Statistical Mechanics: lecture 7

Mean field models

"Mean-field" is the name we give to a series of models which display phase transitions and are also analytically solvable. They provide the starting point: we always start with mean field and then move on to consider more sophisticated models.

Here we shall consider mean field within the context of the Ising model. As the course moves on, we will also see it in other models and contexts.

We will also consider both the classical as well as the quantum Ising models. As we will see, this will serve to introduce us to the notion of quantum phase transitions.

Recommended reading: Salinas, chapter 12.2, 12.3 and 13.2, 13.3.

Currie - Weiss molecular field hypothesis

Consider a system of independent spin 1/2 particles with Hamiltonian

$$H = - h \sum_{i=1}^{N} \sigma_i^z \qquad (1)$$

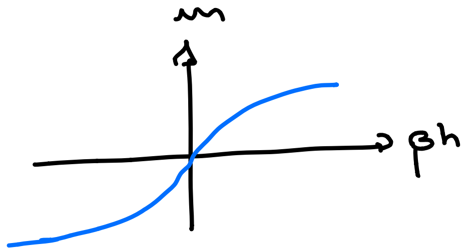
The magnetization per particle

$$m = \frac{1}{N} \sum_{i=1}^{N} \langle \sigma_i^z \rangle \qquad (2)$$

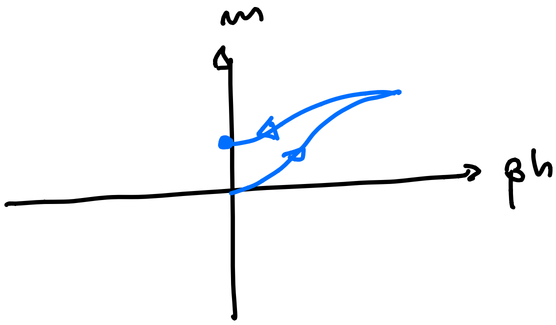
will be

$$m = \tanh(\beta h) \qquad (3)$$

this magnetization describes a paramagnetic material; this means a material where the magnetization responds in the same direction as the field



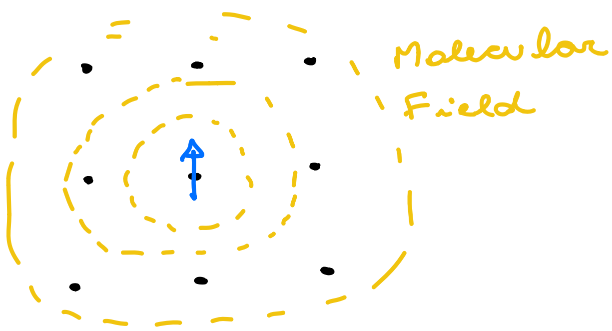
Some materials in nature, however, can have a magnetization even in the absence of a field:



they are called ferromagnetic.

This spontaneous magnetization is a consequence of interactions among the spins. The interaction induces them to align on a common direction.

Curie and Weiss proposed a phenomenological theory to describe this. They argued that each spin also felt the effects of all other spins around it as a kind of effective field. They called this a molecular field. The name "molecular" is now old fashioned. But its meant to be interpreted as "microscopic", because its generated by the surrounding spins.



Curie and Weiss argued that this could be modeled by simply modifying Eq (3) to

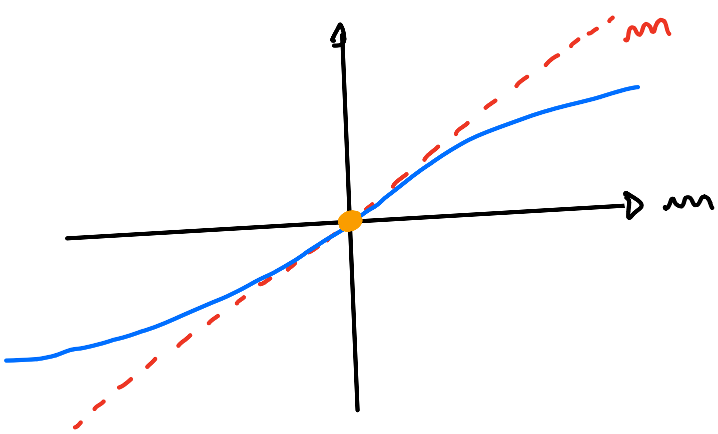
$$m = \tanh[\beta(h + \lambda m)] \tag{4}$$

