Phase transitions and spontaneous symmetry breaking

We have already seen examples of phase transitions in this course. But so far we haven't really discussed them from a formal aspect. This is the goal of these notes.

Phase transitions represent abrupt changes that occur in the equilibrium state of a system at certain specific points. Thermal (or "classical") phase transitions occur at a specific temperature T_c, called the critical temperature. Conversely, quantum phase transitions occur at zero temperature and are driven by some parameter g which measures the relative weight of two competing terms in a Hamiltonian. At a certain critical value g_c, the ground-state of the Hamiltonian changes abruptly.

Phase transitions are all about interactions and the thermodynamic limit. They emerge due to the non-trivial interactions of an enormous number of particles.

As we will learn, phase transitions depend heavily on symmetries and dimensionality. The symmetries of a model greatly determine most of the critical behavior. Moreover, the dimension of the lattice also plays a major role. Most models have no transitions for low d, but do have a transition for higher d.

In statistical mechanics we are often interested in what are the simplest models that may exibit a phase transition. And the winner by far is the <u>Ising model</u>.

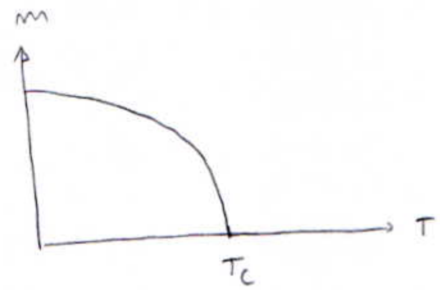
Consider a d-dimensional lattice, where each site is associated with spin 1/2 operators σ_x^i, σ_y^i, σ_z^i. the classical Ising model is then usually written as

$$\boxed{H = -J \sum_{\langle ij \rangle} \sigma_z^i \sigma_z^j} \tag{1}$$

where $\langle ij \rangle$ represents a sum over nearest neighbors. this model has a classical phase transition which can be monitored, for instance, by analyzing the magnetization

$$m = \frac{1}{N} \sum_i \langle \sigma_z^i \rangle = \frac{1}{N} \sum_i \mathrm{tr}\left\{ \sigma_z^i \frac{e^{-\beta H}}{Z} \right\} \tag{2}$$

If plotted as a function of T will usually behave as



The magnetization is called the <u>order parameter</u>; it is different from zero when $T < T_c$ (called the <u>ferromagnetic phase</u>) and identically zero for $T > T_c$ (called the <u>paramagnetic phase</u>)

For $T \lesssim T_c$ the magnetization usually goes to zero algebraically as

$$m \sim |T - T_c|^{\beta} \tag{3}$$

where β is called a <u>critical exponent</u> (This β is not $1/T$. Sorry! Everyone uses β, so I have to keep it here!)

What is remarkable is that β depends <u>only</u> on <u>symmetries</u> and the <u>dimensionality</u> of the model. The prefactors in (3), as well as the value of T_c, depend on a bunch of details. But β depends only on those two things. This is called the <u>universality</u> of phase transitions.

— // —

We can also modify the Ising model (1) so that, in addition to the classical phase transition, it also has a quantum phase transition. This can be accomplished by adding a <u>transverse field</u> g

$$\boxed{H = -g \sum_i \sigma_x^i - J \sum_{\langle ij \rangle} \sigma_z^i \sigma_z^j} \tag{4}$$

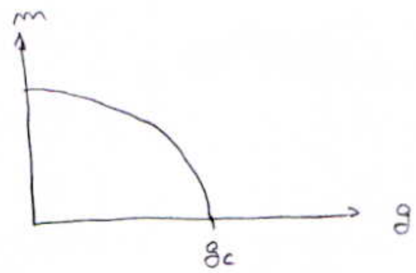
The important point here is that the σ_x guys don't commute with the other terms. Thus, they represent a competition of two effects. On the one hand $\sigma_z^i \sigma_z^j$ tends to align the spins in the z axis. On the other, σ_x^i tends to push them to the x direction.

At zero temperature, the ground-state of (4) will be a function of g : $|\psi(g)\rangle$ (of course, it also depends on J but we suppose J is fixed). There will then be a critical value g_c at which $|\psi(g)\rangle$ will suffer a macroscopic change.

