

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

**PROFESSOR:** OK, let's start. So last time we started thinking about phase transitions. We said that a simple example obtained by taking a piece of magnet such as iron, nickel and seeing what happens as a function of temperature.

And there's a phase transition between a paramagnet at high temperature and a ferromagnet at low temperatures. This transition takes place at a characteristic  $T_c$ . And a nice way to describe what was happening thermodynamically was to also look at the space that included a magnetic field. And then it was clear that various thermodynamic properties had a singularity, had a discontinuity at the line  $h$  equals to 0 for all  $T$  less than or equal to  $T_c$ .

Then we saw that if we looked at characteristic isotherms going from high temperatures going to low temperatures where there was a discontinuity, we more or less had to conclude by continuity that the one that goes along  $T_c$ , the magnetization as a function of field, has to come and hug the axis at 90 degree angle corresponding to having infinite susceptibility.

We also said that once you have infinite susceptibility, you can pretty much conclude that you are going to have long-range correlations across the sample. So that if you make a fluctuation here, the influence is going to be felt at large distances away. OK?

So given with that piece of knowledge, we said what we can do is to basically do some kind of averaging. I can take pieces of the sample. And at each piece, I can find locally what the magnetization is and have this field  $m$  as a function of  $x$  that varies from one point to another point. Presumably, close to this point either below or above, this  $m$  fluctuates across the sample over large distances and is typically

small. So then we use those pieces of information to proceed as follows.

We said all thermodynamic properties of the system can in principle be obtained by looking at a kind of partition function or Gibbs partition function that depends on temperature, let's say, which is obtained by tracing the Hamiltonian that governs all microscopic degrees of freedom that describe this system. The electrons, their spins, nuclei, all kinds of things.

Now naturally, this I cannot do. But I can focus on this magnetization field close to  $T_c$  and say that each configuration of magnetization has some kind of a weight. And in principle, what I can do is I can subdivide all configurations of microscopic degrees of freedom that are consistent with a particular macroscopic weight. Macroscopic field  $m$  of  $x$ . And hence, in principle compute what that weight is.

The analog of tracing over all degrees of freedom would now become integrating over all configurations of this magnetization that I indicate through this symbol of functional integration over all configurations.

Now, clearly I can no more obtain this than I can do the original trace. So what did we do?

We said I can guess what this is going to look like. Because in the absence of the field, it's a function that has rotational symmetry. So what I can do is I can write the log of that probability as something. So far I haven't done anything.

I made the assumption that I can write it as an integral over space of some kind of a density at each point. So that was this kind of quasi-locality assumption. And then I would write anything that comes to my mind that is consistent with rotational symmetry.

Now, since  $m$  is small in the vicinity of this point, it makes sense that I should make something like a Taylor expansion. So the Taylor expansion will start not with a linear term, which violates rotational symmetry, but something that is quadratic. And I can add any even power such as  $m$  to the fourth and higher order terms. And I can add all kinds of gradients that are consistent, again, with rotational symmetry. And

the first of those terms is gradient of  $m$  squared. And there are many more.

Of course, there can be an overall constant term. Maybe I will write it out front. Insert a  $Z$  regular, which means that you-- in the process, you have all kinds of other degrees of freedom that are not reflected in the magnetization. You're going to have phonon degrees of freedom. So there will be a contribution from the phonons to the partition function of the system.

We are not interested in any of those things. We are interested in what becomes singular over here. And the reason I write that as  $Z$  regular is because it's presumably some benign function of temperature as I pass through this point. So it is, indeed, a function of temperature.

It is also worth emphasizing that not only is this parameter that appears outside representing all kinds of other degrees of freedom a function of temperature, that these phenomenological parameters that I introduced here are also functions of temperature. Because the microscopic weight, the true Hamiltonian, is the one that is scaled by  $1$  over  $kt$ .

Just because of analogy, I sometimes call this combination of what is happening here  $\beta H$  or  $-\beta H$  as appearing in the exponent. But that, by no means, indicates that the coefficient here are scaling inversely with temperature like the true microscopic coordinates. Because in order to do this coarse graining, I have to integrate over a lot of different configurations. So there is energy associated with this. There is entropy associated with this. There is all kinds of complicated things that go into this-- these parameters. So that's important to remember.

Finally, if I slightly go away from here just to explore the vicinity of having a finite magnetic field, then I can work in the ensemble where I have added the field here. And the weight here will be modified by an amount that is  $h \cdot m$ . So what is occurring here is this Landau-Ginzburg model that I right now introduced in the context of magnetic systems. But very shortly, I will introduce this in the context of super-fluidity.

I can go back to the original example that we started with the liquid gas phenomena, replace  $m$  with some kind of a density difference, and then it would indicate a phase transition of the liquid gas system. So it is supposed to be very general, applicable to a lot of things because it is constructed on the basis of nothing other than symmetry principles. But at the cost of having no knowledge of how these phenomenological parameters depend on the microscopics of the system-- on temperature. Clearly, for example, even in the context of magnet, I would construct the same theory for iron, nickel, cobalt, et cetera. But presumably these coefficients will be very different from one system as opposed to the other.

All right, so that's supposed to be, as far as we are concerned, sufficient to figure out what is going on in the vicinity of this transition. Maybe I should emphasize one more thing, which is that I said that after all what we are trying to figure out is the nature of singularities in free energy, in phase diagrams, et cetera.

Yet, when I wrote this, I insisted on making an analytical expansion. And the reason making an analytical expansion here is justified is because to get this expansion, I summed over degrees of freedom-- I averaged degrees of freedom over some finite piece of my system.