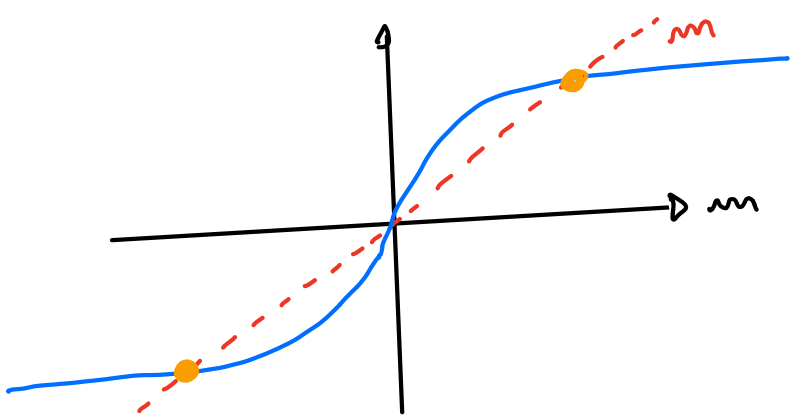
where λ is a constant. The logic is that this molecular field is generated by all spins around a spin. Thus, it should depend on the magnetization itself. The constant λ then measures the strenght of the molecular field.

Eq (4) is called the Curie-Weiss equation. It is a self consistent Eq. to be solved for m. Consider first the case $h=0$:

$$m = \tanh(\beta\lambda m) \tag{5}$$

We can get an idea of the solutions by comparing the left and right-hand sides graphically





close to $x = 0$,

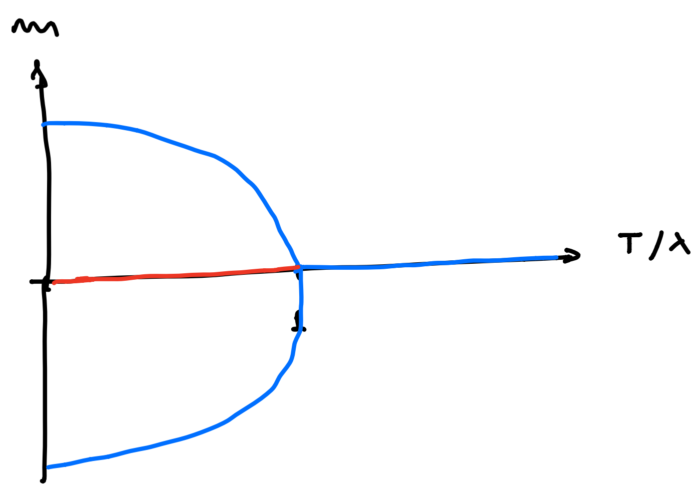
$$\tanh(x) \simeq x \tag{6}$$

Thus, near $m = 0$,

$$\tanh \beta\lambda m \simeq \beta\lambda m \tag{7}$$

If $\beta\lambda < 1$, then $\tanh(\beta\lambda m)$ starts below m, as in the left plot above. In this case the only solution is $m = 0$. But if $\beta J > 1$ then $\tanh \beta J m$ starts above m. Since it then has to eventually curve back, there will be two new solutions with $m \neq 0$.

The solutions of (5) thus look like this:

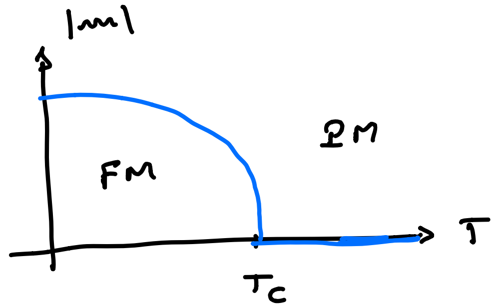


$m = 0$ is always a solution. We will learn later that when $T/\lambda < 1$ the solution with $m = 0$ is actually unstable.

