

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.

PROFESSOR: OK, let's start. So we've been trying to understand critical points. And this refers to the experimental observation that in a number of systems we can be changing some parameters, such as temperature, and you encounter a transition to some other type of behavior at some point.

So the temperature, let's say, in this behavior is the control parameter. And you have to see, for example, this will be normal to superfluid transition. You have one now [INAUDIBLE] change temperature and going through this point.

For other systems, such as magnets, you actually have two knobs. There is also the magnetic field. And there, you have to turn two knobs in order to end at this critical point, also in the case of the liquid gas system in the pressure temperature plane, you have to tune two things to get this point. And the interesting thing was that in the vicinity of this point, the singular parts of various thermodynamic quantities are interestingly independent of the type of material.

So if we, for example, establish a coordinate at t and h describing deviations from this critical point, we have, let's say, the singular part of free energy as a function of t and h has a form like t to the 2 minus α , some scaling function h to the δ , and these exponents, α and δ , other things that are universal.

For example, we could get from that by taking two derivatives with respect to h , the singularity and the divergence of the susceptibility. And we said that the diverging susceptibility also immediately tells you that there is a correlation then that diverges, and in particular, we indicated its divergence to an exponent u . We could for that also establish a scaling form on how the correlation then diverges on approaching this point generally in the ht plane.

So this was the general picture. And building on that, we made one observation last time, which is that any point when you are away from h and t equals to 0, you have a correlation length. And then we concluded that if you are at t and h equals to 0, you have a form of scaling vertices.

And basically what that means is that when you are at that point, you look at your system, it's a fluctuating system, and the fluctuations are such that you can't associate a scale with them. The scale has already gone into the correlation length that is infinite.

And we said that therefore, if I were to look at some kind of a correlation function, such as a magnetization in the case of a magnet, that the only way that it became its separation is as a power of a distance. And this clearly has a property that if we were to rescale x and y by a certain amount, this correlation function nearly gets multiplied by a factor that is dependent on this rescale. And this is after we do the averaging, so it's a kind of statistical self-singularity, as opposed to some factor such as Sierpinski gasket, which are identically and deterministically self-similar in that each piece, if you blow it up, looks like the entire thing.

So what we have in our system is that if we have, let's say, a box which could be containing our liquid gas system at its critical point, or maybe a magnet at its critical point, will have a statistical field, this m of x . And it will fluctuate across the system. So maybe this would be a picture of the density fluctuation.

What I can do is to take a scan along some particular axis-- let's call it x -- and plot what the fluctuations are of this magnetization. Let's say m of x . Now the average will be 0, but it will have fluctuations around the average. And so maybe it will look something like this-- kind of like a picture of a mountain, for example.

Now one thing that we should remember is that this object would be piece of iron or nickel, and clearly I don't really mean that this is what is going on at the scale of a single atom or molecule of my substance. I had to do some kind of averaging in order to get the statistical field that I'm presenting here. So let's keep in mind that there is, in fact, some implicit analog of lattice size or some implicit shortest

distance, shortest wavelength that I allow for my frustrations.

Now I can sort of make this idea of scale invariance of a set of pictures, such as this one, more precise, as follows, by going through a procedure that I will call renormalization that has the following three steps. So the first step, what I will do is to coarse-grain further. And by this, I mean averaging m of x over a scale τa .

So previously, I had done my averaging of whatever means, et cetera. We're giving contribution to the overall magnetization over some number, let's 100 by 100 by 100 spins and a was my scaling distance. Why should I choose 100? Why not choose 200, some factor of what I had originally?

So coarse-graining means increasing this minimum length scale from a to $b a$. And then I define a coarse-grained version of my field. So previously, I had m of x . Now I have \tilde{m} of x , which is obtained by averaging, let's say, over volume around the point x that I had before. And this volume is a box of dimension $b a$ to the d .

And then I basically average over that. I guess let's call it original distance a equals 1, so I don't really have to bother by the dimensionality of y , et cetera. OK? So if I were to apply that to the picture that I have up there, what do I get? I will get an \tilde{m} as a function of x .

