Introduction to Statistical Mechanics
Girish S Setlur
Department of Physics
Indian Institute of Technology – Guwahati

Lecture - 16
Ferromagnetism

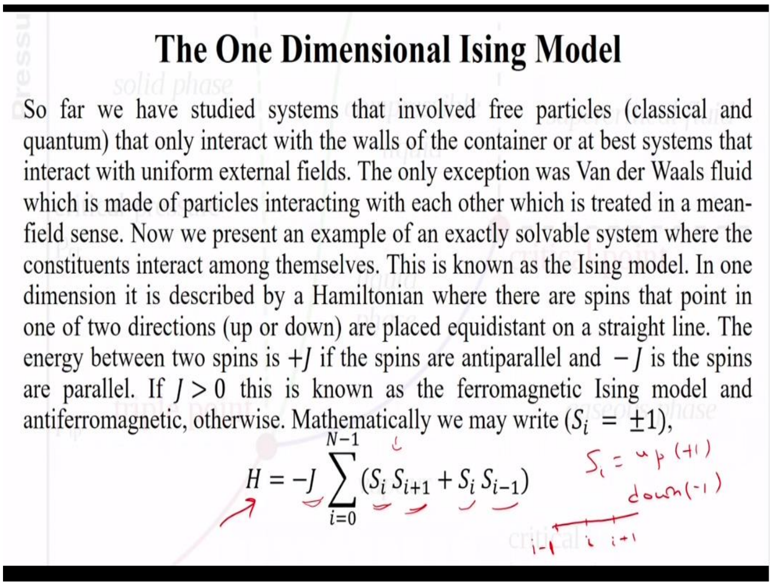
Alright, so we have discussed various types of magnetism till now. So, if you remember, we discussed Landau's diamagnetism, which is present in all materials. Then, we discussed Pauli’s paramagnetism, which is present in materials where the molecules or atoms have an intrinsic magnetic moment. So, now we are going to discuss the other, the third aspect of magnetism which is ferromagnetism.

So, ferromagnetism takes place, is also seen in materials where the magnetic moments not only interact with the external field, but they also interact with themselves, meaning with their neighbors in other words. So, one magnetic moment interacts with its neighbor. So, as a result, it becomes you know a more difficult problem to deal with because the constituent particles are interacting with each other in addition to interacting with the external environment.

So, we have seen for the most part in this course; we have restricted our attention to ideal systems where this does not happen. So, in other words, the constituent particles interact with the external environment only whether it is the walls of the container or the applied field. So, it does not interact with each other, they do not interact with each other. So, the exception of course was this van der Waals fluid which is basically composed of particles or atoms that they kind of have long-range interactions with each other.

So, in other words they are attractive is not necessarily long-range, but they are attractive interactions. So, as a result, they exhibit a richer phase diagram in the sense we saw that there are 2 types of phases, one is liquid, the other is gas. So, we expect something similar to happen in magnetism, when there is ferromagnetism we expect some richer phase diagram.

(Refer Slide Time: 02:45)



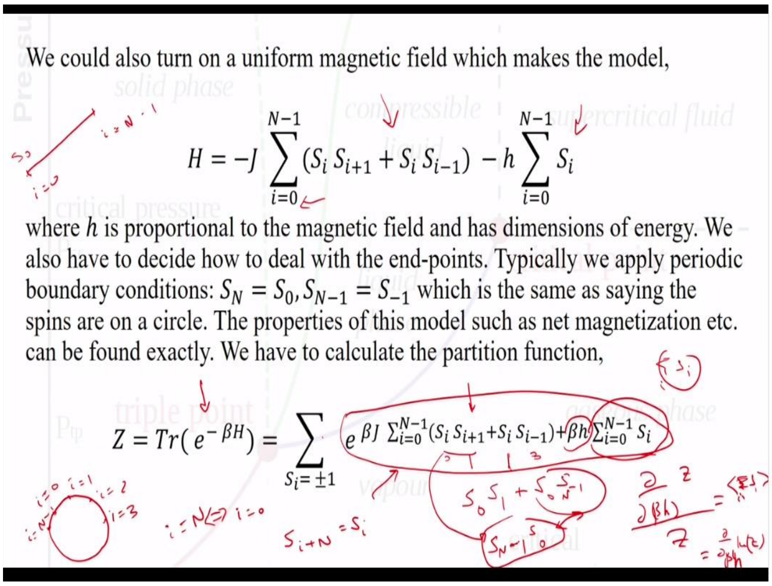
So, in order to model ferromagnetism, we have to take into account the fact that the spins are the magnetic moments coupled to each other in addition to the environment. So, the way you take into account the idea that spins coupled to each other, you start off with something very simple, so you want to minimize the complications caused by other issues. So, you want already there is this complication caused by the fact that the magnetic moments are coupled to each other.