Maybe I took a 100 by 100 by 100 Angstrom cubed block of material and averaged the magnetization of spins in that area, et cetera. And the idea that we also encountered last semester is that as long as you are dealing with a finite system, all of the manipulations that you are doing involve analytic functions such as  $e$  to the minus  $\beta H$  applied to a finite number of integrations. And you simply cannot get a singularity out of such a process.

So the averaging process that goes in this coarse graining must give you analytical functions. That's not the origin of the singularity. The origin of the singularity must come from taking the size of the system, the volume of this piece of iron, essentially to infinity. That's what gives us the singularity.

Also, because of that I expect that at the end of the day, this  $Z$  of  $T$  and  $h$  that I calculate will be something that is proportional to volume or something like that in

the exponent. That is, I will have something like  $e$  to the-- from this component minus  $V$  some  $\beta f$  regular. And actually, let me just forget about the  $\beta$  and just write it as some other function regular. It's an extensive quantity.

And the result of this integration of all of the field configurations should give me another contribution that is proportional to volume. And hopefully, all of the singularities that I expect will come from this piece.

And I am going to expect those singularities to arise only in the limit where  $V$  becomes large. Now, last semester we saw a trick that allowed us to do integrations when the final answer was extensive.

It was the saddle point result. Basically, we said that as we span the integration range, there is a part that corresponds to extremizing whatever you are integrating that gives you overwhelmingly larger weight than any other part of the integration. So let's try to, without justification-- and we'll correct this-- apply that same principle of a saddle point approximation to the functional integration that we are doing over here.

That is, rather than integrating over all functions, let's find the extremum. Where is this function maximized, the integrand is maximized?

So what I need to do is to find maximum of exponent. So what is happening over here.

I am going to make one further statement here before I do that, which is that this term, the thing that is proportional to gradient of  $m$ , has something to do with the way spins know about their neighborhood. And if I am looking at something that is ferromagnet, the tendency is for neighbors to be in the same direction. That's how ferromagnetism emerges. And that corresponds to having a  $K$  that is positive.

So that helps me a lot in finding the extremum of the function because it means that the extremum will occur when  $m$  of  $x$  is uniform across the system. Any variations in  $m$  of  $x$  will give you a cost from that and will reduce the probability. So with  $K$  positive, I know that the optimal solution is going to correspond to a uniform value

across the system. What is this uniform value?

If I put a uniform value, first of all, the integration will simply give me the volume. That's good because that's what I expect over here. So at the end of the day, what do I get?

I will get that  $F$  singular is, in fact, the minimum of  $\psi$  of  $m$  where  $\psi$  of  $m$  is obtained by simply evaluating what is in the integrand at a uniform value of  $m$ .  $t$  over  $2m$  squared  $\mu$   $m$  to the fourth. Potentially higher orders, but no  $K$ . The  $K$  disappeared. Minus  $h$   $m$ .

Actually, I kind of expect also that the uniform solution, if I put a magnetic field, will point along the magnetic field. If there's an up field, the uniform solution will be along the field. If it's 0, then I don't know. If it is down field, it will be the opposite. So it makes sense that if I have a field, the uniform magnetization should point along it, which means that over here this dot product is really just  $h$  times  $m$ . So I reduced the complexity of the problem subject to this saddle point approximation to just minimizing some function.

Now, it is important to note, since I was talking about analyticities, and singularities, and things like that, that whereas the function that is appearing here as  $\psi$  of  $m$  as we discussed is completely analytical, the operation of finding the minimum of a function is something that introduces non-analyticity. So that's how we are going to get that structure that we have over there. Let's explicitly show how that occurs.

So I'm going to show you the shape of this function in the space that is spanned by the parameters that I have here, which are  $t$  and  $h$ . We will come to  $u$  later on. But for the time being, since I want to show things in a plane, let's stick with  $t$  and  $h$ .

So first of all, let's look at the regime where  $t$  is positive. So this is 0. To the right is  $t$  is positive. Right on the axis when  $h$  equals to 0, the function is  $m$  squared plus  $m$  to the fourth, et cetera. The coefficient of  $m$  squared is positive. So basically, right on this axis the function is like a parabola. So what I am plotting here are different forms of  $\psi$  of  $m$ . So this is  $m$  squared is a parabola.

If I go further out,  $m$  to the fourth is something to worry about. But if I'm really looking at case where the emerging magnetization is small, I can stick just with the lowest order term in the expansion. Yes?

**AUDIENCE:** Are we plotting  $h$  or  $\psi$ ? The axis label was  $h$ .

**PROFESSOR:** OK. So there is a two-parameter plane  $t$   $h$ . At each point in this parameter plane, I can plot what  $\psi$  of  $m$  looks like. Unfortunately, I can't bring it out of the plane, so I am going to put it in the two dimensions. So this is why I wrote  $\psi$  of  $m$  with green to distinguish it with the  $h$  that is in white. So this is supposed to be right on the axis. If I go up on the axis, what would  $\psi$  of  $m$  look like?

So now I have a term that is positive  $h$   $m$ . So it starts linearly to go down, and then the  $m$  squared term takes over. So now the minimum is over here. Whereas, right on the axis-- again, by symmetry-- the minimum was at 0.

If I go and look at what happens at the other side when  $h$  is negative, then the slope is in the opposite direction. So the function kind of looks like this. And the minimum has shifted over here. So if I were to plot now what the minimum  $m$  bar is as a function of the field  $h$ , as I scan from some positive  $t$ -- from  $h$  positive to  $h$  negative, what we have is that this sequence then corresponds to a curve in which, essentially,  $m$  is proportional to  $h$ . Something like this. So this is for  $t$  that is positive.