This model therefore has a phase transition! The critical temperature of the model is when $\beta J = 1$, so

$$T_c = J \tag{8}$$

Above T_c the model is in a paramagnetic phase where $m \equiv 0$. But as we cool below T_c, the model enters a ferromagnetic phase, characterized by a spontaneous magnetization:



When $T \sim T_c$, m will be very small, so that we may expand

$$\tanh x \simeq x - \frac{x^3}{3} \tag{9}$$

Eq (5) then becomes

$$m \simeq \beta\lambda m - \frac{(\beta\lambda m)^3}{3} \tag{10}$$

This has a solution $m = 0$. Or, if $m \neq 0$, then

$$\frac{(\beta\lambda)^3}{3} m^2 = \beta\lambda - 1$$

Let us write this in terms of $T_c = \lambda$:

$$m^2 = \left(\frac{3T}{T_c}\right)^3 \left(\frac{T_c}{T} - 1\right)$$

or

$$m = \sqrt{\frac{3T^2}{T_c^3}}\,(T_c - T)^{1/2} \qquad (11)$$

this solution is real only when $T < T_c$. Otherwise, the only solution of (10) is $m = 0$. The important part of this result is the term in blue. It says that m tends to zero algebraically in the FM phase, with an exponent $1/2$. This is called a critical exponent. In fact, quite often we will write (11) only as

$$m \sim |T_c - T|^{1/2} \qquad (11')$$

where "$\sim$" means "the part that actually matters".

Equation of state: going back now to Eq (4), we can also write it as

$$h = -T_c m + T \tanh^{-1}(m) \tag{12}$$

this is the thermodynamic equation of state: h plays the role of pressure (a stimulus) and m plays the role of volume (the response).

Close to the critical point we may expand

$$\tanh^{-1}(m) \simeq m + \frac{m^3}{3} \tag{13}$$

to get

$$h \simeq (T - T_c) m + \frac{T m^3}{3} \tag{14}$$

Exactly at $T = T_c$ we then get

$$m = \left(\frac{3}{T_c}\right)^{1/3} h^{1/3} \tag{15}$$

this is another critical exponent: exactly at $T = T_c$, the magnetization grows algebraically with h with an exponent 1/3.

Susceptibility: the susceptibility at zero field is defined as

$$\chi = \left.\frac{\partial m}{\partial h}\right|_{h=0} \qquad (16)$$

Let us look close to T_c. Differentiating both sides of (14) with respect to h we get

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

$$= \left[(T-T_c) + T m^2\right]\frac{\partial m}{\partial h}$$

Thus (17)

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

If $T > T_c$ then $m=0$ and we get (18)

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

If $T < T_c$ then $m^2 = \frac{3T^2}{T_c^3}(T_c - T)$ [Eq (11)], so that

$$T - T_c + T m^2 = (T_c - T)\left\{-1 + \frac{3T^3}{T_c^3}\right\}$$

$$= (T_c - T)\left\{\frac{3T^3 - T_c^3}{T_c^3}\right\}$$

Since we are close to T_c,

$$\frac{3T^3 - T_c^3}{T_c^3} \simeq 2 + \frac{9(T-T_c)}{T_c}$$

Thus, to leading order in $T-T_c$, we can pick only the "2", leading to

$$\chi = \frac{1}{2(T_c - T)} \qquad T < T_c.$$

Thus, we conclude that both above as well as below T_c

$$\chi \sim \frac{1}{|T-T_c|} \tag{19}$$

The susceptibility diverges at the critical point with exponent -1.