So, you want to minimize the complications due to other factors. So, what you do is, you start off with a one-dimensional system. So, when all the magnetic moments are on a line, so that is the simplest thing you can think of. So, now so that is called a 1-D Ising model and that is described by this Hamiltonian. So, the idea is that there is energy or coupling energy between the spins S_i and S_j.

So, S_i is either up or down, so that is the whole idea. So, that is either up or down and then see a given, so that whole idea is that you have i and they coupled to their neighbors, so that on a line there are 2 neighbors, there is i – 1, i has 2 neighbors, i – 1 and i + 1. So, that is the reason why I have written it here. So, i couples to its right neighbor, which is i + 1 and i also couples to its left neighbor which is i – 1.

And of course, there is some kind of a double counting in the model, but you can always absorb that into J okay, I mean this is sometimes people do not write it this way, they just only do this because eventually it is going to take care of itself as you sum over i you run through all the neighbors anyway.

(Refer Slide Time: 04:46)



So, this is how it is, but the only point is I have to well okay. So, now suppose I will come to that later. So, suppose I in addition have a magnetic field in the Z direction, so the implication here is that S is + 1 and – 1 implies that some kind of a spin angular momentum, a spin half object which either points up or down. So, therefore think of it as Sz.

So, if that is the case, then if I apply a magnetic field in the Z direction, it is going to have a magnetic coupling strength energy which coupling with the magnetic field is going to be – µ times. So, I am going to observe all those µ's and all those multiplying constants into these magnetic fields h and J. So, you should get used to the fact that you know in as you mature in doing physics, you kind of don't you know make a big fuss about dimensions, units and so it is all implied that it is going to take care of itself.

Or you know how to resurrect them when needed. So, they are going to be subsumed into h and J. So, whatever other units that might be required or constants like Bohr magneton etc., they all assumed to be subsumed in J and H. So, the whole idea is now the Hamiltonian is now described in terms of the coupling between the magnetic moments and the coupling with the external field.

So, now the only point you have to address is that what do we do about the end point? So, remember that I said there is the spins lying on a one-dimensional lattice but then the lattice has we have assumed that it has N number of sites. So, I have started labeling it with i = 0.

So, if I label it with i = 0, it is going to go from i = 0 to 1, 2, 3, 4 all the way up to N – 1. So, then I have to ask myself what is going to happen after N – 1.

So, you can take the point of view that that is where it ends in and there is no, so the disadvantage to doing that the disadvantage to ending here is that you see this spin does not have a neighbor to interact with on the right and similarly this spin does not have a neighbor to interact with on the left. So, you will have a situation where the Hamiltonian looks very funny, so all the other spins are treated in a certain way but the endpoints are treated differently.

So, if you want to avoid that, the best way to avoid that is to put everything on a circle. So, if you put everything on a circle, then you end up writing this. So, then so N – 1 is the last point here, so i = N is equivalent to i = 0. So, you know there is a periodicity. So,

$$S_{i+N} = S_i$$

So, there is a periodicity which I am going to make use of. So, if you put everything on a circle that is called periodic boundary conditions, so this is what I am going to assume is the case okay. So, now with all those assumptions in place, now I am called upon to evaluate the partition or the canonical partition function of the system because that is the way to go and I am going to calculate $Tr(e^{-\beta H})$.

