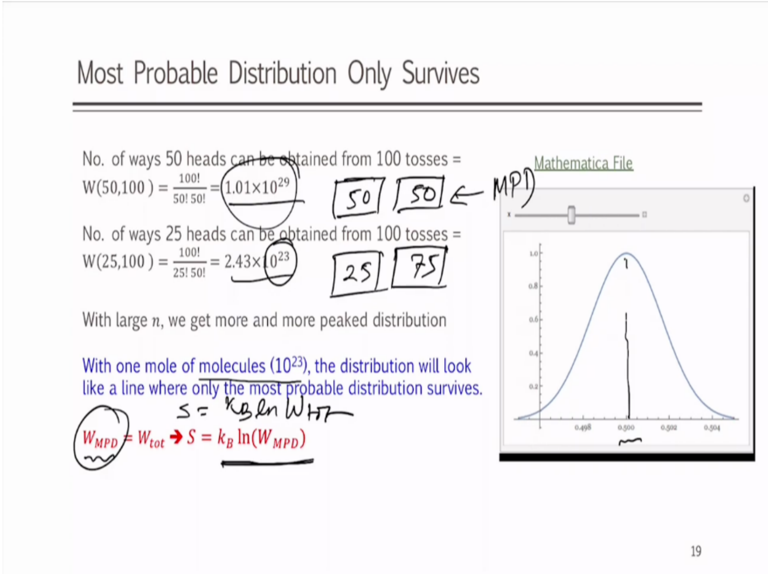
Now, let us go to 1000 particle. Again as you see I'm plotting the W, I will write the formulae here itself so that you can refer to it, n factorial by n1 factorial and n minus n1 factorial. This is the two-level system number of micro states, right. So, here the n is 10, here the n is 100, here the n is 1000, it is peaking up at 500 as you can see. The most probable value or most probable distribution or most probable the maximum W or most probable value of the W is in the middle when it is half at 500 and it has become much narrow at even than the hundred, right?

Let's go to 10,000, here n is 10,000. Again peaking is at 5000 and it has become much narrower. So this 10,000 is just 10 to the power 4, okay. You have to go to 90 more orders of magnitude, in order to reach 10 to the power 23 which is just one mole of particles. Imagine what will happen to this particular graph. Even, .000000001 percent deviation will be extremely small probability.

So therefore we can safely say that it is the most probable distribution, one distribution for two-level system which has half of particle is what is required to calculate the W. Which means we say that our W most probable distribution is all possible distribution when N is large. Because we don't need the other one, because as you can see the graphs drops here, we don't need the other W's, other values. So see for example n1 equal to 2000 is not there no W almost for that, there is no W for even 4000, it drops just around 5000, if you plot at the percentage deviation it will be much more clear.

So percentage deviation from [inaudible 5:11] much more clear. So, according to the percentage deviation is not you know at all changing. Which only means that it is the 5000 like for 10,000 particles is a 5000 which is most probable, so that means half and half its most probable. And that is you need only that and everything else will not be required at all.

(Refer Slide Time: 5:35)



We will show you this through the example, so again number of ways that 50 heads can be obtained from hundred tosses. Which is again 50 particle can be distributed in one box out of hundred particles, is the same thing is 1.01 into 10 to the power 29. If I just do 25, 75 it will be 23 six orders of magnitude less. So it is like the same situation 50 heads out of hundred tosses is same as saying that I give 50 particles in this box and 50 particle in this box and that has a W of this much.

Now if I say that okay I give 25 particle in this box and 75 particle in this box, that has this many possibilities, six orders of magnitude less possibilities than this. So this is our MPD, this is another distribution which is less probable. So with large N more and more this will be prominent such that for a one mole particle the distribution will look like a line which you call as delta function, that's just a straight line, just one line which means that only one value survives and everything else does not.

So one distribution is more than enough and that's why we say that all possibilities, for n particle how many possibilities distributions? n. But all n minus 1 are not important only one is important and that is the most probable distribution. So if we say that our entropy is kbl and WMPD actually entropy is kb entropy should be kblnw tot but since our W tot is almost equal to WMPD we only need one distributions you know most probable distributions to calculate the entropy. And you will see that it has an enormous implications to understand future things later on.