And essentially, let's say if I were to choose a factor of b that was like 2, I would take the average of the fluctuations that they have over 2 of those of those intervals. And so the picture that I would get it would be kind of a smoothed out version of what I have before over there. I will still have some fluctuations, but kind of ironed out.

And basically, essentially, it means that if you were to imagine having taken a photograph, previously you had the pixel size that was 1. Now your pixel size is larger. It is factor of b . So it's this kind of detuning and averaging of the fluctuations that has gone. And so you have here now b .

Now if I were to give you a photograph like that and a photograph like this, you would say that they are not identical. One of them is clearly much grainier than the other. So I say, OK, I can restore some amount of similarity between them by doing

a rescaling.

So I call a new variable x' to be my old variable x divided by a factor of b . So when I do that to this picture, I will get m as a function of x' . x' can go in further less, because all I do is I take this and squeeze it by a factor of b . So I will get a picture that maybe looks something like this.

Now if I were to look at this picture and this picture, you would also see a difference. That is, there is a contrast. So here, there would be, let's say, black and white. And as you scan the picture, you sort of see some variation of black and white.

If you look at this, you say the contrast is just too big. You have big fluctuations as you go across compared to what I had over there. So there's another step, which is called renormalize, which is that you define m' to be m by a change of contrast factor ζ . So you take a knob that corresponds to contrast and you reduce it until you see pictures that kind of statistically look like what you started with.

So in order to sort of generate pictures that are self-similar, you have this one knob. Basically, scaling variance means the change of size. But there is associated with change of size a change of contrast for whatever variable you are looking at. It turns out that that change of contrast would eventually map to one of these exponents that we have over there. Yes.

STUDENT: Are you using m or n tilda?

PROFESSOR: m tilda, thank you. So I guess the green is m tilda of x' , and the pink is m' of x' . So what I have done mathematically is as follows. I have defined an m' of x' , which is 1 over ζ , this contrast factor b to the d because of the averaging over a volume that involved b to the d pixels of the original field centered at a location that was $b x' + y$.

So in principle, I can go and generate lots and lots of configurations of my magnetization, or lots and lots of pictures of a system at the liquid gas critical point,

or magnetic systems at their critical point. I can generate lots and lots of these pictures and construct this transformation. And associated with this transformation is a change of probability, because there was some probability-- let's call it P old, that was describing my original configurations m of x . Let's forget the vector notation for the time being. Then there will be, after this transformation, probability that describes these configurations m prime of x prime.

Now you know that averaging is not something that you can reverse. So this transformation going from here and here, I cannot go back. There are many configurations over here that would correspond to the same average, like up, down or down, up would give you the same average, right? So a number of possibilities here have to be summed up to generate for you this object.

Now the statement of self-similarity presumably is that this weight is the same as this weight. You can't tell apart that you generated configurations before or after that scaling. So this is same at critical point. I've not constructed either weight, so it really doesn't amount to much.

But Kadanoff introduced this concept of doing this and thinking of it as a kind of group operation called renormalization group that I describe a little bit better and evolve the description as we go along. So if I look at my original system, I said that self-similarity occurs, let's say, exactly at this point that corresponds to t and h equals to 0.

Now presumably, I can, in some sense, force these things, if I were to take its log, for example. I can construct some kind of a weight that is associated with m , and this would be a new weight that is associated with m prime. Presumably, right at the critical point, these two would be the same weight, and it would be the same Hamiltonian.

What happens, if I do this procedure, to a system that is initially away from the critical point? So my initial system is characterized by deviations t and h from this scale in variant ways, which means that over here I have a correlation length.

Now I go through all of these transformations. I can do those transformations also for a point that is not at the critical point. But at the end of the day, I certainly will not get back my original weight, because I look at the picture after this transformation. Before the transformation, I had a long correlation length, let's say a mile. When I do this transformation, that correlation length is reduced by a factor of b .