So, in other words, calculating trace simply means that you have to $Tr(e^{-\beta H})$ and sum over all the spin configurations. So, you see, remember that each spin at each site can be either + 1 or – 1. So, you have a whole bunch of different configurations. So, all the spins can be +1 or you know half of them can be + 1, half of them can be – 1 or the first one can be + 1, second one - 1 and the third can be + 1.

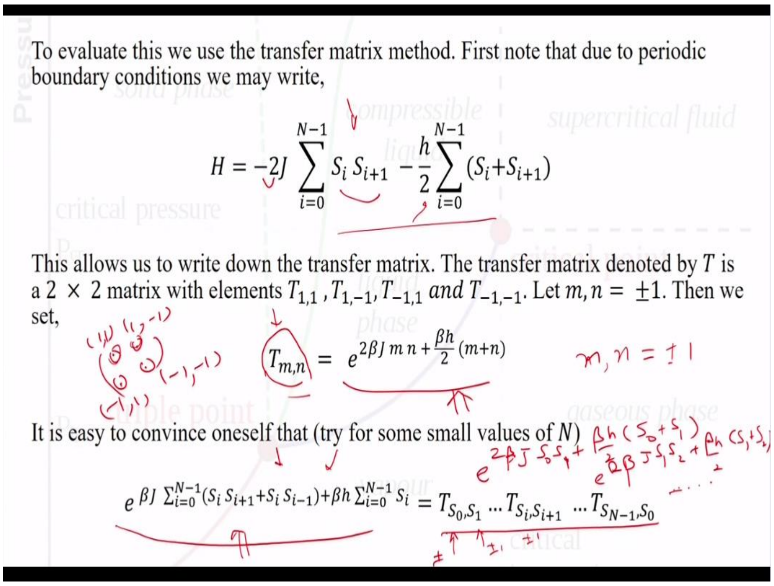
So, you have a whole bunch of different possibilities and you have to sum over all those possibilities. So, that is what trace means in this case. So, in order to do that it is not going to be easy mainly because there is this link. So, i is linked to i + 1 and also to i – 1. So, it is not going to be straightforward to do the trace that way. So, fortunately there is a simple way to do the trace in one dimension which I am going to explain now.

It is called the transfer matrix method. So, this is a peculiar method that works predominantly only in one dimension. It is not easily generalizable to more than one dimensions for reasons that I may or may not have time to go into. So, it so happens that in one dimension, there is a simple method to deal with this and find the trace in an exact analytical way. So, the question is how do you do that?

So, before you do that what we are going to do is that because we have imposed periodic boundary conditions, I am entitled to rewrite this in this fashion. So, I am going to ask you to convince yourself that this is the case. It is not surprising because you see if you have S_0, so it is this is going to become S_0 S_1 but then this is going to become S_0 S_{-1} but then remember that S_{-1} is the same as S_{N-1}.

So, each term so then you see then if this was N – 1, this is going to be N. So, if suppose this was N – 1, so this is going to be N – 1 + 1, so S_N but S_N itself is S_0. So, you will find that each such term appears twice. So, this will also appear twice. So, if I put here for example if I put i = 1, this will become 0. Here, if I put 0, this becomes 1. If I put 1, this becomes 0. So, each term appears twice.

(Refer Slide Time: 11:29)



So, as a result, I can write twice times i times i + 1, so because of periodic boundary condition that takes care of itself. The other thing is that this I can also rewrite as the average of i and i + 1 because of again because of periodic boundary condition the sum over S_i is the same as sum over i + 1 because you are going to come back to the same point. So, you should convince yourself that this is true.