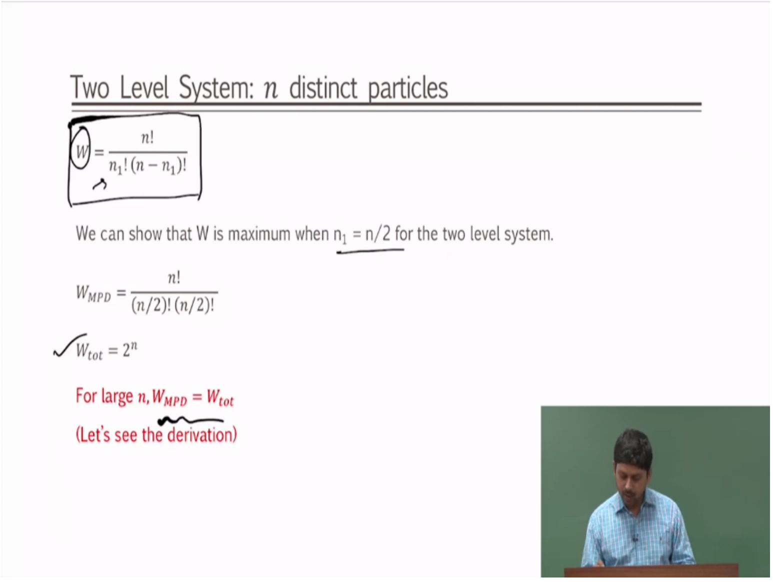
We will show you that it is not possible for us to calculate all possible distributions for realistic system. However we can still calculate MPD and since we know almost probable distribution and since we know the most probable distribution is all that we required to calculate our entropy we are safe. Had we required all possible distributions, we could not calculate entropy for realistic systems.

We will show you that, that it is very important to know that most probable distribution is the only distribution that is required for us to calculate the entropy and most probable distributions can be calculated for any realistic systems due to Boltzmann, we will come to that. Now, here you can see I'm showing in Mathematica that how with number of N's the distribution getting sharper and sharper, but I'm going to show you some calculations in Mathematica as well.

So, as you can see in Mathematica we have used, there is a software if in which you can calculate this big combination problems and then you can plot that and we have shown that how with the percentage deviation 0.5 as the number increases the distribution getting sharper and sharper, you see here we have use the relative deviation from the half. So 0.5, the maximum is always maintained that at this value as you can see but with N it is shrinking, and that's all is required.

So when the number becomes very large, here we have not used much larger like 10,000 or something. If it goes very very high then the distribution will be much much sharper.

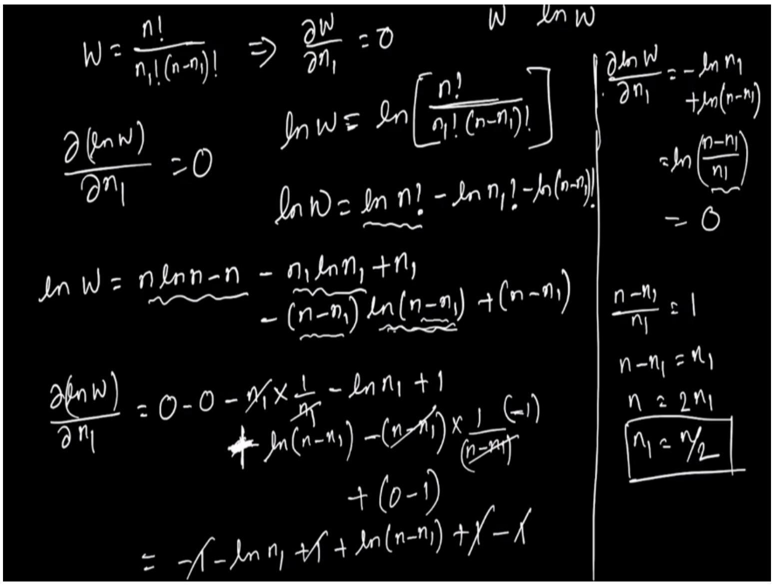
(Refer Slide Time: 9:18)



Now, we can mathematically show that, we really don't have to rely on this graph to show you that the most probable distribution survives. So, here the W denotes some you know number of micro sets from for a particular distribution, meaning for n1 in one box and n minus n1 in other box.

And we can show that W is maximum when n1 is half, we will first show that, and the. So why that is half we can show that and then all possible distributions as we said for an distinguishable particles in distinguishable boxes, all possible distribution is 2 to the power n, we have discussed that, that is all possibilities. And we will show that for large n this becomes equal to that, okay. So we are going to show first that this quantity W become maximum when n1 becomes n by 2. So we need more space here and we're going to use this space here.