This can, as before, also be captured by the magnetization along the z axis, which instead of (2), now reads

$$m = \frac{1}{N} \sum_i \langle \psi(g) | \sigma_z^i | \psi(g) \rangle \tag{5}$$

If $g = 0$ then m will be large because the spins will be aligned along the z axis. But if g is really large $m = 0$ because they will be aligned in the x-axis. Thus we will get something like



Again, close to g_c the magnetization will decay algebraically as

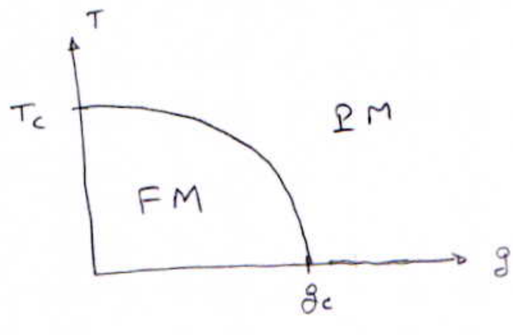
$$m \sim |g - g_c|^{\beta'} \tag{6}$$

with a different exponent β'. Quite interestingly, for the Ising model β' for d dimensions equals β for $d+1$

$$\beta'(d) = \beta(d+1) \tag{7}$$

There is a cool reason for this, as we will learn.

Lastly, we can consider the transverse field Ising model (TFIM) (4) at $T \neq 0$. Then we will have both a classical and a quantum phase transition. So we can plot a phase diagram which will look like this



Understanding this interplay between classical and quantum phase transitions is among the most widely studied topics in current research.

Mean-field approximation

The problem with phase transitions is that, precisely because they involve interacting systems, they are very difficult to study. A very common approach to be able to "do something" is to perform a mean-field approximation. Here is how it works.

Thinking about the Ising model, for concreteness, define the fluctuation operator

$$\delta\sigma_z^i = \sigma_z^i - \langle\sigma_z^i\rangle \tag{8}$$

By construction $\langle\delta\sigma_z^i\rangle = 0$. Moreover, we will assume that the system is translationally invariant, so that $\langle\sigma_z^i\rangle$ is independent of i. Then all will have the same value, which is precisely m in Eq (2):

$\langle\sigma_z^i\rangle = m$

We will now substitute in $\sigma_z^i\sigma_z^j$ the expression $\sigma_z^i = m + \delta\sigma_z^i$. This yields

$$\begin{aligned}\sigma_z^i\sigma_z^j &= (m + \delta\sigma_z^i)(m + \delta\sigma_z^j)\\ &= m^2 + m(\delta\sigma_z^i + \delta\sigma_z^j) + \delta\sigma_z^i\,\delta\sigma_z^j\end{aligned} \tag{9}$$

So far this is exact. Now comes the point about the mean-field approximation: it assumes that the fluctuations around the average are small compared to m, so that we may neglect the last term in (9)

Thus, mean-field is essentially the statement that $\delta\sigma_z^i \delta\sigma_z^j$ can be neglect. Whether this is a good approximation or not is something we don't have the tools yet to evaluate. So let's just keep going (rock-n-roll style) and see where this leads us.

The next step is to go back to σ_z^i

$$\sigma_z^i \sigma_z^j \simeq m^2 + m(\delta\sigma_z^i + \delta\sigma_z^j)$$

$$= m^2 + m(\sigma_z^i - m + \sigma_z^j - m)$$

$$= m(\sigma_z^i + \sigma_z^j) - m^2$$

Thus, we can summarize the mean-field approximation as

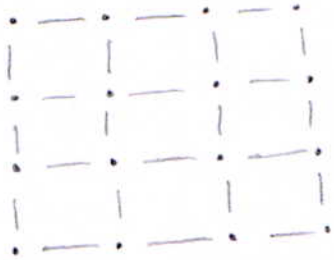
$$\boxed{\sigma_z^i \sigma_z^j \simeq m(\sigma_z^i + \sigma_z^i) - m^2} \qquad (10)$$

Next we plug this in Eq (5):

$$\sum_{\langle ij \rangle} \sigma_z^i \sigma_z^j = m \sum_{\langle ij \rangle} (\sigma_z^i + \sigma_z^j) - \sum_{\langle ij \rangle} m^2 \qquad (11)$$

Dealing with these sums can be a bit confusing, so let's practice with a specific example

Suppose we have a 2D square lattice



Now, $\langle ij \rangle$ means a sum over nearest neighbors. that is, it is a sum over all bonds. To carry out this sum, we can run over each of the N sites and then sum over left, right, up and down. But if we do this we will be counting each bond twice, so we have to divide by 2.

Based on this example, define

$$\nu = \text{number of nearest neighbors} \tag{12}$$

For the 2D square lattice $\nu = 4$. In fact, for d-dimensional cubic lattices, $\nu = 2d$. (For instance, $d = 3$ has 6 nearest neighbors). Then

$$\sum_{\langle ij \rangle} \text{ will have } \frac{N\nu}{2} \text{ terms} \tag{13}$$

Returning them to Eq (11), we get

$$\sum_{\langle ij \rangle} m^2 = \frac{N\nu}{2} m^2 \tag{14}$$

Similarly

$$\sum_{\langle ij\rangle} m\sigma_z^i = \frac{\nu}{2}\sum_i m\sigma_z^i$$

$$\sum_{\langle ij\rangle} m\sigma_z^j = \frac{\nu}{2}\sum_j m\sigma_z^j = \frac{\nu}{2}\sum_i m\,\sigma_z^i$$

Thus (11) becomes, in the mean-field approximation

$$\boxed{\sum_{\langle ij\rangle} \sigma_z^i\sigma_z^j \simeq -\frac{N\nu}{2}m^2 + \sum_i \nu m\sigma_z^i} \qquad (15)$$

What is important about this result is that the RHS is now a sum of independent terms. We have effectively decoupled the interacting system into N non-interacting systems. Of course, the price we pay for this is that now there is an extra parameter $m = \langle\sigma_z^i\rangle$ hanging around in the model.