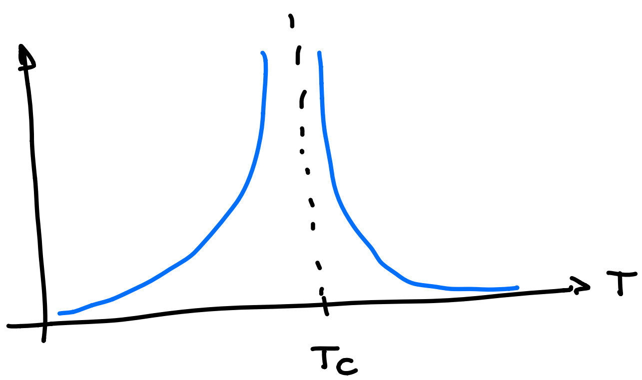


Table of critical exponents for the Ising model

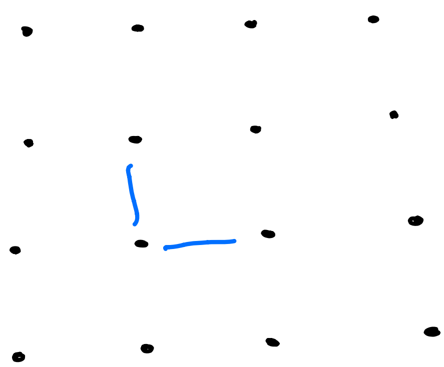
		2D Ising	3D Ising	Mean-field
$C \sim \frac{1}{\lvert T-T_c \rvert^{\alpha}}$	α	0	0.1108	0
$m \sim \lvert T-T_c \rvert^{\beta}$	β	1/8	0.3264	1/2
$\chi \sim \frac{1}{\lvert T-T_c \rvert^{\gamma}}$	γ	7/4	1.2370	1
$m \sim h^{1/\delta}$ @ $T = T_c$	δ	15	4.789	3
$\langle \sigma_m^z \sigma_{m+r}^z \rangle \sim \frac{1}{r^{d+\eta-2}}$ @ $T = T_c$	η	1/4	0.0362	0
$\langle \sigma_m^z \sigma_{m+r}^z \rangle \sim e^{-r/\xi}$, $\xi \sim \frac{1}{\lvert T-T_c \rvert^{\nu}}$	ν	1	0.62997	1/2

Counting bonds

This is a small parenthesis to facilitate what will come next. Consider a classical Ising model:

$$H = -J \sum_{\langle i,j \rangle} \sigma_i^z \sigma_j^z - h \sum_i \sigma_i^z \tag{20}$$

where $\langle i,j \rangle$ is a sum over nearest neighbors. For concreteness, let us assume we are on a d-dimensional cubic lattice with periodic BCs.



One way to run over all n.n. $\langle i,j \rangle$ is to run over all sites and, for each site, consider the interaction "above and to the right". This way we never count each bond twice. Alternatively, we can count all bonds for each site (4 in the 2D lattice). Then we will be counting each bond twice, so we have to divide by 2 in the end.

In any case, we conclude from this that in 2D there will be $2N$ bonds, where N is the total number of sites. In 3D there will be $3N$ and in d-dimensions, dN.

This serves to define the coordination number of a lattice:

$$\text{coordination number} := \gamma = \text{number of nearest neighbors.} \tag{21}$$

In a d-dimensional lattice,

$$\gamma = 2d \tag{22}$$

(think about the square lattice). The number of bonds will then be

$$n^\circ \text{ of bonds} = \frac{\gamma N}{2} \tag{23}$$

Mean-field approximation

Considering now the Ising model (20), let us introduce fluctuation operators

$$\delta\sigma_z^i = \sigma_z^i - \langle \sigma_z^i \rangle \tag{24}$$

they represent the fluctuations of σ_z^i around the average. If we assume that the system is translation invariant, then all $\langle \sigma_z^i \rangle$ will be equal

$$\langle \sigma_z^i \rangle = m \tag{25}$$

The interaction term in (20) then becomes

$$\begin{aligned} \sigma_z^i \sigma_z^j &= (m + \delta\sigma_z^i)(m + \delta\sigma_z^j) \\ &= m^2 + m(\delta\sigma_z^i + \delta\sigma_z^j) + \delta\sigma_z^i \delta\sigma_z^j \end{aligned} \tag{26}$$

The idea of the mean-field approximation is that the last term is squared in the fluctuations. If the fluctuations are small, then this term can be neglected, leading to

$$\sigma_z^i \sigma_z^j \simeq m^2 + m(\delta\sigma_z^i + \delta\sigma_z^j) + \underbrace{\delta\sigma_z^i \delta\sigma_z^j}_{\simeq 0 \text{ (MF approx)}} \tag{27}$$

We can now go back to working with σ_z^i:

$$\sigma_z^i \sigma_z^j \simeq m^2 + m\left[\sigma_z^i + \sigma_z^j - 2m\right]$$

$$= -m^2 + m\left(\sigma_z^i + \sigma_z^j\right) \qquad (28)$$

Eq (20) then becomes

$$H \simeq -J \sum_{\langle i,j \rangle} \left[-m^2 + m(\sigma_z^i + \sigma_z^j)\right] - h \sum_i \sigma_z^i \qquad (29)$$

In the first term we use (23) to obtain

$$\sum_{\langle i,j \rangle} m^2 = m^2 \frac{N\nu}{2} \qquad (30)$$

where ν is the coordination number. In the 2^{nd} term of (29), we have a sum over all bonds, but the summand depends on i. We can thus loop over each site i and, for each site, there will be $\nu/2$ values of j:

$$\sum_{\langle i,j \rangle} m\sigma_z^i = \frac{m\nu}{2} \sum_i \sigma_z^i \qquad (31)$$

The sum over j is the same

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

whence, Eq (29) becomes

$$H = \frac{NJrm^2}{2} - \sum_{i=1}^{N} (Jrm + h)\,\sigma_z^i \tag{32}$$

We therefore see that within the mean-field approximation, the Hamiltonian reduces to that of independent spins subject to an effective magnetic field

$$h_{eff} = h + Jrm \tag{33}$$

But there is a catch: the field is a function of $m = \langle \sigma_z^i \rangle$. The problem is thus self-consistent.