So the new system has deviated more from the critical point. Because the further you go away from the critical point, you have a larger correlation length. So the idea is that right at the critical point, the two weights are the same. Deviation from the critical point is described by these two parameters, t and h .

And if you do the renormalization procedure on a Hamiltonian that deviates, you will get a Hamiltonian that more deviates, still describable by parameters t and h that have changed. So again, this says that c was, in fact, b times c of t prime and h prime, and t prime and h prime are further away.

Now the next thing that Kadanoff said was, OK, therefore there is a transformation that tells me after I do a rescaling by a factor of b how the new t and the new h depend on the old t and the old h . So there is a mapping in this space. So a point that was here will go over there. Maybe a point that is here will map over there. A point that is here will map over here. So there is a mapping that tells you how th get transformed under this procedure.

Actually the reason this is called a renormalization group, groups we are really thinking usually in terms of operations that are invertible. This transformation is not invertible. But this is a mapping. So potentially this mapping is invertible. You can say that if this point came from this point under inversion, it will go back to the original point, and so forth.

The next part of the argument is what did we do over here? We got rid of some short wavelength fluctuations. Now one of the things that I said right at the beginning was that as long as you are getting rid of short scale fluctuations, you are summing over a cube that has 100^2 , 200^3 . It doesn't matter, 100^3 , 200^3 -- you are doing some analytical function.

So the transformation that relates these to these, the old to new, should be analytical, and hence you should be able to write a Taylor series for it. So let's try to make a Taylor series for this. Taylor series start with a constant. But we know that the constant has to be 0 in both cases because the starting point was the point that was scale invariant and was mapping onto itself.

So the first thing that I can write down are linear terms. So there could be a term that is proportional to t . There could be a term that is proportional to h . There could be a term here that is proportional to h . There could be a term that is proportional-- well, let's call this t . Let's call this h . And then there will be terms that will be order of t squared and higher.

So I just did an analytical expansion, justified by this summing over just finite degrees of freedom at short scale. Now if I have a structure, such as the one that I have over there, I also know some things on the basis of symmetry. Like if I'm on the line that corresponds to h equals to 0, there is no difference between up and down. Under rescaling, I still don't know the difference between up and down.

So I should not generate an h if h was originally 0 just because t deviated from 0. So by symmetry, that has to be absent. And similarly, by symmetry, there is no difference between h positive and h negative. As far as t is concerned, h and minus h should behave the same. So this series should start at order of h squared and not h , so that term should be absent. So at this level, we have a nice separation into t prime is at and h Prime. Is dh .

Now we know something more, which is that the procedure that we are doing has some kind of a group character, in that if I, let's say, originally transform by some factor b_1 , change by a factor of 2, then change by a factor of 3, the answer is equivalent to changing by a factor of 2 times 3, or 3 times 2. Doesn't matter in which order I do them.

So also, I would get, if I were to do b_1 first and b_2 later, it would be the same thing. So what does that imply? That if I do two of these transformation, I find that my new

t is obtained in one case by the product, in the other case by the product of the two a's. So that's, again, some kind of a group character.

And furthermore, if I don't change the length scale, everything should stay where it is. So you glance at those, and you find that there is only one possibility, that a as a function of b should be b to some power. So you know therefore that at the lowest order under rescaling by a factor of b, t prime should be b to some y -- I called it y -- times t plus higher orders, while h prime is b to some other power of y times h plus higher orders.

And you say, OK, fine. What's this good for? Well, let's take a look at what we did over there. We said that I take some bunch of initial configurations, sum their weights to get the weight of the new configuration.

What happens if I sum over all initial configurations? Well, if I sum over all initial configuration, I will get the partition function. Now essentially, all the original configurations I regrouped and put into these coarse-grained configurations that are weighted this way.

So there could be an overall constant that emerges from this. But this really implies that the singular part of $\log z$, and presumably this depends on how far away I am from the critical point, is the same as $\log z$ that singular after I do this t prime and h prime.