This m is a bit weird because it is an average over a state $e^{-\beta H}$ which itself depends on m through H. Hence, m will have to be determined self-consistently. We will see how to do this next

Classical phase transition

To start, let us consider the simple Ising model (5), but let's add a magnetic field

$$H = -h\sum_i \sigma_z^i - J\sum_{\langle ij\rangle} \sigma_z^i \sigma_z^j \tag{16}$$

Using the MF approx. (15) this becomes

$$H \simeq \frac{N\nu}{2} J m^2 - \sum_i (h + J\nu m)\sigma_z^i \tag{17}$$

these are now independent spin 1/2 particles, so it doesn't matter if we work with N spins or with just one. In any case, recall that

$$\operatorname{tr}(e^{\beta h \sigma_z}) = 2\cosh(\beta h) \tag{18}$$

thus

$$Z = \operatorname{tr}(e^{-\beta H}) = e^{-\beta\frac{N\nu}{2}Jm^2}\left[2\cosh\left(\beta(h+J\nu m)\right)\right]^N \tag{19}$$

the free energy per particle will then be

$$\boxed{f = -\frac{T}{N}\ln Z = \frac{J\nu}{2}m^2 - T\ln\left\{2\cosh\left[\beta(h+J\nu m)\right]\right\}} \tag{20}$$

From this we easily compute the magnetization as

$$m = -\frac{\partial f}{\partial h} \tag{21}$$

which gives

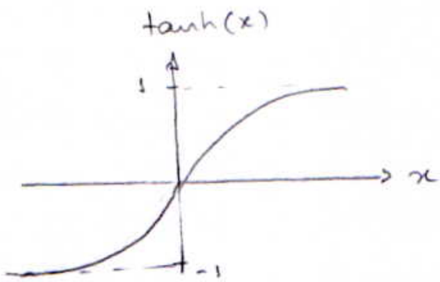
$$\boxed{m = \tanh\left(\frac{h + J\nu m}{T}\right)} \tag{22}$$

this is called the Curie - Weiss equation. It is a self-consistent equation for m.

We can understand it's solution as follows. First let's assume $h=0$. then we get

$$m = \tanh\left(\frac{J\nu m}{T}\right) \tag{23}$$

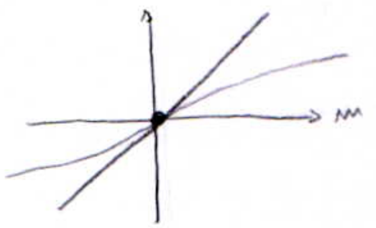
the function $\tanh(x)$ has the following form

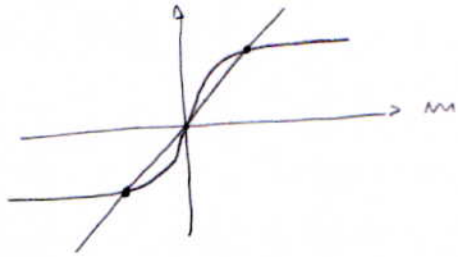


close to $x = 0$ it behaves linearly

$$\tanh(x) \simeq x$$

Thus, we see that there are two possibilities related to whether $\frac{Jvm}{T}$ initially has a slope which is larger or smaller than m.





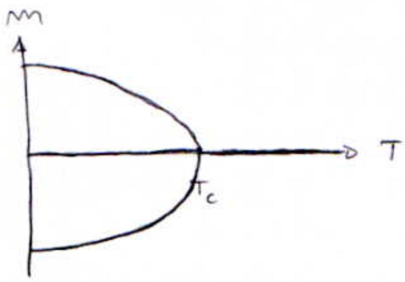
If $Jv/T > 1$ then $\tanh\left(\frac{Jvm}{T}\right)$ initially shoots up higher than m. But then it has to eventually curve back since $\tanh(x) \in [-1, 1]$. Thus, there will eventually be two points $m \neq 0$ in which the curve $\tanh(Jvm/T)$ will cross the curve m. Conversely, if $\frac{Jv}{T} < 1$ then the only solution is $m = 0$.