Now, what happens if I try to draw these same curves for  $t$  that is negative? Let's stick first again to the analog of this curve, but now for  $h$  equal to 0. So for  $h$  equals to 0, in plotting  $\psi$  if  $m$ , whereas previously the parabola had the positive coefficient, now the parabola has a negative coefficient. So it kind of looks like this.

Now, one of the things that I know for sure in my system-- again, mathematics does not tell us that, physics tells us that. That I have a piece of iron and I know that the typical configurations of magnetization are some small number. So if the mathematics says that I have a function such as this, it is wrong because it would say that the extremum would go towards infinity. But that's where the  $u$ -term comes into play.

So I need now to have a  $u$ -term with a positive sign to ensure that the function does not have an extremum that goes to minus plus infinity. Physics tells us that the function that I get if  $t$  is negative should have a positive  $u$ . Now, what happens if we take this curve and go to  $h$  positive?

Well, now again, like we did over here, I start to shift it in one direction. In particular, what I find is that the minimum to the left goes up and the minimum to the right goes down. What did I do?

Drew this poorly. I start to go like this. So now there is a well-defined minimum over here.

If I go to negative fields, I will get the opposite Where I have this. So now if I follow a path that corresponds to these  $t$  negative structures, what do I get?

What I get is that the extremum is at some positive value. So this is for  $t$  that is negative. As I scan  $h$  from positive values, I am tracking this minimum. But that minimum ends up over here at 0. So basically, I come to here.

Whereas, if I come from the negative side, I am basically following the opposite curve. And then I have a discontinuity exactly at the  $h$  equals [INAUDIBLE].

Again, by continuity if I am exactly at  $t$  equals to 0, the function is kind of-- rather than the parabola is  $m$  to the fourth, and a little bit of thought convinces you that the shape of the curve kind of looks like this. We will quantify what that shape is shortly for  $t$  equals to 0. So what did we do?

We were able to reproduce exactly the structure of the isotherms that we are getting for the case of the ferromagnet from this simple theory. In order to just match exactly this with what is going on with iron or nickel, what do I have to do?

I just have to ensure that this parameter  $t$  goes to 0 at  $T_c$  of whatever the material is. Now, remember that I said that my  $t$  is really a function of temperature. So I can certainly make an expansion of it around any point. Let's see the  $T_c$  of the material. So what I require is that the first term in that expansion has to be 0. Then I will have



something that is linear in  $T$ , and then quadratic in  $T - T_c$  squared and so forth. So there is one condition that I have to impose, that that function of  $t$ -- which again, by all of the arguments that I mentioned, completely has to be an analytic function of temperature. Hence, expandable in a Taylor series. The 0 to order term in that Taylor series has to be 0. What else?

That I can expand any other function.  $U$  of  $t$  is  $u_0$  plus  $u_1 T - T_c$  and so forth.  $K$  of  $t$  is  $k_0$  plus  $k_1 T - T_c$  and so forth.

And I don't really care about any of these coefficients. The only things that I know are the signs.  $a$  has to be positive because the high temperature side corresponds to paramagnet.  $u_0$  has to be positive because I require the stability types of things that I mentioned. And  $k_0$  has to be positive because I want to have this kind of ferromagnetic behavior.

Apart from that, I really don't know much. So the Landau-Ginzburg Hamiltonian with one condition reproduces the phenomenology of the magnet and all the other phase transitions, which as one of you-- I think it was David-- was pointing out to me, is very much like dealing with a branch cut singularity in the mathematical sense along this axis terminating at that point. So this branch cut singularity is a consequence of this minimization procedure of a purely analytical function.

And the statement of universality at this level is that pick any analytical function and do this minimization. You will always get the same mathematical branch cut that we will now explain in terms of the exponents.

So we said that experimentally these phase transitions, their universality was characterized by looking at singularities of various quantities and looking at the exponent, the functional forms. So let's first look at the magnetization. So if I do the extremization of this, what do I do?

I have to have  $d\psi$  by  $dm$  equals to 0.  $d\psi$  by  $dm$  is  $t_m$  bar plus  $4u$   $m$  bar cubed minus  $h$ .

For  $h$  equals to 0 along the symmetry axis, then I have two solutions. Because that I can write as  $t$  plus  $4u$   $m$  bar squared times  $m$  bar equals to 0. If I state at that equation, I immediately see that the possible solutions are  $m$  bar is 0. And that's really the only solution that I have for any  $t$  that is positive.

While for  $t$  negative, in addition to the solution at 0, which is clearly unphysical because it corresponds to a maximum and not a minimum, I have solutions at  $\text{minus plus square root of minus } t \text{ over } 4u$ .

And so if I were to plot this magnetization as a function of  $t$ , essentially I have a kind of coexistence curve because I have nothing above and below I have a square root singularity. So this corresponds to the exponent  $\beta$  being  $1/2$ . So that's the prediction [? from each. ?] What about the shape of this green curve, the isotherm that you have at  $t$  equals to  $T_c$ ?

Well,  $t$  equals to  $T_c$ , in our language, corresponds to small  $t$  equals to 0, which means that I have to look at the equation  $4u$   $m$  bar cubed equal to  $h$  or  $m$  bar is proportional to-- well, let's write it.  $h$  over  $4u$  to the power of  $1/3$ . So this green curve that comes with infinite slope corresponds to a  $1/3$  singularity. The exponent  $\delta$  was defined to be the inverse of this, so this corresponds to having an exponent  $\delta$  that is 3.

We had behavior of susceptibility characterized by another set of exponent. Now, the susceptibility, quite generally, is the response of the magnetization if you change the field. And typically, we were interested in the limit where we measured a response if you are just at  $h$  equals to 0 and then you put a little bit more.