Now there is one other issue, which is extensivity. Up to signs, factors of beta, et cetera, this is b times an intensive free energy, which is a function of t and h . So this is the same as v prime, because the volume shrunk. I took all of my scales and shrunk it by a factor of v , v prime, f of t prime and h prime.

So now let's go this way. Note that v prime is the original v divided by b to the d scaling factor. So you do the divisions here, and you find that f as a function of t and h is the ratio of v prime to v , which is b to the minus d , f as a function of t prime and h prime. But t prime we said to lowest order is b to the y t . h prime is b to the y h .

This is actually the more correct form of writing a homogeneous function. So

previously in last lecture, we assumed that the free energy had a homogeneous form. Now subject to these conditions and assumptions of renormalization group, we have concluded that it should have that homogeneous form.

Now you say this homogeneous form does not look like the homogeneous forms that I had written for you before. I say, OK. Presumably this is true for any factor of b that I want to choose. Let me choose a b , a rescaling factor such that v to the yt is of the order of 1. Could be 1, could be π . I don't care. Which means that I chose a factor of b that will scale with t as t to the minus 1 over yt .

I put this b -- this expression is true for all choices of b . If I chose that particular value, what I get is t to the d over yt , some function. First argument has now become 1 or some constant. Really it only depends on the second argument in the combination h and t to the power of yh over yt .

So you can see that this is, in fact, the same as the first line that I have above. And I have identified that $2 - \alpha$ would be related to this factor of yt , which is how you would scale under renormalization, the parameters t and h . And the gap exponent is related to the ratio of yh over yt .

Similarly, we had that the correlation length-- I have a line there. Ψ of t and h is b times Ψ of t' and h' . So I have that Ψ as a function of t and h is b times Ψ as a function of b to the yt , b to the yh . So that's also correct.

I can again choose this value of v , substitute it over there. What do I get? I get that Ψ as a function of t and h would be t to the minus 1 over yt , some scaling function-- let's call it g Ψ -- of, again, h to the power of yh over yt .

So I have got an answer that ν should be 1 over yt . I can get the scaling form for the correlation length. I identify the divergence of correlation length with inverse of this. And by the way, I get, if I substitute ν as 1 over yt here, the Josephson hyperscale in relation to minus α equals to b .

I can go further if I want. I can calculate magnetization as a function of t and h , would correspond to basically the behaviors that we identify with exponents β or

delta as $d \log z$. Yeah, let's say f by dh . If I take a derivative over there, you can immediately see that what that gives me is b to the power of y_h minus d , and then some scaling function which is the derivative of this scaling function b the y_t , b to the y_h .

And again, if I make this choice, then this goes over to t to the power of d minus y_h over y_t , and then some scaling function of h t to the δ . So I can continue with my table. And for example, I will have β to be d minus y_h divided by y_t . I can go and calculate delta, et cetera.

Actually I was a little bit careless with this factor ζ , which presumably is implicit in all of these transformations that I have. And I have to do special things to figure out what ζ is so that I will get self-similarity right at the critical point. But we can see that already we have the analog of a rescaling for m . And so it is easy to sort of look at those two equations and identify that my ζ should be precisely this one.

So the ζ is not independent of the relevance of the magnetic field. And if you think about it, the field and the magnetization are conjugate variables in the sense that in the weight here, I will have a term that is like hm -- integrated, of course. And so hm integrated, you can see that up to a factor of b to the d from integration, the dimensionality that I assign to h and the dimensionality that I assign to m should be related. And not only for the magnetization, but for any pair of variables that are so conjugate-- there's some f , and there's some x -- there will be a corresponding relation between what would happen to this x at the critical point and this factor f when I deviate from the critical point.

So all of this is kind of nice, but it's a little bit hand waving. I essentially traded one set of assumptions about homogeneity and scaling of free energy correlation length to some other set of assumptions about two parameters moving away from a scale invariant critical point. I didn't calculate anything about what the scale invariant probability is. I didn't show that, indeed, two parameters are sufficient, that this kind of scaling takes place, et cetera. So we need to be much more precise if we want to do, ultimately, calculations that give us what these numbers y_t and y_h are. So let's

try to put this hand waving on a little bit more firm setting.