Thus, we see that this allows us to identify the critical temperature

$$\boxed{T_c = Jv} \tag{24}$$

For $T > T_c$ the only solution is $m = 0$ (PM phase). But if $T < T_c$ then there will be solutions with $m \neq 0$ (FM phase) ($m = 0$ continues to be a solution.

the magnetization will therefore look like this



the solutions with $m \neq 0$ are symmetric $m, -m$. This is a consequence of the symmetry in our model. When $h=0$, the Hamiltonian (16) is invariant under the Z_2 symmetry

$$\sigma_z^i \longrightarrow -\sigma_z^i \qquad (25)$$

the solutions reflect this.

For future reference, let us use (24) to get rid of $J\nu$ in our results. Thus (20) becomes

$$\boxed{f = \frac{T_c m^2}{2} - T \ln \left\{ 2\cosh\left(\frac{h + T_c m}{T}\right)\right\}} \qquad (26)$$

and (22) becomes

$$\boxed{m = \tanh\left(\frac{h + T_c m}{T}\right)} \qquad (27)$$

Thermodynamic properties

Now let's explore a bit more the thermodynamic properties of this system. It is convenient to write (27) as

$$\boxed{h = -T_c m + T \tanh^{-1}(m)} \tag{28}$$

Fun fact: this is the equation of state of the system, like $pV = N k_B T$. Here h plays the role of volume (an externally controlable parameter) and m plays the roll of pressure (the system's response). It is also useful to know that for small m, we can expand $\tanh^{-1}(m)$ as

$$\tanh^{-1}(m) = m + \frac{m^3}{3} \tag{29}$$

Thus, in the vicinity of the critical point (where m is small) Eq (28) becomes

$$h \simeq (T - T_c) m + \frac{T m^3}{3} \tag{30}$$

Now, if $h = 0$ we get either $m = 0$ or

$$m = \pm \sqrt{\frac{3}{T}(T_c - T)} \tag{31}$$

which, of course, is real only if $T < T_c$

We now see that, forgetting about the prefactors, m behaves as

$$m \sim |T - T_c|^{1/2} \tag{32}$$

so that the critical exponent in Eq (3) is

$$\beta = 1/2 \tag{33}$$

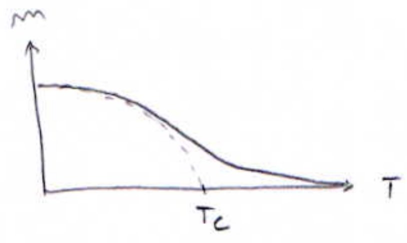
Going back to that idea of universality, this number 1/2 is a signature of a mean-field model. Literally all mean-field models have $\beta = 1/2$. If you are ever working on a model and find $\beta = 1/2$ for the order parameter, you may (without fear) be fancy and say "this model is therefore in the mean-field universality class"

This result is also an indication of how dramatic is the MF approx: note that β doesn't even depend on the number of nearest neighbors v, which is where the dimensionality comes in.

It is possible to show that MF works better for higher dimensions. In fact, every model has what is known as an upper critical dimension d_U at which the exact critical exponents become the same as the MF ones. For the Ising model this is $d_U = 4$. That is, if you could solve the Ising model in 4D exactly, it would give $\beta = 1/2$

(Just for curiosity, there is also a lower critical dimension d_L, below which no phase transition occurs. For the Ising this is $d_L = 2$. (there is no transition in 1D)).

Going back to (30), another common exponent to look at is how m scales with h exactly at T_c. In general, the curve $m(T)$ is modified by h in the following way



Thus, exactly at T_c, h will lift m from zero. From (30) we see that this will have the form

$$m = \left(\frac{3h}{T_c} \right)^{1/3} \tag{34}$$

This therefore defines another critical exponent

$$m \sim h^{1/3} \quad \text{at} \quad T = T_c \tag{35}$$

Usually this exponent is called $1/\delta$

$$m \sim h^{1/\delta} \tag{36}$$

so for MF $\delta = 3$.

Next we turn to the susceptibility:

$$\chi = \frac{\partial m}{\partial h} \tag{37}$$

Differentiating (30) with respect to h we get

$$1 = (T - T_c)\frac{\partial m}{\partial h} + T m^2 \frac{\partial m}{\partial h}$$

$$= \left[(T - T_c) + T m^2 \right] \chi$$

Thus

$$\chi = \frac{1}{T - T_c + T m^2} \tag{38}$$

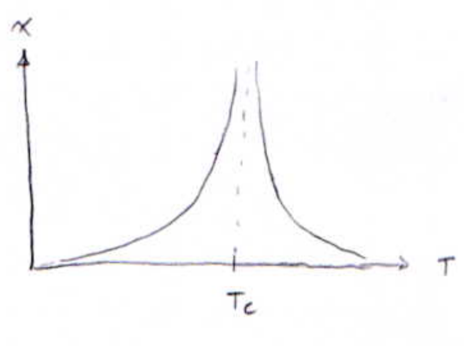
If $T > T_c$ then $m = 0$ so we get

$$\chi = \frac{1}{T - T_c} \tag{39}$$

which diverges at $T = T_c$. If $T < T_c$ then from (31) we get

$$\chi = \frac{1}{(T - T_c) + 3(T_c - T)} = \frac{1}{2(T_c - T)} \tag{40}$$

which also diverges. Thus, χ behaves as



Recall that χ measures the fluctuations in the magnetization. Thus, we see that at the critical point the fluctuations become extremely large