The equation that we have that relates  $h$  and  $m$  bar is simply that  $h$  equals to  $t$   $m$  bar plus  $4u$   $m$  bar cubed. Rather than taking  $dm$  by  $dh$ , let me evaluate its inverse, which is  $dh$  by  $dm$ .  $dh$  by  $dm$  is  $t$  plus  $12u$   $m$  bar squared. And so this inverse susceptibility, if I am for  $t$  positive,  $m$  bar is also 0. So the inverse susceptibility is  $t$ .

If I am for  $t$  negative,  $m$  bar squared is  $\text{minus } t \text{ over } 4u$ . I put a  $\text{minus } t \text{ over } 4u$  here. And it becomes  $t$  minus  $3t$ . So it becomes  $\text{minus } 2t$ . Again, nicely positive. Response

functions have to be positive.

If I were to plot this susceptibility, therefore as a function of temperature, or  $t$  that is related to  $T$  minus  $T_c$ , what do I get?

I will get a divergence that is inverse of  $1$  over  $T$  minus  $T_c$ -- on both sides. So basically, I get something like this.

And we said that the divergence of the susceptibility we characterize by exponent  $\gamma$ . And as I had promised, we explicitly see that, in this case, the  $\gamma$ s on both sides of the transition are the same and equal to unity. The inverse vanishes linearly, so the susceptibility diverges with unit exponent. But actually, we are making-- we can make an additional statement here that experimentalists can go and check that I hadn't told you before.

Now, you see in all of these other cases the only thing that I can say is universal side is the functional form. This is where the exponent  $\beta$  came from. But the amplitude of what is happening there, or the amplitude of what is happening here, these are things that depend on you and all of these things that I have no idea about.

Similarly here, because I don't know what the relationship between  $t$  minus  $T_c$  and the parameter  $t$  is, it involves this number  $a$  that I don't know of. But one thing that I notice is that the ratio of these two things is a pure number. So I say, OK, what you have is if you measure the susceptibility on the two sides of the transition, you will see amplitudes. And then  $T$  minus  $T_c$  absolute value to minus  $\gamma$  plus or  $\gamma$  minus.

I have told you that the  $\gamma$ s are the same. I don't know what the amplitudes are, but I can tell for sure that the ratio of amplitudes-- if this is the theory that describes things-- is a pure number of  $2$ . So that's another thing that you can go and say the experimentalists can check. They can check the divergence, and then see that the amplitude ratio is a universal object.

OK, there's one other response function that I had mentioned. There was the

exponent alpha that came from the heat capacity. So how do I calculate heat capacity? So the heat capacity, which is a function of temperature, let's focus only in the case where  $h$  equals to 0, is obtained by taking a temperature derivative of the internal energy of the system.

Now, the energy, on the other hand, is obtained by taking a  $d$  by  $d\beta$  of  $\log$  of the partition function. Now, I have all of my answers in terms of these parameters  $t$ ,  $u$ , et cetera. But I know that to lowest order, there is a linear relationship between small  $t$  and the real temperature that I have to put into these expressions.

So in particular, I can do the following. I can say that something like  $d$  by  $dT$   $d$  by  $d\beta$ , where  $\beta$  is  $1$  over  $kt$ , is something approximately when you look at the linear regime that is close to  $T_c$  of the order--  $1$  over  $\beta$  is going to be  $kb t$  squared. And then I would have  $d$  by  $dT$ , and then another  $d$  by  $dT$ .

I have to evaluate all of these things eventually in the vicinity of the critical point. Everything else is going to be a correction. So to lowest order, I will do that. And then I note that this is related-- derivatives between temperature and small  $t$  are related through a factor of  $a$ . So I do this and I put a squared up here. Doesn't matter.

The only reason I do that is because through the process that we have described here, I get an idea for what  $\log$  of  $Z$  is. So in particular,  $\log$  of  $Z$  has a part that comes from all of the regular degrees of freedom and a part that comes from this additional minimization that we are doing. So we have a minus  $V$  times the minimum of the function, which is  $t$  over  $2m$  bar squared plus  $u$   $m$  bar to the fourth when evaluated at  $h$  equals to 0.

So this is  $\log Z$ , which is some regular function of temperature, and hence  $t$ . Why is this part singular?

Because for  $t$  that is positive,  $m$  is 0. So this is going to give me 0 contribution for  $t$  that is positive. Whereas, for  $t$  that is negative, I have to substitute the value of  $m$  bar squared as I found above, which is minus  $t$  over  $4u$ . So I will get  $t$  over  $2$  times

minus  $t$  over  $4u$ . So that's going to give me minus  $t$  squared over  $8u$ .

Here I will have-- once I substitute that formula, plus  $t$  squared over  $16u$ . The overall thing would be minus  $t$  squared over  $16u$ . So I have to take two derivatives of this function. You can see that the function will give me 0 above for  $t$  positive and it will give me a constant for  $t$  negative.

So if I were to plot two derivatives of this function as  $T$  or  $T$  minus  $T_c$  is varied. Well, there is a background part that comes from all the other degrees of freedom. So there is basically some kind of behavior you would have had normally.

What we find is that above  $t$  positive, that normal behavior is the only thing that you have. And when you are below, taking two derivatives of this, something is added to this. So basically, the prediction is that the heat capacity of the system as a function of temperature should have a jump discontinuity.

Now, I said that in a number of cases we see that the heat capacity actually diverges and we introduce then exponent  $\alpha$  to parametrize that divergence. Since we don't have this divergence, people have resorted to indicating this with  $\alpha$  equals to 0. But since  $\alpha$  equals to 0 is ambiguous, putting a discontinuity in addition to be precise about what is happening.