And because both are the same, I am counting it twice, so I have to divide by 2 to get back my original result. So, now this form is very convenient for me to use what is known as the transfer matrix method. So, let me define what the transfer matrix is, the transfer matrix is a 2 by 2 matrix whose elements are whose indices are described by the pair of numbers m, n where m is either + 1 or - 1 and n is also either + 1 or – 1.

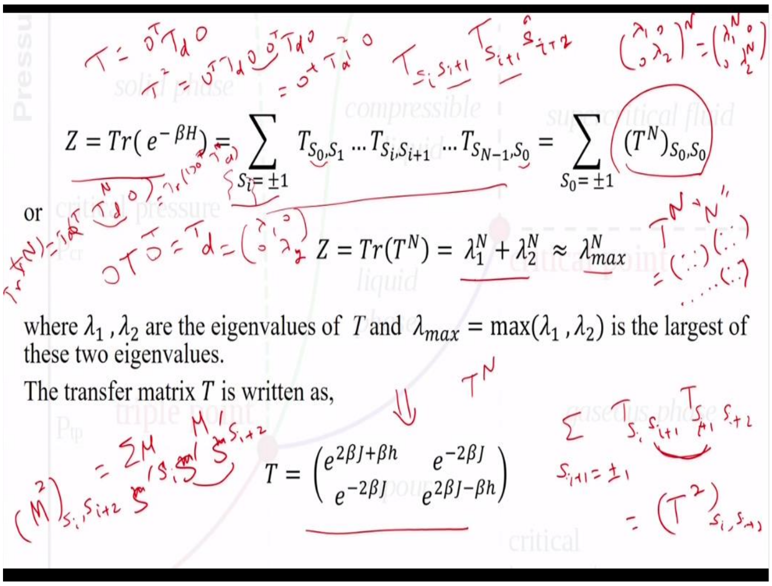
So, now the components of that matrix $T_{m,n}$ may be written in this fashion. So, this is by definition. So, this is the transfer matrix for this particular problem and you might be wondering why I did that because I mean superficially it is because it resembles $e^{-\beta H}$ but then I have to really convince you that this has any role to play in this problem. So, it does because this so recall that what we have to do is this is the quantity that I have to trace out over.

So, in other words, have to sum over all the spin configurations of this particular quantity and this particular quantity may be written in this fashion. So, you write this as the transfer matrix of S_0. So, remember that m and n are either + 1 or – 1, but so are all the S's, so the S_0 is either + 1 or – 1, S_1 is either + 1 or - 1 and S_i is either + 1. So, this makes perfect sense. So, I am going to convince you that it is possible to write this object in this fashion.

So, it is just a mathematical identity, there is nothing mysterious about this. It is just take this definition of $T_{m,n}$ and substitute the values that I have suggested here. So, that is going to be $\exp[2\ \beta J\ S_0\ S_1 + \beta h/2\ (S_0 + S_1)]\ \exp[2\ \beta J\ S_1\ S_2 + \beta h/2\ (S_1 + S_2)]$ and so on and so forth and then you will see that that is precisely this, this quantity, so it just you know you put them all in the exponent and sum over you get back this, so that is what this is.

So, you can write this quantity which is $e^{-\beta H}$ which is a function of a whole bunch of spin configurations in terms of whole bunch of transfer matrix components so namely this. So, the question is why is this useful? So, why would I be doing this? Because so in other words, we will be successful in doing this.

(Refer Slide Time: 15:08)



So, what we will have to do is that now that we have written $e^{-\beta H}$ as a product of a whole bunch of T's, so now we are tracing over the spin configurations simply implies summing over all the S's of this huge product of all the T's. So, now you see, look at any typical pair of T's sitting here in this huge product.

So, if you take for example, this one S_i and S_{i+1} and you look at the next one which is S_{i+1} S_{i+2}. So, then you see, you will be called upon or you are being called upon this is not just i but a whole bunch of I mean all of them right. So, this being summed over all the S's, so S_0, S_1, S_2 all the way up S_{N-1}. So, then you see when you sum over S_{i+1} also I mean you are going to be summing over all the S's including S_{i+1}.