In either case, this defines a critical exponent γ as

$$\chi \sim \frac{1}{|T-T_c|^{\gamma}} \tag{41}$$

For MF we thus see that $\gamma = 1$.

We could go on and compute other things. Most notably is the specific heat, which usually also diverges as

$$C \sim \frac{1}{|T-T_c|^{\alpha}} \tag{42}$$

And the other really important one is the correlation length. It is defined from 2-point correlation functions

$$\langle \sigma_z^{i} \, \sigma_z^{i+r} \rangle \sim e^{-r/\xi} \tag{43}$$

thus, the correlation length ξ measures how quickly the correlations at 2 different parts of the system decay with their distance. Close to the critical point the correlation length diverges as

$$\xi = \frac{1}{|T-T_c|^{\nu}} \tag{44}$$

Thus, at $T = T_c$, even parts very far away from each other become strongly correlated.

Unfortunately for the MF Ising model it is not possible to capture this effect, since we loose all spatial information in the MF approx.

Here is a summary of the critical exponents for the Ising model in different dimensions

		$d = 2$	$d = 3$	$d=4$ (Mean-field)
$C \sim \dfrac{1}{\lvert T - T_c \rvert^{\alpha}}$	α	0	0.11008	0
$m \sim \lvert T - T_c \rvert^{\beta}$	β	1/8	0.3264	1/2
$\chi \sim \dfrac{1}{\lvert T - T_c \rvert^{\gamma}}$	γ	7/4	1.2370	1
$m \sim h^{1/\delta}$ @ $T = T_c$	δ	15	4.789	3
$\langle \sigma_z^i \sigma_z^{i+r} \rangle \sim \dfrac{1}{r^{d+\eta-2}}$ @ $T = T_c$	η	1/4	0.0362	0
$\langle \sigma_z^i \sigma_z^{i+r} \rangle \sim e^{-r/\xi}$, $\xi = \dfrac{1}{\lvert T - T_c \rvert^{\nu}}$	ν	1	0.62997	1/2

Quantum phase transition

Let's now use the MF approx (15) within the transverse field Ising model (4). As before, we also add a longitudinal field (in the z direction):

$$H = -h \sum_i \sigma_z^i - g \sum_i \sigma_x^i - J \sum_{\langle ij \rangle} \sigma_z^i \sigma_z^j \tag{45}$$

Applying (15) we get

$$H \approx \frac{N \nu J}{2} m^2 - \sum_i \left\{ (h + J \nu m) \sigma_z^i + g \sigma_x^i \right\}$$

As before, let us define $T_c = J\nu$. then we get

$$H = \frac{N T_c m^2}{2} - \sum_i \left\{ (h + T_c m) \sigma_z^i + g \sigma_x^i \right\} \tag{46}$$

This Hamiltonian is not yet diagonal. But it is easy to diagonalize it, as it is simply a sum of independent spin 1/2 particles.

Looking only on the reduced Hamiltonian of each spin,

$$H_1 = -h_e \sigma_z - g \sigma_x \quad , \quad h_e = h + T_c m \tag{47}$$

it's eigenvalues will be given by

$$\pm \sqrt{g^2 + h_e^2} \tag{48}$$

Thus, the partition function of each spin will be

$$Z_1 = \mathrm{tr}\, e^{-\beta H_1} = 2\cosh\left(\beta\sqrt{g^2 + h_e^2}\right) \tag{49}$$

and the full partition function will be $Z = (Z_1)^N$. The free energy per particle $f = -\frac{T}{N}\ln Z$ will then be

$$\boxed{f = \frac{T_c m^2}{2} - T\ln\left\{2\cosh\left[\frac{\sqrt{(h + T_c m)^2 + g^2}}{T}\right]\right\}} \tag{50}$$

which clearly reduces to the classical result (26) when $g = 0$.