So the predictions of the saddle point method applied to this field theory are the set of exponents and functional forms  $\beta$  equals to  $1/2$ ,  $\gamma$  as being 1, et cetera. A nice set of predictions. And of course, the test is, do they agree?

It turns out that there is one and only one case where you do the experiments and you get these precise exponents. And that's something like a superconductor. And the picture that I showed you last time for the gas, et cetera, corresponds to totally different set of exponents. So at this point, we have to face one of two alternatives.

One, the starting point is wrong. We put everything we could think of in the starting point. Maybe we forget something, but it seems OK. The other is, maybe we didn't do the analysis right when we did the saddle point approximation. And we'll gradually build the case that that is, indeed, the case and that we should treat the

problem in a slightly better fashion. Any questions? OK. Yes?

**AUDIENCE:** Just to remind me, the saddle point approximation was saying  $m$  was continuous, a continuous number across the substrate?

**PROFESSOR:** The saddle point approximation is to evaluate the functional integral, which corresponds to looking at all configurations, replacing that integration with the value of the integrand at the point that is most probable.

In this case, the most probable point was the uniform case. But maybe in some other case, the most probable configuration would be something else. The saddle point is to replace the entire functional integral with just one value of the integral. Yes?

**AUDIENCE:** So your second claims that analysis is probably wrong somewhere. It is most likely when we are trying to compute the energy of the system from all of field  $m$  and we just assumed something incorrectly, and that's why we get incorrect exponents?

**PROFESSOR:** No. My claim is that up to the place that I say Landau-Ginzburg, I have been extremely general. It may be that I missed something, but I will convince you that that's not the case. Then the line below that says saddle point approximation. My claim is that that's where the error came.

**AUDIENCE:** Also, we can do a similar kind of analysis for liquid gas transition in critical [INAUDIBLE]?

**PROFESSOR:** Yes.

**AUDIENCE:** And then, how would it be reasonable to assume uniform density? Because I guess the whole point of behavior--

**PROFESSOR:** We did exactly that approximation in 8.3.3.3. I wrote down some theory for the liquid gas transition out of which came the van der Waals equation.

**AUDIENCE:** Yes.

**PROFESSOR:** And the assumption for that was that the density in the grand canonical ensemble, in the grand canonical ensemble was uniform. So that you either got the density for the liquid or the density for the gas. But I made there the saddle point approximation also. I assumed that there was a uniform density that was--

**AUDIENCE:** OK. So was just likely to be the point where--

**PROFESSOR:** Yes.

**AUDIENCE:** --something breaks up.

**PROFESSOR:** Yes. So before doing that, let me point out to something interesting that happened. And it's just a matter of terminology.

Note that we constructed the Landau-Ginzburg Hamiltonian for  $h$  equals to 0 on the basis that we should have rotational symmetry. Nonetheless, even for  $h$  equals to 0, what we find are solutions where the magnetization is pointing in one direction or the other. So it is possible to have the state that emerges as a result of a weight that has some symmetry to not have that symmetry. So the symmetry is spontaneously broken and the direction in space is selected.

Now, of course, what that means is that if you apply the rotation operation to one of these ground states, then you will generate another equally good ground state. You can take everything that is pointed along the z-axis and make them all point along the x-axis. That's an equally good ground state.

So essentially, you have a manifold of possible states. And making a change from one state deforming to another ground state does not cost you any energy. So one consequence of that is that slow deformations should cost little energy. What do I mean by that?

So let's imagine that I start with a state where after I minimize, I find that all of my magnetizations are pointing up. Now, as I said, I could rotate everybody into this direction and the formation of my state would cost no energy. That's a uniform deformation. What if I took a deformation that is very slow? So I gradually rotate

from one to the other state.

Then in the limit where the wavelength of this deformation becomes of the order of the size of your system, you should have no energy cost. And it kind of makes sense that in the limit where you have long wavelengths, you should have little energy cost. So you should have slow wavelength, no energy distortions or modes, called Goldstone modes. But you can only have this for a broken continuous symmetry such as what I have depicted there, where all orientations are equally likely.

But if I had the liquid gas system, the density was either above average or below average. If I had uniaxial magnet, the spin would be either pointing up or down. Then I can't deform slowly from one to the other. So for discrete symmetries, you don't have these modes. For continuous symmetries, you have these modes.

And actually, we've already seen one set of those modes. These were the phonons. When in the first lecture I was constructing this theory of elasticity, I said if we take the whole deformation and move it uniformly, there is no cost. And then we were able, based on that, to conclude that long wavelength phonons have little cost. And we wrote their dispersion, relation, et cetera. So phonons are an example of Goldstone modes. These kinds of rotations of spins in a magnet-- magnons are another example of these modes.

But something else that we said, therefore in the first lecture, is something that we should start to think about. Which is that we said that because these modes exist and they have so little energy cost, if I am at some finite temperature, I will be able to excite them. So I know for sure that if I'm at finite temperature, there are at least these fluctuations that are going on in my system. And maybe in lieu of that, I should be wary of assuming that only the state where everything is uniform is the thing that is contributing. What about the fluctuations?

So let's think about these fluctuations that are easiest and most easily generated and look at their thermal excitations and consequences for the phase in phase transition. And let's do that in the context of superfluid.



So we saw the problem of super fluidity towards the end of 8.3.3.3. You had helium that was an ordinary liquid. We cooled it below 2.8 degrees and suddenly it became a new form of matter that has this ability to flow through capillaries, et cetera. And we pointed out that there was some kind of quite likely quantum origin to that because of the similarities that it showed to Bose-Einstein condensation.

And basically, it was in this context that Landau introduced something like this The theory that we write down where he chose as order parameter as the analog of the  $m$  of  $x$  that we have over there, a complex function  $\psi$  of  $x$ . And very roughly, you can regard this-- and again, this is very rough-- as overlap of wave function with the ground state at position  $x$  in some coarse-grained sense.