(Refer Slide Time: 10:29)



W is n factorial by n1 factorial and n minus n1 factorial. Now we have to maximise the W with respect to n1, right? So it is a problem of maximising any function FX with respect to X. And what we did that in that function we say that if it is maximum, then dW with respect to dn1 should be equal to 0 and that gives us either maximum or minimum, and by doing the second derivative we figure out whether it is maximum or minimum. Here we're just going to do the derivations of W.

Now, since this problem is associated with something called factorial, maximising W is difficult. But we know that maximising W is essentially same as maximising ln of W. So, we are going to do, we're going to maximise ln of W with respect to n1 and put it equal to 0. So what is ln of w? ln of W is then ln of n factorial by n1 factorial and n minus n1 factorial, so that is the ln.

Now, in this particular case we are going to use something called Stirling approximation. We are going to write it here and that is true for large n. Remember, so it is not true for all possible n's, it is only true for large N. So Stirling's approximation says that ln n factorial is nlnn minus n, we need to use that. And I will show you using Mathematica that this quantity that means ln n factorial divide by this right-hand side nlnn minus N is is approaching one when n becomes large.

Since we deal with the situations of very large n, since already specified but that's for small n most probable distribution is not equal to all possible distributions, only for large n most probable distribution is almost equal to all possible distribution, therefore we can use Stirling approximation because we are working in the regime of large n only.

So let us simplify ln W then little bit. It is ln n factorial minus ln n1 factorial minus ln n minus n1 factorial that is my ln W. So, I will simplify it further, so ln W is equal to nlnn minus n that is my first term, my second term then n1lnn1 minus instead of minus it will be plus n1 and my third term which is minus n minus n1 ln n minus n1 plus n minus n1. So now my W is complete, now what I'm going to do is that I'm going to take a derivative with respect to dlnW by dn1.

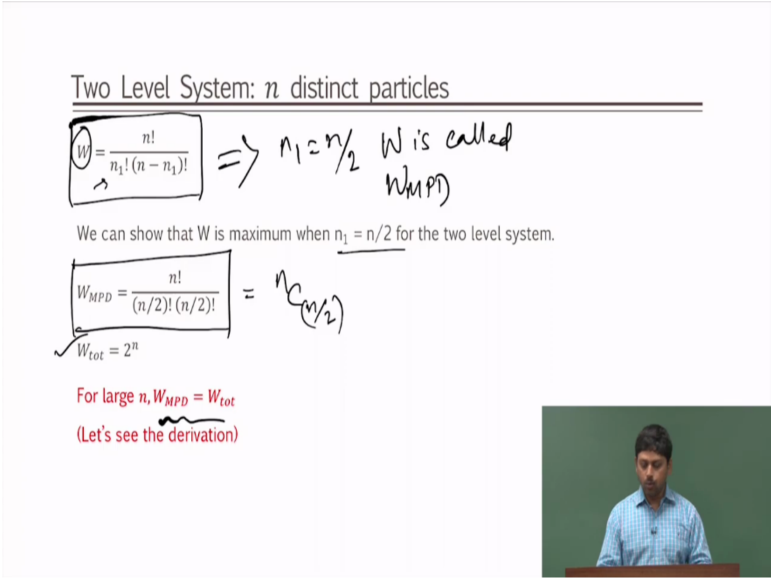
So the derivative of this term is going to be zero because n is a fixed quantity. So first term is zero minus second term is also zero now the third term, and now this term. So this is a product I can do as a product way, n1 into derivative of lnn1 is 1 by n1 minus lnn1 because derivative of n1 is 1, so I am just doing a jump but you can check it yourself. Now, derivative of n1 is 1, now this quantity, okay.

Now this quantity I have a minus here, this quantity this term I have a minus. Now, this minus has n minus n1, so first one will give me zero and second one will give me minus 1 and minus minus will actually become plus so plus was fine ln n minus n1. Now, I need to take a derivative of the second one. So, now I say n minus n1, ln of this quantity, right-hand side this particular quantity, which is 1 by n – n1 and then I have to take a derivative of the 1 that is a here.

So that will be -1. Now, the last term plus that will give me 0 minus 1. Now let us write it down what we got, 0 minus 0 is 0 so just I will not write it, first two terms I will not write it I will start from this term, so this n1 and this one cancels and gives me minus 1 minus ln n1 plus 1 plus lnn minus n1, this cancels and this minus and this minus cancels gives me plus 1 and minus 1.