So let's see how we should proceed. We start with some experimental system, critical point. So I tell you that somebody in the experiment, the liquid gas system, they saw a diverging correlation length, critical opalescence, et cetera.

So then I associate with that some kind of a statistical field. And let's kind of stick with the notation that we have for the magnet. Let's call it m of x .

And in general, this would be the part where one needs to put in a lot of thinking. That is, the experimentalist comes and tells you that I see a system that undergoes a phase transition. There are some response functions that are divergent, et cetera. You have to put in some thinking to think about what the appropriate order parameter is.

And based on that order parameter or statistical field, you construct the most general weight consistent with symmetries, with not only asymmetries but the kind of assumptions that we have been putting in play. So we put in assumptions about locality, symmetry. Stability is, of course, paramount. But there is a list of things that you have to think about.

So once you do that you say, OK, I associate with my configurations m of x some set of probabilities. Probabilities are certainly positive. So I can take its log, call its minus its log to be some kind of a weight, βh , that governs these m of x 's.

If I say that I'm obeying locality, then I would write the answer, for example, like this. But it doesn't have to be. I have to write some particular example. But you may construct your example depending on the system of interest.

And let's say we are looking at something like a superfluid, maybe, that we don't even have the analog of magnetic field, and we go and construct terms that are symmetric and made for a two component m . And I will write a few of these terms to emphasize that this is, in principle, a long list. There's coefficient of m to the sixth. We saw that the gradient terms could start with this k . But maybe there's a higher order gradient, and there's essentially an infinity of terms that you can write down

that are consistent with these assumptions that you have made so far.

So you say, OK. Now I take this, and implicit in all of these calculations is, indeed, some kind of a short scale cutoff. To construct the statistical field, I do apply the three steps of RG-- renormalization group, as I described before. And this will give me a new configuration for each of the old configurations through the formula that I gave you over there.

So in principle, this is just a transformation from one set of variables to a new set of variables. So if I do this transformation, I can calculate the weight of the new configurations, m' of x' . I can take minus the log of that.

And again, up to some constant, it will be the same as a probability. So there could be, in this procedure, some set of constants that are generated that don't depend on m . And then there will be a function that depends on m' of x' .

Now the statement is that since I wrote the most general function over here, whatever I put here will have to have exactly the same form, because I said put anything that you can think of that is consistent with symmetries over here. So you put everything there. What I put here should have exactly the same functional form, but with coefficients that have changed.

So you basically prime everything, but you have this whole thing. Now this may seem like truly difficult thing. But we will actually do this. We will carry out this transformation explicitly in particular cases.

And we will show that this transformation amounts to constructing a rescaling of each one of these parameters-- t' , u' , v' , k' , l' , and so forth-- as functions of the old parameters. So this is, if you like, a mapping. You take some set of parameters-- t , u , v , k , l , blah, blah, blah-- and you construct a mapping, s' , which is some function of the original set of parameters.

So this is a huge dimensional space. Any points that you start on the transformation will go to another point. But the key is that we wrote the most general form that we could, so we had to stay within this space.

So why are you doing this? Well, I started by saying that the key to this whole thing is have to having a handle as to what this self-similar scale invariant probability is. I can't construct that just by guessing. But I can do what we usually do, let's say, in constructing wave functions in quantum mechanics that have some particular symmetry.

Maybe you start with some wave function that doesn't have the full symmetry, and then you rotate it and rotate it again, and you average over all of them, and you end up with some function that has the right symmetry. So we start with a weight that I don't know whether it has the property that I want. And I apply the action of the group, which is this scaling variance, to see what happens to it under that transformation.

But the point that I am interested, or the behavior that I am interested, is where I basically get the same probability back. So I'm very interested at the point where, under the transformation, I go back to myself. And that's called a fixed point.