Now, anything quantum mechanical we saw has an amplitude and a phase. So this is actually a number plus a phase. Or if you like, it has a real part and an imaginary part. And there is no way that we know anything about the phase. The phase is not an unobservable. So the probability that when we scan the system we have identified some  $\psi$  of  $x$  that the probability should depend on the phase is meaningless. It's not an observable. So this functional should only depend on things like absolute value of  $\psi$ .

If, like Landau, we assume that it is a local form, then the kinds of terms that we can write are absolute value of  $\psi$  squared, absolute value of  $\psi$  to the fourth power. And the tendency for the order to expand across the system you would put through a term such as gradient of  $\psi$  squared.

Now, for the case of the superfluid, there is no physical field that corresponds to the  $\hbar$ . That just you don't have that field.

You can convince yourself that if you write  $\psi$  to be  $\psi_1$  plus  $i\psi_2$ ,  $i\psi_2$ -- real and imaginary part. And put that in this formula, that corresponds exactly to the theory that we wrote over there as long as we choose a two-component magnetization. So these two theories are identical.

Now, if I look at this system for  $t$  negative and try to find a minimum of the functional

that I have over there, then the shape of the function functional  $\psi$ -- poor choice of notation.  $\psi$  both being the wave function as well as the function that I have to extremize. But let's stick with it. It has a minimum that goes along a circuit.

So basically, take this picture that we have over here that corresponds to essentially one direction and rotate it. And what you will get is what is sometimes called the wine bottle type of shape, or the Mexican hat potential, or whatever. But essentially, it means that there is a ring of possible ground states. So the minimum occurs for  $\psi$  of  $x$  having some particular magnitude which corresponds to the location of this ring-- how far away it is from the center. And that will be given by the square root of  $t$  formula that we have up there.

But then there is a phase that is something that you don't know. Now, let's ask the question. Suppose I allow this phase to vary from one part of the sample to another part of the sample. So that's the analog of this slow distortion that I was making up there. So essentially, as I go from one part of the sample to another part of the sample, I slowly move around this bottom of this with potential. And I ask, what is the cost of this distortion that I impose on the system?

If I calculate  $\beta H$  for  $\bar{\psi}$  the  $i$   $\theta$   $x$ , what I get is whatever I have put over here, such as this function which minimizes the-- which is the location of the minimum. But because of the variation in  $\theta$  that I have allowed, there is a cost. So let's write that as  $\beta H_0$  plus this additional cost.

The additional cost comes from this term. If I simply put  $\bar{\psi}$  in to the  $i$   $\theta$  over there, I will get an integral  $d x k \bar{\psi}^2$  over  $2$  gradient of  $\theta$  squared. So there is an additional energy cost. This is very similar to the energy cost that we had for phonons. Because if I Fourier transform, you can see that I get a  $k$  squared just like we got for the case of phonons.

And just like for the case of phonons, I expect that at some finite temperature, these kinds of modes are thermally excited. So in reality, I expect that if I'm at some finite temperature, this phase will fluctuate across my system.

And maybe I should note that whereas I'm here thinking in terms of thermal fluctuations, by appropriate boundary conditions one can establish a gradient of theta that is uniform across the system. And that actually corresponds to a superfluid flow. So the case of a superfluid flow can be regarded by having a gradient of theta being proportional to velocity, and then this is something like the kinetic energy. But that's a different story. Let's just stick with the fact that this is the cost of making these distortions. And I want to know, what's the probability of having one of these distorted shapes? Yes.

**AUDIENCE:** Question. When you're introducing the psi as another parameter and you call it overlap of [INAUDIBLE] what is boundaries and what values this sort of parameter can take? So I basically wonder if we have this Mexican hat-shaped potential with minima on the ring of radius at 1, can the value of potential-- of the other parameter principally be further than that?

**PROFESSOR:** OK. So again, here we are trying to phenomenologically explain an observation that there is a transition between a case where there is no super fluidity and right when a certain small amount of super fluidity has been established in the system. The question that you asked over here is legitimate, but you could have asked it also for the case of the magnetization. So you could have said, why not to have that potential with the minimum somewhere else? But that does not explain the phenomena that we are trying to explain.

The phenomena that we are trying to explain is the observation that I go from having nothing to having a little bit of something. And I choose the mathematical form that is capable of describing that.

**AUDIENCE:** My question basically was, when we were talking about magnetization, if you take a piece of metal, you can magnetize it from 0 to a pretty large value.

**PROFESSOR:** No.

**AUDIENCE:** If we are interested in something not too large--

**PROFESSOR:** No.

**AUDIENCE:** And as I say, [INAUDIBLE].

**PROFESSOR:** What is your scale? What is very large? For a magnet, there is a maximum magnetization--

**AUDIENCE:** In this case, I mean that spontaneous magnetization for a magnet would be lower than saturation, no?

**PROFESSOR:** What is saturation?

**AUDIENCE:** When all spins are same direction.

**PROFESSOR:** OK. So you have a microscopic picture in mind. Now, the place that we are is far from that saturation magnetization. Similarly, in this case, presumably if I go to 0 temperature, there is some uniformity. And if I call this an overlap, the maximum of it will be 1.

**AUDIENCE:** Yeah. But basically, I just don't understand-- what is overlap of wave function?

**PROFESSOR:** Well, that's why I didn't want to go into that detail. But basically, the overlap of two wave functions would be the  $\psi_1^* \psi_2$  of  $x$ . And if you are thinking about the ground state, let's say that I have normalized this function to have a maximum of 1.