The partition function is

$$Z = e^{-\beta N J r m^2/2} \left[2\cosh(\beta h_{eff}) \right]^N \tag{34}$$

so

$$F = -T \ln Z = \frac{NJrm^2}{2} - T \ln\left[2\cosh(\beta h_{eff})\right] \tag{35}$$

or, considering the free energy per particle, $f = F/N$,

$$f = \frac{Jrm^2}{2} - T \ln\left[2\cosh(\beta h + \beta J r m)\right] \tag{36}$$

The self-consistency part comes now. From f we compute the magnetization as

$$m = -\frac{\partial f}{\partial h} \tag{37}$$

yielding

$$m = \tanh(\beta h + \beta J \nu m) \tag{38}$$

We now recognize this as being precisely the Curie-Weiss equation (4), with the molecular field being given by

$$\lambda = J\nu \tag{39}$$

This makes sense: the molecular field depends on the interaction J as well as the number of nearest neighbors. More neighbors means a stronger molecular field.

We also know what will be the critical temperature

$$T_c = \lambda = J\nu \tag{40}$$

For future reference, let us rewrite the free energy in terms of T_c:

$$\boxed{f = \frac{T_c m^2}{2} - T \ln\left[2\cosh\left(\frac{h + T_c m}{T}\right)\right]} \tag{41}$$

This is a very important formula. We will come back to it many times.

The Landau free energy

Let us analyze (41) in more detail when we are close to T_c. In this case we can expand (41) for small m. Assume first that $h=0$. Then we can use

$$\ln[\cosh(x)] \simeq \frac{x^2}{2} - \frac{x^4}{12} \qquad x \ll 1 \tag{42}$$

This yields

$$f \simeq \frac{T_c m^2}{2} - T\left[\frac{1}{2}\left(\frac{T_c m}{T}\right)^2 - \frac{1}{12}\left(\frac{T_c m}{T}\right)^4\right] - T\ln 2$$

$$= \frac{T_c}{2}\left[1 - \frac{T_c}{T}\right]m^2 + \frac{T_c^4}{12T^3}m^4 - T\ln 2$$

$$= \frac{T_c}{2T}(T-T_c)m^2 + \frac{T_c^4}{12T^3}m^4 - T\ln 2 \tag{43}$$

The last term is a constant, so we can forget about it. The 1st two terms we now write as