So S is a shorthand for this infinite vector of parameters. I want to find the point s^* in this parameter space. Actually, let me call this transformation R and indicate that I'm renormalizing by a scale b , such that, when I renormalize by a scale b , my original set of parameters, if I am at this fixed point, I will end up at that point.

So clearly, this is a system that has exactly these properties that I was harping in at the beginning. This is the point that is truly scale invariant. That's the point that I want to get at.

So again, once we have done this transformation in a specific case, we'll figure out what this fixed point is. But for the time being, let's think a little bit away from this and deviate from fixed point.

So I start with an initial point S that is, let's write it, $S^* + \text{a little bit away}$. Just like in the picture that I have here, I started with a fixed point, and I said I go away by an amount that I had parameterized by t and h . Now I have essentially a whole

line of deviations forming a vector.

I act with R_b on this, and I note that if δS goes to 0, then I should go back to S^* . But if δS is small, maybe I can look at the $\delta S'$, which is a linearized version of these transformations. So basically these transformations are highly nonlinear just as the transformation over here, in principle, would have been highly nonlinear. But then I expanded it around the point t and h equals to 0.

Similarly, I'm assuming that this δS is small, and therefore $\delta S'$ can be related to δS through the action of a matrix that is a linearized version. Let's call it here RL of b . So this is a linearized transformation, which means that it's really a matrix.

In this particular case, in principle, I started with a 2 by 2 matrix. The off diagonal terms were 0, so it was only the diagonal terms that mattered. But in general, it would be a matrix, which would be the square of whatever the size of the parameter space is that I am looking at.

Now then you have a matrix, it's good always to think in terms of its eigenvalues and eigendirection. In this problem that I had over here, symmetries had already diagonalized the matrix. I didn't have off diagonal terms. But I don't know here. It could be all kinds of off diagonal terms. So the properties are captured by diagonalize, RL , which means that I find a set of vectors in this space-- let's call them O_i -- such that under action of this, I will get λO_i , λ_i . Of course, the transformation depends on the rescaling parameter, so there should be a b here.

Now of course, you will get a totally different matrix for each b . So is it really hopeless that for each b I have to look at a new matrix, new diagonalization, et cetera? Well, exactly this thing that we had over here now comes into play, because I know that if I make a transformation size b_1 followed by a transformation size b_2 , the answer is a transformation size b_1, b_2 . And it doesn't matter in which order I do it.

AUDIENCE: Can't you just mix notation, because L used to be [INAUDIBLE]?

PROFESSOR: Sorry. So in particular, I see that these linearized matrices commute with each other for different values of b . And again, from your quantum mechanics, you probably know that if matrices commute then they have the same eigenvectors. So essentially, I was correct here in putting no index b on these eigenvectors, because it's independent of eigenvector, whereas the eigenvalues, in principle, depend on b . And how they depend on b is also determined by this transformation, that is λ_i of b_1 , λ_i of b_2 should be the same thing as λ_{b_1} , λ_{b_2} .

And of course, λ_i of 1 should be 1. If you don't change scale, nothing should change. And this is exactly the same set of conditions as we have over here, which means that we know that the eigenvalue's λ_i can be written as b to the power of some set of y_i . So we just generalized what we had done before, now to this space that includes many parameters.

So the story is now something like this. There is this multi-dimensional space with lots and lots of parameters-- t , u , v , blah, blah, blah, many of them. And somewhere in this space of parameters, presumably there is a fixed point, S^* . Now in the vicinity of that S^* , I have established that there are some particular directions that I can obtain by diagonalizing this.

So let's imagine that this is one direction, this is another direction, this is a third direction. And that if I start with a β h-- well, actually, let's do this. That is, if I start with an S that is S^* plus whatever is a projection of my components are along these different dimensions, so let's call them, let's say, a_i along these \hat{O}_i -- just make sure we kind of think of them as vectors-- that under rescaling, I will go to S' , which is S^* plus sum over i , $a_i b^{y_i} \hat{O}_i$.