For completeness, I will also write down the reduced density matrix of each spin:

$$\rho_1 = \frac{e^{-\beta H_1}}{Z_1} = \frac{1}{2} + \frac{1}{2}\frac{1}{\sqrt{h_e^2 + g^2}}\tanh\left(\frac{\sqrt{g^2 + h_e^2}}{T}\right)\left[h_e\sigma_z + g\sigma_x\right] \tag{51}$$

From (50) we get the (z-direction) magnetization as $m = -\partial f/\partial h$, or

$$\boxed{m = \frac{h + T_c m}{\sqrt{(h+T_c m)^2 + g^2}} \tanh\left(\frac{\sqrt{(h+T_c m)^2 + g^2}}{T}\right)} \tag{52}$$

Again, this reduces to the Curie-Weiss equation (27) when $g = 0$. We are particularly interested in $h = 0$, where this simplifies to

$$m = \frac{T_c m}{\sqrt{T_c^2 m^2 + g^2}} \tanh\left(\frac{\sqrt{T_c^2 m^2 + g^2}}{T}\right) \tag{53}$$

Again, there is always a solution with $m = 0$. Assuming $m \neq 0$ we get

$$\frac{\sqrt{T_c^2 m^2 + g^2}}{T_c} = \tanh\left(\frac{\sqrt{T_c^2 m^2 + g^2}}{T}\right) \tag{54}$$

If $g = 0$ we get back the classical solution (23). Instead, let us see what happens when $T = 0$. In this case $\tanh(\,) = 1$ and we get

$$T_c^2 m^2 + g^2 = T_c^2$$

or

$$m = \pm\sqrt{\frac{T_c^2 - g^2}{T_c^2}} \tag{55}$$

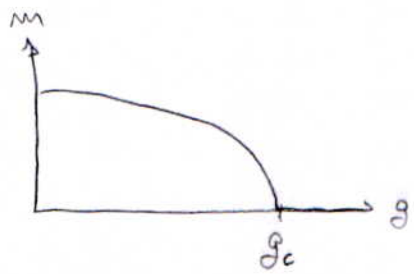
This motivates us to define the quantum critical parameter

$$\boxed{g_c = T_c = J\gamma} \tag{56}$$

Of course, the fact that $g_c = T_c$ is just a coincidence for this model. But in any case, we see that

$$m \sim (g_c^2 - g^2)^{1/2} \tag{57}$$

provided $g < g_c$. For $g > g_c$ the only solution is $m = 0$. Thus, we see that at $T = 0$ the system also undergoes a transition at a critical value of g



The fact that m depends on $g_c^2 - g^2$ instead of $g - g_c$ is also important and is related to the fact that $\langle \sigma_z^i \rangle$ doesn't care if the field is $g\sigma_x^i$ or $-g\sigma_x^i$. But we can also write (55) as

$$m = \pm \sqrt{\frac{(g_c - g)(g_c + g)}{g_c^2}}$$

If $g > 0$ then the only important part of the scaling behavior is

$$\boxed{m \sim |g_c - g|^{1/2}}$$

If $g < 0$ then it would be $g + g_c$

Going back now to (53) and (54), we can try to determine what is the relation between T and g such that there is a non-zero solution for m. This is how we determine the critical line in the T vs g diagram, which separates the two phases.

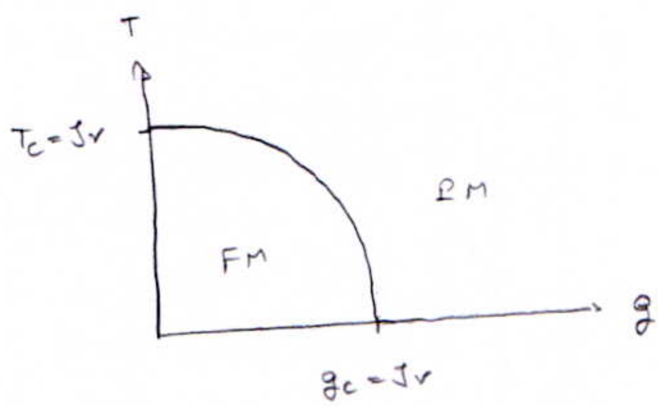
If m = 0 then we use (53). But in (54) we are already assuming $m \neq 0$. However, we can assume $m \neq 0$ but super tiny, so that we can set $m \simeq 0$ in (54), yielding

$$\frac{|g|}{T_c} = \tanh\left(\frac{|g|}{T}\right)$$

or

$$\boxed{T = \frac{|g|}{\tanh^{-1}(|g|/T_c)}} \tag{58}$$

A plot of this function looks like this



This is the typical phase diagram we discussed in page 5.

Quantum-classical mapping

I want to finish these notes with a calculation I really really like. I want to show that the quantum phase transition of the TFIM in d dimensions is equivalent to the classical phase transition of the classical Ising model in d+1 dimensions.

For concreteness I will work with $d=1$, the result is general. Thus, consider a 1D lattice with N sites

and described by the TFIM

$$H = -J \sum_{\langle ij \rangle} \sigma_z^i \sigma_z^j - g \sum_i \sigma_x^i \tag{59}$$

$$:= H_0 + V$$

We want to study the quantum phase transition (which we already solved for exactly a few lectures ago). One way to do this is to study the usual partition function

$$Z = \mathrm{tr}\left(e^{-\beta H}\right) \tag{60}$$

but then take the limit $T \to 0$, or $\beta \to \infty$.