$$\boxed{f = \frac{a m^2}{2} + \frac{b m^4}{4}} \tag{44}$$

where

$$a = \frac{T_c}{T}(T-T_c) \tag{45}$$

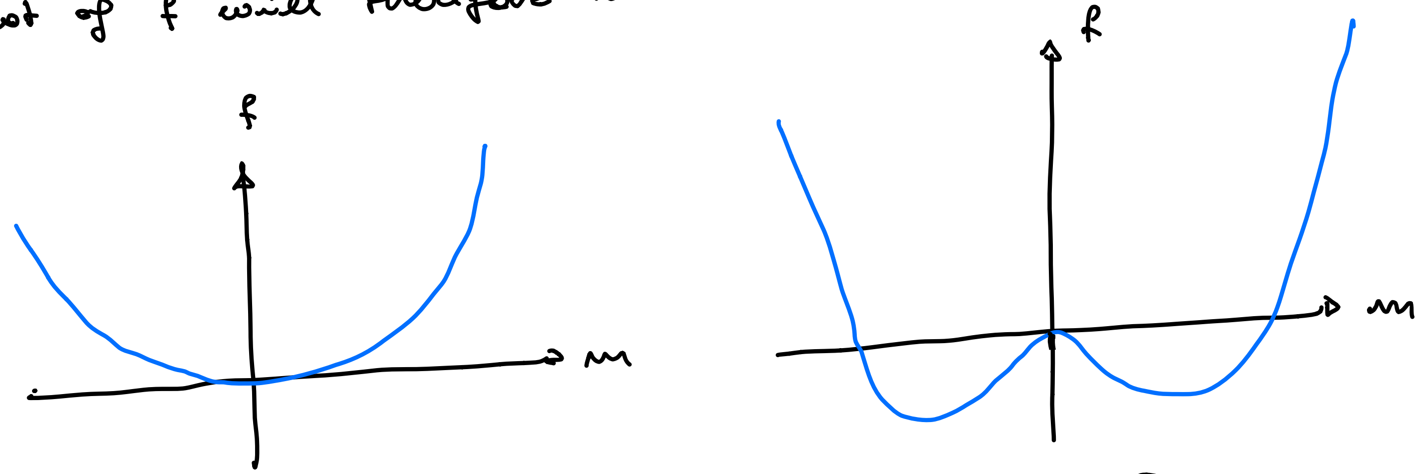
$$b = \frac{T_c^4}{3T^3}$$

This is called the Landau free energy.

Close to criticality, $f(m)$ is a simple polynomial in m. The coefficient b is always positive. But a, on the other hand, changes sign at the critical point:

$$\begin{aligned} a &< 0 \qquad T < T_c \\ a &> 0 \qquad T > T_c \end{aligned} \tag{46}$$

A plot of f will therefore look like this:



$T > T_c$ $\qquad$ $T < T_c$

For $T > T_c$ there is only one minimum at $m = 0$. But when $T < T_c$, two new minima appear at $m \neq 0$ and $m = 0$ becomes a maximum.

The importance of these minima is that we can view the solutions of the Curie-Weiss equation as those states which minimize the free energy: going back to (41), consider

$$\frac{\partial f}{\partial m} = T_c m - T_c \tanh(\beta h + \beta T_c m)$$

Thus we see that

$$\frac{\partial f}{\partial m} = 0 \iff m = \tanh(\beta h + \beta T_c m) \qquad (47)$$

The solutions are those which minimize the free energy. We already knew this of course: equilibrium is the state which minimizes the free energy.

This explains now why $m = 0$ becomes an unstable solution when $T < T_c$: it corresponds to a maximum, instead of minimum, of f.

Spontaneous symmetry breaking

The Landau free energy (44) preserves the symmetry of the original Ising Hamiltonian. Namely, it is invariant under Z_2 symmetry

$$m \to -m \tag{48}$$

(f only contains even powers of m).

Below T_c, however, the system will tend to either one of the minima of f:

$$m^* = \pm\sqrt{\frac{-a}{b}} \tag{49}$$

If the system is in such a state, then the symmetry (48) is broken. If the system is at $+\sqrt{-a/b}$ it cannot jump to $-\sqrt{-a/b}$ because there is a huge energy barrier between them. So if it is put in one minimum, it stays there.

The FM phase thus has a lower symmetry than the PM phase, because (48) is lost.

When the system undergoes a phase transition, a symmetry is spontaneously broken. This is what phase transitions are all about.

Quantum phase transitions

We can make the Ising model quantum by adding a transverse field

$$H = -J \sum_{\langle i,j \rangle} \sigma_z^i \sigma_z^j - h \sum_i \sigma_z^i - g \sum_i \sigma_x^i \tag{50}$$

The reason why this is quantum is because the different parts of the Hamiltonian no longer commute, so that it is not trivially diagonal. In fact, we will see how to diagonalize this model in 1D. In 2D or 3D we do not know how to do it.

We are going to treat (50) here at the mean-field level. The "black" part of Eq (50) is modified as before. Using Eq (32) we then get

$$H = \frac{N T_c m^2}{2} - \sum_{i=1}^{N} \left\{ (T_c m + h) \sigma_z^i + g \sigma_x^i \right\} \tag{51}$$

where $m = \langle \sigma_z^i \rangle$. This is again the Hamiltonian for N independent spins. But each Hamiltonian is not yet diagonal.