So, now if you sum over the S_{i+1}, what does this tell you? I mean what does this resemble

$$\sum_{S_{i+1}=\pm 1} T_{S_i,S_{i+1}} T_{S_{i+1},S_{i+2}}$$

So, I am summing over these two. So, what this resembles basically is the matrix product. So, remember recall how you define the matrix product. So, if you have M^2 and I want to calculate you know some matrix elements of this quantity, so suppose I want to calculate the matrix element, this is going to be.

$$(M^2)_{S_i,S_{i+2}} = \sum_{S'} M_{S_i,S'} M_{S',S_{i+2}}$$

Then, something else, then something else, so in other words it is I have to do this right, I mean basically it is whatever this is S' if you like right. So, I have to sum over this S' and this

S' is precisely this S_{i+1} here. So, that is going to be and this this M is nothing but this T. So, this is nothing but

$$\sum_{S_{i+1}=\pm 1} T_{S_i,S_{i+1}} T_{S_{i+1},S_{i+2}} = (T^2)_{S_i,S_{i+2}}$$

So, as a result, what will happen is that, this when you do the sum over all this, all that is going to happen so if you sum over S_1, S_2, S_3 all the way up to $S_{N-1.}$

So, you leave out S_0 because they are at the extreme ends and you sum over all the S's in between, so what is going to happen is that you will end up with that quantity which is T^N , the component of T^N or the diagonal component of T^N at the column or the row number or equal to column number equals S_0. So, that is what it is going to end up being, but then eventually you will also have to sum over S_0.

$$Z = \sum_{S_0} T^N_{S_0,S_0}$$

Because you have called upon to sum over all the S's, so now if you sum over S_1, S_2, all the way up to S_{N-1}, you end up with this quantity with just T^N which is the transfer matrix, remember the T itself is a 2 by 2 matrix. So, T^N is the product of N 2 X 2 matrices. So, it is going to T^N is going to be like this. So, they are going to be N of these, so T^N is again a 2 X 2 matrix.

And what we have to do now is calculate the diagonal matrix element namely S_0, S_0 and then we have to finally sum over those S_0 as well. So, when you do that you get the answer for Z. So, I have not told you what this T is. I have told you what the components of T are. So, in other words, $T_{m,n}$ is this, so that means $T_{1,1}$ is m and n are + 1 or – 1. So, $T_{1,1}$ could be so you can think of this as 1, 1 1, – 1 – 1, 1 and – 1, - 1.

So, this is the labels of these rows and columns. So, $T_{1,1}$ would be first row first column, $T_{1,-1}$ would be first row second column, $T_{-1,1}$ would be second row first column and $T_{-1,-1}$ is second row second column. So, if that is the case, I have told you what all sits inside the small circles here and that is given by this $T_{m,n}$. So, that is what is going to sit here and then in that case if I decide to take these $T_{m,n}$'s and write it in a matrix form, it is going to look like this.

So, because when m and n are equal to + 1, it is going to look like this. If m is 1 and n is – 1, this is going to go away; it is going to become - 1 and so on and so forth. So, this is my T. So, now I am going to now being called upon to take T and raise it to the Nth power and so you might be wondering that is not easy because now there is a 2 by 2 matrix and there are no, all the entries are non-zero, but then I have to multiply N of them and I do not even know what N is, N is some huge number.

N is the number of sites. So, fortunately it is not T^N that I am interested in, but rather it is trace. So, in mathematics, there is a simple formula for the trace of a matrix and the trace of a matrix just happens to be. So, if you have a 2 X 2 matrix or a N X N matrix or whatever so if you have a square matrix, the trace of a square matrix is by definition the sum of all the diagonal elements.