To compute the trace in (60) we use the standard computational basis,

$$|\sigma\rangle = |\sigma_1 \ldots \sigma_N\rangle \tag{61}$$

which are the eigenstates of σ_z^i

$$\sigma_z^i |\sigma\rangle = \sigma_i |\sigma\rangle \tag{62}$$

Eq (60) then becomes a sum over 2^N states,

$$Z = \sum_\sigma \langle\sigma| e^{-\beta H} |\sigma\rangle \tag{63}$$

If $g = 0$ in (59) then H_0 would already be diagonal in the $|\sigma\rangle$ basis.

$$H_0 |\sigma\rangle = E_0(\sigma) |\sigma\rangle \tag{64}$$

where

$$E_0(\sigma) = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j \tag{65}$$

are the eigenenergies. But when $g \neq 0$ this is not the case, so computing the sandwich in (63) is very difficult

But now we will introduce the following trick. We are going to divide the number β into M parts

$$\beta = M\tau \qquad \tau = \beta/M \tag{66}$$

where M is supposed to be a very large integer. We then write

$$e^{-\beta H} = e^{-\tau H} e^{-\tau H} \dots e^{-\tau H} \qquad (M \text{ terms}) \tag{67}$$

Here we are allowed to split the exponentials because H commutes with itself. Eq (63) then becomes

$$Z = \sum_{\sigma} \langle \sigma | e^{-\tau H} \dots e^{-\tau H} | \sigma \rangle \tag{68}$$

We are now going to introduce a bunch of completeness relations between each exponential

$$1 = \sum_{\sigma} |\sigma\rangle\langle\sigma| \tag{69}$$

We then get

$$Z = \sum_{\sigma, \sigma_1 \dots \sigma_{M-1}} \langle \sigma | e^{-\tau H} |\sigma_1\rangle\langle\sigma_1| e^{-\tau H} |\sigma_2\rangle\langle\sigma_2| e^{-\tau H} |\sigma_3\rangle\langle\sigma_3| e^{-\tau H} \dots$$

$$\dots |\sigma_{M-1}\rangle\langle\sigma_{M-1}| e^{-\tau H} |\sigma\rangle$$

This is known as a Trotter decomposition

To make things more symmetrical, let us relabel $\sigma = \sigma_M$.

(its a dummy variable since we are summing over it). We then get

$$Z = \sum_{\sigma_1 \ldots \sigma_M} \langle \sigma_M | e^{-\delta H} | \sigma_1 \rangle \langle \sigma_1 | e^{-\delta H} | \sigma_2 \rangle \ldots \langle \sigma_{M-1} | e^{-\delta H} | \sigma_M \rangle \qquad (70)$$

Now there are 2^{N+M} terms in the sum. It's fun to see how the cyclic property of the trace introduces a kind of periodic boundary condition on σ_M.

Now comes the fun part. If M is large enough then δ will be very small. In this case we can approximate

$$e^{-\delta(H_0 + V)} \simeq e^{-\delta H_0} e^{-\delta V} \qquad (71)$$

since the corrections will be of the order δ^2. Thus, the sandwiches that appear in (70) will have the form

$$\begin{aligned} \langle \sigma_m | e^{-\delta H} | \sigma_{m+1} \rangle &\simeq \langle \sigma_m | e^{-\delta H_0} e^{-\delta V} | \sigma_{m+1} \rangle \\ &= e^{-\delta E_0(\sigma_m)} \langle \sigma_m | e^{-\delta V} | \sigma_{m+1} \rangle \end{aligned} \qquad (72)$$

where here

$$|\sigma_m\rangle = |\sigma_{m1}, \ldots, \sigma_{mN}\rangle \qquad (73)$$

(Sorry that the notation is a bit ambiguous)

The remaining sandwich in (30) can now be computed rather easily:

$$\langle \sigma_m | e^{-\tau V} | \sigma_{m+1} \rangle = \langle \sigma_{m,1} | e^{\tau g \sigma_x^1} | \sigma_{m+1,1} \rangle \langle \sigma_{m,2} | e^{\tau g \sigma_x^2} | \sigma_{m+1,2} \rangle \dots \langle \sigma_{m,N} | e^{\tau g \sigma_x^N} | \sigma_{m+1,N} \rangle \qquad (74)$$

These sandwiches refer now only to 2×2 matrices. They look something like $\langle \sigma | e^{\tau g \sigma_x} | \sigma' \rangle$ where $\sigma, \sigma' = \pm 1$ and

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad (75)$$

Since

$$e^{\tau g \sigma_x} = \begin{pmatrix} \cosh(\tau g) & \sinh(\tau g) \\ \sinh(\tau g) & \cosh(\tau g) \end{pmatrix} \qquad (76)$$

we may write

$$\langle \sigma | e^{\tau g \sigma_x} | \sigma' \rangle = \cosh(\tau g) \left(\frac{1 + \sigma\sigma'}{2} \right) + \sinh(\tau g) \left(\frac{1 - \sigma\sigma'}{2} \right) \qquad (77)$$

This is one of those naughty tricks with binary variables. If $\sigma = \sigma'$ (either $(1,1)$ or $(-1,-1)$) we get cosh and if $\sigma = -\sigma'$ (either $(1,-1)$ or $(-1,1)$) we get sinh.