That is, some of these directions, the component will get stretched if y_i is positive. It will get diminished if y_i is negative. And so now some notation comes into play.

If y_i is positive, the corresponding direction is called relevant. Eigendirection is relevant. If y_i is negative, the corresponding eigendirection is irrelevant. And very

occasionally, we may run into the case where y_i is 0. And there is a terminology. The corresponding eigendirection is marginal.

And what that means is that I need to resort to higher order terms to see whether it is attracted or repelled by the fixed point. So we need higher orders. After all, so far I have only linearized the transformation.

Now the set of irrelevant directions to this particular S fixed point, S star, defines basing of attraction of S star. So let me go back to the picture that I have over here and be precise and use the arrow going away as an indication that the corresponding b is positive, and I'm forced out along this direction. Let me choose going in as an indicator that the corresponding y is negative. And as I make b larger and larger, I shrink along this axis.

So in this three dimensional representation that I have over there, I have one relevant direction and two irrelevant directions. The two irrelevant directions will define the plane in this three dimensional space, which is the basing of attraction. So basically these two define a surface, and presumably any point that is in this surface in the three dimensional picture under looking at larger and larger things will get attracted to the fixed point. If you are away from the surface, maybe you will approach here, and then you will be pushed out. All right, fine.

Now let's go and look at the following. We have a formula, ψ of t and h . Or quite generally, ψ under rescaling is b times the new ψ . Or the new ψ under any one of these transformation, ψ prime, is the old ψ divided by the old correlation length divided by a factor of b .

So if I look at the fixed point-- so if I ask what is ψ at the fixed point-- then under the transformation, I have the same parameters. So ψ at the fixed point should be the ψ of the fixed point divided by b . There are only two solutions to this. Either ψ of S star is 0 or ψ of S star is infinite.

Now we introduce physics. ψ being 0 means that I have units that are completely uncorrelated to each other. Each one of them does whatever it wants.

So this describes essentially, let's say, a system of infinite temperature. Every degree of freedom does whatever it wants. Well, I should say this corresponds to disordered or ordered phases. Because after all, we said that when we go to the ordered states also, there is an overall magnetization, but fluctuations around the overall magnetization have only a finite correlation length. And as you go further and further into the ordered phase, that correlation length shrinks to 0.

So there is a similarity between what goes on at very high temperature and what goes on at very low temperature as far as the correlation of fluctuations is concerned. There is, of course, a long range order in one case that is absent in the other. But the correlation of fluctuations in both of those cases basically becomes finite, and under rescaling, goes all the way to 0. And clearly this is the interesting case, where it corresponds to critical point.

So we've established that, once we found this fixed point, that those set of parameters are what can give us the scale invariant behavioral that we want. Now this list is hundreds of parameters. So this point corresponds to a very special point in this hundreds of parameter space.

So let's say there is one point somewhere there which is the fixed point. And then you take your magnet and you change your critical temperature, are we going to hit that point? The answer is, no. Generically, you are not going to hit that point.

But that's no problem. Why? Because if this basing of attraction. Because for any point on basing of attraction, I do rescaling, and I find that ψ prime is ψ over b . It becomes smaller. So you generically tend to become smaller.

But ultimately, you end up at this point. And this point, the correlation length is infinite. So any point on this basing of attraction, in fact, has infinite correlation length. So every point on the basis of ψ prime equals to ψ , and hence ψ has to be infinite. Yes.

AUDIENCE: Question. Why should there be only one fixed point?

PROFESSOR: There is no reason.

AUDIENCE: OK. So this is just an example?

PROFESSOR: Yeah. So locally, let's say that we found such a fixed point. Maybe globally, there is hundreds of them. I don't know. So that will always be a question in our minds. So if I just write down for you the most general set of transformations, who knows what's happening?

Ultimately, we have to be guided by physics. We have to say that, if in the space of all parametrization, there are some that have no physical correspondence, we throw them out, we seek things that can be matched to our physical system. Yes?