But you can also convince yourself by diagonalizing that, so this is something you should ask your maths teacher to explain and that is that the trace of our square matrix is also the sum of all its eigenvalues. So, in other words, it is sufficient for you, so you do not have to really suppose if somebody tells you what the eigenvalues are, so you do not even know, you do not even have to know what that matrix is.

See the matrix itself will be quite huge. Suppose it is say M by M matrix, there will be M rows and M columns and M can be quite huge, it can be 1000 X 1000 matrix. So, you do not have to I mean strictly speaking somebody has to tell you all the diagonal elements and you have to add them all up but then it is possible that somebody has told you all the eigenvalues but they have not told you what the matrix itself is.

But then they want you to find the trace of that matrix. So, it is not going to be difficult because the trace of that square matrix is just the sum of all the eigenvalues. So, now it so happens that if you have T^N, it is very hard to write it down explicitly, but it is fortunately it is very easy to write down its eigenvalues. So, you cannot really write down the matrix itself very easily but you can write down its eigenvalues very easily.

And the reason is because the eigenvalues of, so if suppose T is diagonalizable, so in other words, what does that mean, what that means is that there is a orthogonal matrix okay, so that orthogonal matrix which can kind of render this diagonal. So, in other words, I can always find two matrices O and O^t such that I can rewrite this in this fashion and by definition, the diagonal entries are precisely the eigenvalues of T.

$$OTO^t = T_d$$

So, if that is the case, then you can see that T itself can be written as $T = O^t T_d O$. So, now therefore $T^2 = O^t T_d O \, O^t T_d O$. So, then OO^t is 1 because it is an orthogonal matrix and so this becomes $T^2 = O^t T_d^2 O$. So, then T^N, therefore can be written as $T^N = O^t T_d^N O$. So, you might be wondering why is this useful.

Because then the trace of this is nothing but trace of this and then because of the cyclic property of that trace, I can take this O on this side and this becomes Tr(O^t $T^N{}_d$). So, I am sorry about this messy. So, I will probably explain this, you know in greater detail, you know in the tutorial that is going to be appended to these lectures. So, right now, please bear with me.

So, I am going to assume you know some of this. So, this is linear algebra. So, I cannot teach you linear algebra in this course, so I am going to assume that you know some of this. So, this is just to you know refresh or jog your memory. So, this is one. So, then the trace of T raise to N is nothing but the trace of the diagonal version of T raise to N, but then the diagonal version of T is just the λ_1, λ_2 on the diagonal and zero on the off diagonal.

And then if it is very easy to raise a diagonal matrix to the Nth power, so it so happens that if you have λ_1 and λ_2, it is very easy to raise it to the Nth power and that is going to be λ_1^N and λ_2^N and 0. So, it is as simple as that. This you can easily verify. So, if you have a matrix like this where it is not diagonal, it is not easy to raise it to the Nth power.

But if you have a diagonal matrix, it is very easy to raise it to the Nth power. So, fortunately the Nth power of this T is same as the trace of the diagonal version of T raise to the Nth power. So, the diagonal version of T is basically the diagonalized T using an orthogonal transformation and so as a result, the trace of T^N is just the Nth power of the diagonal version of T which is $\begin{pmatrix} \lambda_1^N & 0 \\ 0 & \lambda_2^N \end{pmatrix}$ where you sum over all the diagonal elements.