The point is that what that maximum is, is folded into all of these parameters--  $a$ ,  $u$ , et cetera-- and is pretty irrelevant to the nature of the transition.

**AUDIENCE:** OK.

**PROFESSOR:** OK? So the probability of a particular configuration of  $\theta$  across the system is given by this formula. I can unpack that a little bit better just like we did for the case of phonons by writing  $\theta$  in terms of Fourier modes.  $e^{i \mathbf{q} \cdot \mathbf{x} \theta(\mathbf{q})}$ , which for the time being, I assume I have discretized appropriate values of  $\mathbf{q}$ . I choose this normalization  $\sqrt{V}$  in this context.

If I substitute that over here, what I find is that  $\beta H$  as a function of the collection of  $\theta$   $\mathbf{q}$ 's is, again, some  $\beta H_0$ , which is not important. I put gradient of  $\theta$ .

Once I do gradient of theta, you can see that I get a factor of  $iq$ . I will have two of them, so the answer is going to be proportional to  $q$  squared.

Let's call this combination  $k \bar{\psi}^2 k \bar{\psi}$  so that I don't have to write it again and again.  $k \bar{\psi}^2$ . And then theta of  $q$  squared. OK?

So the probability of some particular combination of these Fourier amplitudes is proportional to exponential of this. And therefore, a product of independent Gaussian-distributed quantities.  $q$  squared. Yes.

**AUDIENCE:** When you plug in your summation of your-- well, your Fourier series into the gradient and then you square that, why don't you get interactions between the different Fourier amplitudes?

**PROFESSOR:** OK, let's do it explicitly. I have  $\int dx$  gradient of theta gradient of theta. Gradient of theta is  $i q \sum$ -- OK. Is  $i \sum$  over  $q$   $q e^{i q \cdot x}$  theta tilde of  $q$ . And I have to repeat that twice. So I have  $i \sum$  over  $q$  prime  $q$  prime into the  $i q$  prime dot  $x$  theta tilde of  $q$  prime. So basically, this went into that. This went into that. OK?

So I have a sum over  $q$ , sum over  $q$  prime, and an integral over  $x$ . What is the integral over  $x$  of  $e^{i q \text{ plus } q \text{ prime dot } x}$ ?

It is 0, unless  $q$  and  $q$  prime add up to 0. This is delta function. So you put it there.

Only one sum survives. Actually, I had introduced the normalizations that were  $\sqrt{V}$ . The normalizations get rid of this factor of  $V$ . That's why I had normalized it with the  $\sqrt{V}$ . And I will get one factor of  $q$  remaining. Since  $q$  prime is minus  $q$   $iq$  prime becomes  $q$  squared. And then I have theta tilde of  $q$  theta tilde of minus  $q$ , which gives me this.

So each mode is independently distributed according to a Gaussian, which immediately tells me the average of theta of  $q$  tilde is, of course, 0. Let's be careful and put the tildes all over the place. While the average of theta tilde of  $q$  squared is  $1 / k \bar{\psi} q$  squared.

Again, all that says is that as you go to long wavelength modes and  $q$  goes to 0, the fluctuations become larger because the energy cost is smaller. OK, so that's understandable. But now, let's look at what is happening in real space.

I pick two points,  $x$  and  $x'$ . And I ask, how do the fluctuations vary from one point to another point? So I'm interested in  $\langle \theta(x) \theta(x') \rangle$ . Let's say  $x' = x - \theta$ . Let's calculate the following first,  $\langle \theta(x) \theta(x') \rangle$  just because the algebra is slightly easier.

Now,  $\langle \theta(x) \theta(x') \rangle$  I can write in terms of  $\tilde{\theta}(q)$ . So this becomes  $\sum_{q, q'} \frac{1}{V} e^{iq \cdot x} e^{-iq' \cdot x'}$ . There is a factor of  $1/V$  that comes from the normalization I chose. And then the average of  $\tilde{\theta}(q) \tilde{\theta}(q')$ .

But we just established that the different modes are independent of each other. So basically, this gives me a delta function that forces  $q$  and  $q'$  to add up to 0. If they do add up to 0, the expectation that I get is  $1/k^2$ . And so what I find is that this becomes related to a sum over  $q$   $1/V e^{iq \cdot x - iq \cdot x'}$  divided by  $k^2$ .

If I go to the continuum limit where the sum over  $q$  I replace with an integral over  $q$ , then I have to introduce the density of states. And so then I find that  $\langle \theta(x) \theta(x') \rangle$  is  $1/k^2 \int \frac{d^d q}{(2\pi)^d} e^{iq \cdot (x - x')}$  the Fourier transform of  $1/q^2$ .

Now, the Fourier transform of  $1/q^2$  is something that appears all over physics. So let's give it a name. So we're going to call the integral  $\int \frac{d^d q}{(2\pi)^d} e^{iq \cdot x} / q^2$ . And let's put a minus sign in front of that. And I'll give it the name the Coulomb potential in  $d$  dimensional space.

And for those of you who haven't seen this, the reason this is the Coulomb potential is because if I take two derivatives and construct the Laplacian of that function-- so I take two derivatives with respect to  $x$  here, the two derivatives will go inside the integral. And what they do is they bring down two factors of  $iq$  divided by  $q^2$ .

The minus sign disappears,  $q$  squared over  $q$  squared goes to 1. Fourier transform of  $e$  to the  $iqx$  is simply the delta function. So this  $Cd$  of  $x$  is the potential that would emerge from a unit charge at the origin at a distance  $x$ .

So again, for those who have forgotten this or not seen it, let's calculate it explicitly using Gauss' theorem. The potential due to a unit charge is going to be spherically symmetric, so it's only a function of the magnitude of  $x$ . It doesn't depend on the orientation.