AUDIENCE: If there are multiple fixed points, do the planes of the basins of attraction have to be parallel to each other?

PROFESSOR: They may have to have some conditions on non-intersecting or whatever. These are only linear in the vicinity of the fixed point. So in principle, they could be highly curved surfaces with all kinds of structures and things that I don't know. Yes?

AUDIENCE: Is there any reason why you might or might not have an attracting point that is actually a more complicated structure, like say, a limit cycle or even a [INAUDIBLE]?

PROFESSOR: Yeah. So again, we are governed ultimately by physics. When I write these equations, they are as general as equations as the people in dynamical systems use that also includes cycles, chaotic attractors, all kinds of strange things. And we have to hope that when we apply this procedure to an appropriate physical system, the kind of equations that we get are such that their behavior is indicative of the physics.

So there is one case I know where people sort of found chaotic renormalization group trajectories for some kind of a [INAUDIBLE] system. But always, again, this is a very general procedure. We have to limit mathematics, ultimately, by what the physical process is. So it's good that you know that these equations can do all kinds of strange things. But then we take a particular physical system, we have to beat on them until they behave properly.

So let's imagine that we have a situation, such as this, where we have three parameters. Two of them are irrelevant. One of them is relevant.

Then presumably, I take my physical system at some temperature and it would correspond to being on some point in this phase diagram. Some color that we don't have. Let's say over here. And I change the temperature. And I will take some trajectory-- in this case, three dimensional space. And this is a line in this three dimensional space.

And experimentally, I've been told that if I take, let's say, my piece of iron and I change temperature, at some point I go through a point that has infinite correlations. So I have to conclude that my trajectory for iron will intersect with surface at some point.

And I'll say, OK, I take nickel. Nickel would be something else. And I change temperature of nickel, and I will be doing something completely different. But that experimentalist also has a point where you have ferromagnetic transition, so it must hit this surface. Then you do cobalt, where some other trajectory comes and hits off the surface.

Now what we now know is that when we rescale the system sufficiently, all of them ultimately are described at the point where they have infinite correlation length by what is going on over here. So if I take iron, nickel, cobalt, clearly at the level of atoms and molecules, they are very different from each other. And the difference between ironness, nickelness, cobaltness is really in all of these irrelevant parameters.

And as I go and look at larger and larger scale, they all diminish and go away. And at large scale, I see the same thing, where all of the individual details has been washed out.

So this is able to capture the idea of universality. But there is a very important caveat to this, which is that the experimental system, whether you take iron or cobalt or some mixture of these different elements, you change one parameter

temperature, and you always see a transition from, let's say, paramagnetic to ferromagnetic behavior.

Now if I have, say, a line here in three dimensional space and I draw another line that corresponds to change in temperature, I will not intersect it. I have to do something very special to intersect that line. So in order that genetically I have a phase transition-- which is what my experimentalist friends tell me-- I know that I can only have one relevant direction, because the dimensionality of the basing of attraction is the dimensionality of the space minus however many relevant directions I have.

And I've been told by experimentalists that they exchange one parameters, and generically they hit the surface. So that's part of the story. I better find a theory that, at the end of the day, when I do all of this, I find a fixed point that not only is well-behaved and is not a limit cycle, but also a fixed point that has one and only one relevant direction, if that's the physical system that I'm describing.

Now of course, maybe that was for the superfluid, where they could only change temperature, and you have a situation where the magnet comes into play and they say, oh, actually we also have the magnetic field. And we really have to go to the space of zero field. And then if I expand my space of parameters here to include terms that break the symmetry, in that generalized space, I should only have two relevant directions.

So it is kind of strange story, that all we are doing here is mathematics. But at the end of the day, we have to get the mathematics to have very specific properties that are dictated by very rough things about experiments.

So this was kind of conceptually rich. So I'll let you digest that for a while. And next lecture, we will start actually doing this procedure and finding these kinds of [INAUDIBLE] relations.