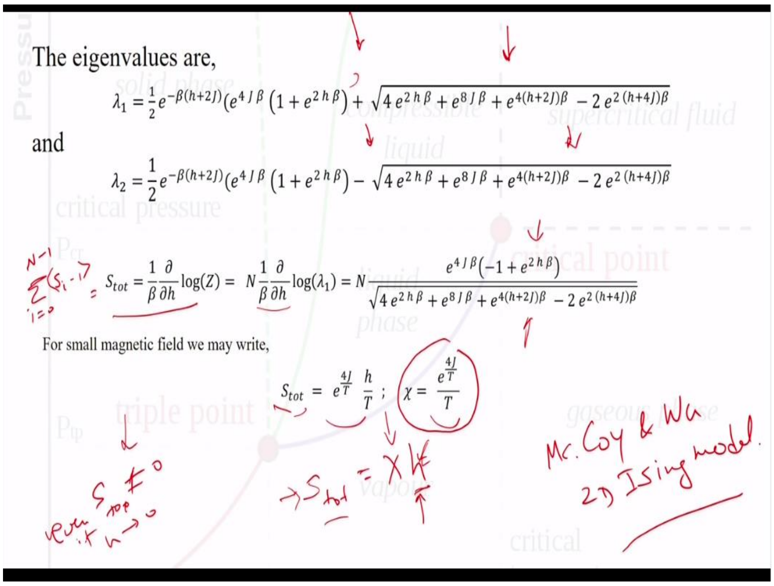
And thus be λ_1^N and λ_2^N added up. So, that is what you get for the partition function but recall that you know in this whole course, we are going to study systems in the thermodynamic limit namely that we are going to assume the system sizes are huge. So, as a result, it is only then you can apply effectively the methods of statistical mechanics because otherwise you will have to take into account fluctuations which we have ignored throughout this course.

So, remember that fluctuations may be ignored only if the system sizes are large because those fluctuations are of the order of the square inverse of the square root of the system size. So, if system sizes are not small, then you have to reckon with the fluctuations as well. So, we are going to assume system sizes are large, so I am going to ignore fluctuations and because system sizes are large, N is very huge.

So, even though there are 2 eigenvalues, the eigenvalue that is going to dominate is really the largest eigenvalue which is largest of these two. So, remember that this is this T itself is a symmetric matrix, so it has real eigenvalues and because the eigenvalues are real, one of them will be larger than the other and because one of them is the larger than the other, the larger one I can denote by λ_{max} and this is approximately λ_{max}^N N is very large.

So, that is the story of, well I am going to continue with the story, but basically this is how you find the partition function using the transfer matrix method for the 1-D Ising model and this transfer matrix method unfortunately works only in 1-D and with great difficulty in 2-D also but without a magnetic field and that is called the Onsager solution.

(Refer Slide Time: 28:48)



And I am not going to discuss that in this course unfortunately, although it is very interesting because it is somewhat technical alright. So, the Onsager solution is something you should you know you should muster some courage and look it up, so there is a book on the 2-D Ising model which can be solved exactly using the transfer matrix method and the book's title is 2-D Ising model and the author's names are McCoy and Wu.

So, these are the authors who have written a quite extensive treatise on the 2-D Ising model and I hope some enterprising students who are taking this course are going to you know get hold of this book and read it and maybe you know discuss it with me. So, I will be very happy to discuss the contents of this book with you alright. So, coming back to the 1-D Ising model, the eigenvalues of this transfer matrix are this and this.

So, it looks nasty, these days you know you do not have to do anything by hand. So, these computer algebra packages are there, my favorite is called Mathematica. So, that package I mean that is a software, which does you know I think of Mathematica as a calculator for doing calculus. So, normal calculators do arithmetic, but these kinds of calculators do calculus, they can do algebra, linear algebra, they can do whole bunch of things related to mathematics and allied disciplines.

So, it is worthwhile getting your hands on some of these packages and making use of them, that is what I did and the eigenvalues are λ_1 and λ_2 and of these the larger one is the one with a positive sign on the square root. So, I am going to use that and because then I can go ahead and so recall that suppose I wanted to find the average of this, what would I have to do?

I have to differentiate with respect to B, I have to differentiate Z with respect to βh. So, that is going to bring in total S_i downstairs there and then I divide by Z, so this is going to be the average of all the S_i. So, this is going to be the net magnetic moment. So, this is nothing but $\frac{\partial \ln(Z)}{\partial \beta h}$ okay. So, the average of all the sum over all the magnetic moments. So, this is the formula for the net magnetic moment okay.