Next let us try to write (77) as

$$\langle \sigma | e^{\tilde{\beta} g \sigma_x} | \sigma' \rangle = A\, e^{K \sigma \sigma'} \qquad (78)$$

for some constants A and K. Equating (77) and (78) we get that, if $\sigma = \sigma'$,

$$A\, e^{K} = \cosh(\tilde{\beta} g) \qquad (79a)$$

whereas if $\sigma = -\sigma'$

$$A\, e^{-K} = \sinh(\tilde{\beta} g) \qquad (79b)$$

Dividing (79a) by (79b) we get $e^{-2K} = \tanh(\tilde{\beta} g)$, or

$$K = -\frac{1}{2} \ln \tanh(\tilde{\beta} g) \qquad (80)$$

And multiplying (79a) and (79b) we get

$$A = \sqrt{\sinh(\tilde{\beta} g) \cosh(\tilde{\beta} g)} \qquad (81)$$

Now that we have the compact expression (78) for the 2×2 sandwich, we can write Eq (74) as

$$\langle \sigma_m | e^{-\tau V} | \sigma_{m+1} \rangle = A^N \exp\left\{ K \sum_{i=1}^{N} \sigma_{m,i}\, \sigma_{m+1,i} \right\} \qquad (82)$$

Thus, the sandwich (72) of the full $e^{-\tau H}$ becomes

$$\langle \sigma_m | e^{-\tau H} | \sigma_{m+1} \rangle = A^N \exp\left\{ \sum_{i=1}^{N} \left[\tau J\, \sigma_{m,i}\, \sigma_{m,i+1} + K\, \sigma_{m,i}\, \sigma_{m+1,i} \right] \right\} \qquad (83)$$

where I used Eq (65) for $E_0(\sigma_m)$.

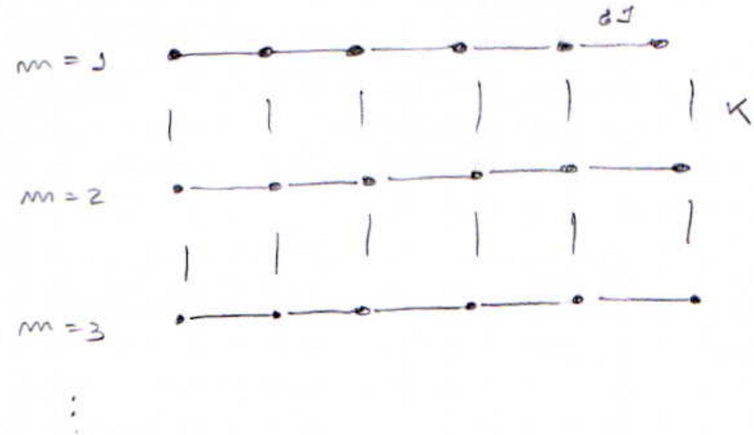
Finally, plugging this in the partition function (70) yields

$$\boxed{Z = A^{NM} \sum_{\{\sigma\}} \exp\left\{ \sum_{m=1}^{M} \sum_{i=1}^{N} \left[\tau J\, \sigma_{m,i}\, \sigma_{m,i+1} + K\, \sigma_{m,i}\, \sigma_{m+1,i} \right] \right\}} \qquad (83)$$

where $\{\sigma\}$ means a sum over all 2^{NM} spin configurations $\sigma_{m,i}$. Moreover, I left implicit a periodic boundary condition in m. That is, $\sigma_{M+1,i} = \sigma_{1,i}$.

What we see in (83) is the partition function of the classical Ising model, but in 2 dimensions! Thus, the physics of the quantum Ising chain in d dimensions is equivalent to the classical Ising model in $d+1$.

It's something like this



Note also that even though we had to do some approximations to get here, these approximations become exact as $M \to \infty$.

The type of calculation we just did is very powerful and the same idea is used in many situations. For instance, this is exactly the same procedure used to construct the Feynman path integral: a quantum model in d dimensions is mapped into a classical model in $d+1$. There this extra dimension is time. Here it is β. But the two are not so different: β is just an imaginary time because if we set $\beta = it$ then $e^{-\beta H}$ becomes the time evolution operator e^{-itH}.