And Gauss' law states that the integral of Laplacian over volume is the same as the integral of the analog of the electric field, which is the gradient over the surface. So I have the surface integral of gradient.

Now, for the case that we are dealing with, the left-hand side Laplacian is a delta function. So when we integrate that over the sphere, I simply get 1. So this gives me 1. What do I get on the right-hand side?

It's just like the flux of the electric field that you have calculated. It is the magnitude of the electric field times the surface area. And I am doing this generally in  $d$  dimensions. So the surface area in  $d$  dimension grows like  $x$  to the  $d$  minus 1. And then there is a factor such as  $2\pi$ ,  $4\pi$ , et cetera, which is the solid angle that you would have to put in  $d$  dimensions.

And to remind you, the solid angle in  $d$  dimensions is  $2\pi$  to the  $d$  over  $2$   $d$  over 2 minus 1 factorial. So the magnitude of this derivative  $dC$  by  $dx$  following from that is simply  $1$  divided by  $x$  to the  $d$  minus 1, or  $x$  to the  $1$  minus  $d$  divided by  $Sd$ . So this is generalizing how you would calculate Gauss' law in three dimensions.

So now I just integrate that and I find that the  $d$  dimensional Coulomb potential is  $x$  to the  $2$  minus  $d$  divided by  $2$  minus  $d$   $Sd$ . And of course, there could be some constant of integration. So it reproduces the familiar  $1$  over  $x$  law in three dimensions.

But the thing that is important to note is how much this Coulomb potential depends

on dimensions. It determines these angle-- angle fluctuation correlations. And typically, here in this context we want to know something about large distances. If I make a fluctuation here, how far away is the influence of that fluctuation felt?

So I would be interested in the limit of this when the separation is large-- goes to infinity. And we can see that the answer very much depends on the dimensions.

So we find that for  $d$  that is greater than 2, like the Coulomb potential in three dimensions, you basically go to a constant. While in  $d$  less than 2, it is something that grows as a function of distance as  $x$  to the  $2 - d$ . And actually, write at the borderline dimension of  $d$  equals to 2, it also grows at large distances. If you do the integration correctly, you will find that it is  $1$  over  $2\pi$   $\log x$  over some distance or [INAUDIBLE].

Now, what you are really interested-- actually, this thing that I wrote down is not particularly meaningful. The thing that you are interested in is what I had originally written, which is that if I look at the angle that I have at  $x$  and then I go far away-- because the angle itself is not an observable, but angle differences are. So the average of this quantity will be 0. But there will be some variance to it, so I can look at this quantity.

And that quantity-- I can expand this-- is twice the average of  $\theta$  at some particular location. Presumably, it doesn't matter which location I look at. So it's the variance locally that you have in the angles. And then minus twice  $\theta$  of  $x$  minus  $\theta$  of  $x$  prime, which is the quantity that I calculated for you above.

So all I need to do is to take the Coulomb potential that I calculated, multiply it by a factor of 2, divide by a factor of  $k$  that I basically indicated as part of the definition. So this object is going to be  $2x^{2-d}$  divided by  $k$   $x^{2-d}$ . And actually, the reason I do this is because now I can indicate the overall constant as follows.

Remember that all of our statistical field theories are obtained by averaging. And I shouldn't believe any of these formulas when I look at very short wavelengths. So I



shouldn't really believe any answer that I got from those formulas when the points  $x$  prime and  $x$  come too close to each other. So there is something of the order of a lattice spacing, averaging distance, et cetera, that I'll call  $a$ .

And by the time I get to  $a$ , I expect that my fluctuations vanish because that's the distance over which I'm doing the average. So I manage to get rid of whatever this constant is by substituting [INAUDIBLE] with the scale over which I expect my theory to cease to be valid.

But again, what I find is that if I look at the fluctuation between two points at large distances, if I am in dimensions greater than 2, this thing eventually goes to a constant. Which means that if I'm in three dimensions, the fluctuation in phase between one place and another place are not necessarily small or large because I don't know what the magnitude of this constant is, but they are not getting bigger as I go further and further along.

Whereas, no matter what I do in  $d$  that is less than or equal to 2, this thing at large distances blows up. So I thought that I had a system where I had broken spontaneous symmetry and all of my spins, all of my phases were pointing in one direction. But I see that when I put these fluctuations, no matter how small I make the amplitude, the amplitude doesn't matter.

If I go to far enough distances, fluctuations will tell me that I don't know what the phase is from here to here because it has gone over many multiples of  $2\pi$ , so that it has become divergent. So what that really means is that because of fluctuations, you cannot have long-range order. So destroy continuous long-range order in dimensions that are less than or equal to 2. This is called the Mermin-Wagner theorem.

So you shouldn't have any, for example, super fluidity, magnetization, anything in two dimensions or one dimension. If you go to long enough, you will see that fluctuations have destroyed your order. So we can see already how important fluctuations are.

This  $d$  equal to 2 is called the lower critical dimension. It is this phenomena of symmetry breaking, ordering, phase transition, et cetera, that we are discussing for continuous systems-- for continuous symmetry breaking can only exist in three dimensions but not in two dimensions. We'll see that for discrete symmetries, you can have ordering in two dimensions, but not one dimensions. So there, the lower critical dimension is 1. Yes.

**AUDIENCE:** And does this hold for any  $n$ ? This example you were doing--

**PROFESSOR:** Yes.

**AUDIENCE:** --it just has two components.

**PROFESSOR:** Any  $n$ .  $n$  equals to 2, 3, 4, anything. We'll see later on, towards the end of the course, that there is a slight proviso for the case of  $n$  equals to 2, but that we'll leave for later.