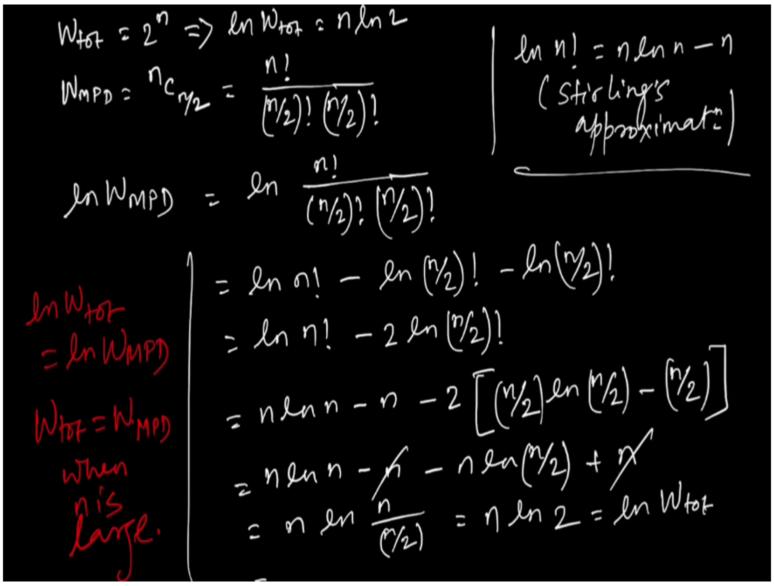
Let me see minus 1 plus 1 cancels minus 1 plus 1 cancels giving me I will have to have more space so I will use this side. So Dell lnW by Dell n1 is equal to minus ln n1 plus ln n minus n1, okay, which is we know that in that case it is nothing but n minus n1 by n1 and that is zero. So when is ln zero when this quantity is one. So, n minus n1 by n1 is equal to 1 which is n minus n1 is equal to n1 which is n equal to 2n1 and n1 equal to n by 2. So you can, we have seen that when n1 equal to nby 2 then W will be maximum.

(Refer Slide Time: 16:55)



So now we will go back to the point where. So we have shown that this is, we have shown that WMPD that means most probable distribution, that means the maximum value of W is when n equal to n1. So from here we get that, if n1 equal to n by 2 the W is called W MPD. Now, once we get that we have to show that WMPD. So WMPD is nothing but ncn by 2, that is nothing but 2 to the power n, so let us do that.

(Refer Slide Time: 17:32)



So, W tot equal to 2 to the power n that is all possible distribution, right. n distinguishable particles, 2 distinguishable boxes we know that how many possibilities will be there, 2 to the power n because 1st particle can be any of the 2 boxes 2 ways, 2nd particle can be any of the 2 boxes 2 ways and 2 the power n. It's like coin tosses also, so if I toss n coins, then the number of possible heads and tails the string can be there are 2 to the power n possibilities.

Now we are saying that MPD the most probable situations will be when there will be half head and half tail, half particle on this side half particle on that side equivalent statement. And when the n is large, then it will be almost it will be same. That means basically all possibilities and half head half tail are same. But that is very difficult to really imagine, we are going to show that through our derivations that that is indeed the case.

So, let us take from here ln of W tot which is nln2. Now, ln of MPD, no, first of all we write that MPD most probable distribution this is this and this is this. So, ln of WMPD is equal to ln of n factorial by n by 2 factorial multiplied by n by 2 factorial, equal to lnn factorial minus lnn by 2 factorial minus lnn by 2 factorial, which is lnn factorial minus 2 into ln n by 2 factorial.

Now, again we are talking about large n because it is not equal for a small values of n. And therefore you can use Stirling approximation, so I will remind you Stirling approximation is lnn factorial is nlnn minus n, which is called Stirling's approximation. As you have seen in my Mathematica file that they are actually said. So now let us see, so this is nlnn minus n minus 2 n by 2 lnn by 2 minus n by 2.

So nlnn minus n minus nln n by 2 minus minus plus n, cancels each other, I take n common, I get lnn by n by 2, which gives me nln2 isn't it same as ln W tot, which means then ln of W tot is equal to ln of WMPD, so therefore W tot equal to WMPD when n is large. So, it **it** clearly says that using one distribution which is the most probable distribution is enough to calculate the number of micro-states and therefore and in subsequently to calculate the entropy.

This is extremely helpful, without that our life would have been difficult and you will come to know later that why is that. So the question is that we have considered that the particles have no interactions and that is why they distribute equally. What if the particle would have had some interactions and you going to see later on that whether there is interaction between the particles or whether there is some preference of some boxes than the other box will actually change this statistics from n by 2 to n by 2 or that means equal to something else and that will become clear later on that how it is.