So, $\sum_{i=0}^{N-1} < S_{i-1} >$ okay, that is what that is. So, if I differentiate with respect to h, it is going to look like this and then for of course I mean this is the general answer and you will have to deal with it, but it so happens that if h is small, I can expand in powers of h and write the net magnetization as a quantity proportional to h and the proportionality constant depends on temperature and that is called the ferromagnetic susceptibility basically the susceptibility of the system.

So, the question is it is, see ferromagnetism is a peculiar phenomenon where the spins spontaneously align with each other even in the absence of the magnetic field. So, you see, so now the question is how is that likely to happen? Suppose, so if you make the magnetic field smaller and smaller, so the point is that there should be some temperature at which S_{tot} is not 0 even though h -> 0.

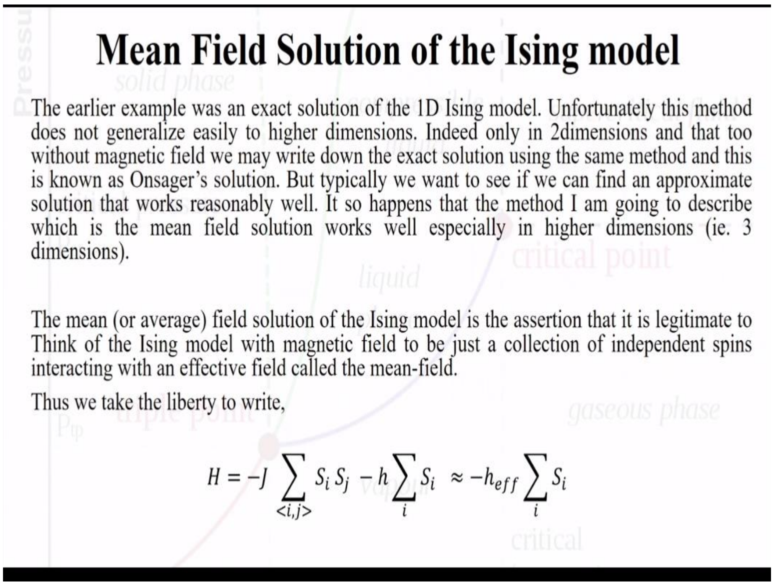
So, even if h -> 0, so even if h -> 0 , S_{tot} should not tend to 0. So, when can that happen? So, if h is very small, I have told you that we are in the linear regime whereas total is proportional to h. So, now it looks like when h -> 0, S_{tot} will also tends to 0. When is it going to happen that if h -> 0, S_{tot} need not tend to 0 that is going to be the case if χ itself becomes infinity for certain temperatures.

So, if for certain temperatures, if the susceptibility diverges, then there is a chance that you have ferromagnetism, but now if you ask yourself for what temperature does the susceptibility diverge, you will see unfortunately it only diverges at T = 0. So, that means it is only at absolute zero you can expect ferromagnetism. So, the 1-D Ising model is a disappointment even though we expected ferromagnetism.

Because we had incorporated you know spins that coupled to each other and we knew that that is source of ferromagnetism. Unfortunately, it led to an exact answer which shows that ferromagnetism is possible only at absolute zero which is an uninteresting result and inconsistent with what we see in nature. Of course, we do not see 1-D systems in nature anyway; it is not surprising we were disappointed.

So, what we really should be doing is studying something more realistic like a 2-D system. Unfortunately, then there is the transfer matrix method is either not feasible or sufficiently technical to prevent us from doing that. So, what is going to happen is that we will be forced to use certain approximate methods to study this problem, so to get an interesting result to show that there is ferromagnetism in such a system where the spins are coupled to each other in more than one dimension requires some approximate methods.

(Refer Slide Time: 35:33)



So, that is the topic of the next lecture, which is the mean field theory of the mean field solution of the Ising model. So, I am going to stop here right now. In the next class, I am going to discuss the mean field solution of the Ising model. So, I hope you enjoyed this lecture. So, let us meet for the mean field solution okay.