Let us focus on a single spin 1/2 with Hamiltonian

$$H = -h_e \sigma_z - g\sigma_x \qquad (52)$$

where, for us, $h_e = T_c m + h$. This Hamiltonian is diagonalized by a rotation

$$U = e^{-i\theta\sigma_y/2} = \begin{pmatrix} \cos\theta/2 & -\sin\theta/2 \\ \sin\theta/2 & \cos\theta/2 \end{pmatrix} \qquad (53)$$

with

$$\tan\theta = g/h_e. \qquad (54)$$

We then get

$$U^\dagger H U = \sqrt{h_e^2+g^2}\,\sigma_z = \begin{pmatrix} \sqrt{h_e^2+g^2} & 0 \\ 0 & -\sqrt{h_e^2+g^2} \end{pmatrix} \qquad (55)$$

Going back now to Eq (51), we introduce a global rotation

$$U = e^{-\frac{i\theta}{2}\sum_i \sigma_y^i} \qquad (56)$$

which will diagonalize all parts of the Hamiltonian

$$U^\dagger H U = \frac{N T_c m^2}{2} - \sum_i \sqrt{h_e^2+g^2}\,\sigma_z^i \qquad (57)$$

We are then back to Eq (32).

the partition function is

$$z = \mathrm{tr}\, e^{-\beta H}$$

Using the cyclic property of the trace and $UU^{\dagger} = 1$, we can also write this as

$$z = \mathrm{tr}\left\{ e^{-\beta H} U U^{\dagger} \right\} = \mathrm{tr}\left\{ U^{\dagger} e^{-\beta H} U \right\} = \mathrm{tr}\left\{ e^{-\beta U^{\dagger} H U} \right\} \tag{58}$$

Thus, we are back to (32) and (34):

$$z = e^{-\beta N T_c m^2/2} \left\{ 2 \cosh\left(\beta \sqrt{h_e^2 + g^2}\, \right) \right\}^N \tag{59}$$

The free energy per particle will thus be

$$\boxed{f = \frac{T_c m^2}{2} - T \ln \left\{ 2 \cosh\left(\beta \sqrt{(h + T_c m)^2 + g^2}\, \right) \right\}} \tag{60}$$

As a sanity check, this clearly reduces to the classical case (41) when $g = 0$.

The magnetization in the z-direction is

$$m = -\frac{\partial f}{\partial h} = \frac{h + T_c m}{\sqrt{(h+T_c m)^2 + g^2}} \tanh\left(\frac{\sqrt{(h+T_c m)^2 + g^2}}{T}\right) \tag{61}$$

which is a modified Curie-Weiss equation. Let us focus on $h=0$:

$$\boxed{m = \frac{T_c m}{\sqrt{T_c^2 m^2 + g^2}} \tanh\left(\frac{\sqrt{T_c^2 m^2 + g^2}}{T}\right)} \tag{62}$$

Clearly, $m=0$ is still a solution. If $m \neq 0$ then we get

$$\frac{\sqrt{T_c^2 m^2 + g^2}}{T_c} = \tanh\left(\frac{\sqrt{T_c^2 m^2 + g^2}}{T}\right) \tag{63}$$

If $g=0$ this reduces to Eq (5). Instead, let us consider what happens when $T \to 0$. In this case

$$\tanh(\cdots) \to 1$$

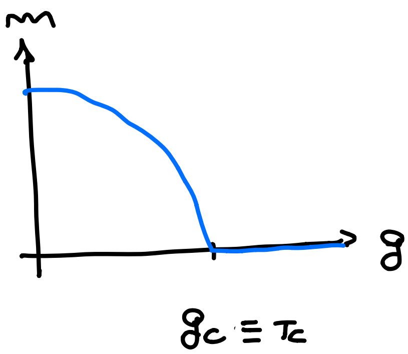
so we are left with

$$\frac{\sqrt{T_c^2 m^2 + g^2}}{T_c} = 1$$

or

$$m = \pm \frac{\sqrt{T_c^2 - g^2}}{T_c} \tag{64}$$

This behaves like



this is very similar to the classical transition, except that here it occurs as a function of g instead, and at $T=0$. this is the idea of a quantum phase transition. there is a critical parameter g_c (which coincidentally equals T_c for this model) where m changes from being non-zero to being identically zero.

the parameter g represents the transverse field (applied in the x direction). this field directly competes with the interaction J in (50), which wants the spins to be aligned in the z direction. For sufficiently strong g, J eventually capitulates.

To see this competition even more clearly, we can compute the magnetization in the x-direction

$$\nu := \langle \sigma_x^i \rangle \tag{65}$$

Since g is the transverse field

$$\nu = -\frac{\partial f}{\partial g} = \frac{g}{\sqrt{T_c^2 m^2 + g^2}} \tanh\left(\frac{\sqrt{T_c^2 m^2 + g^2}}{T}\right) \tag{66}$$

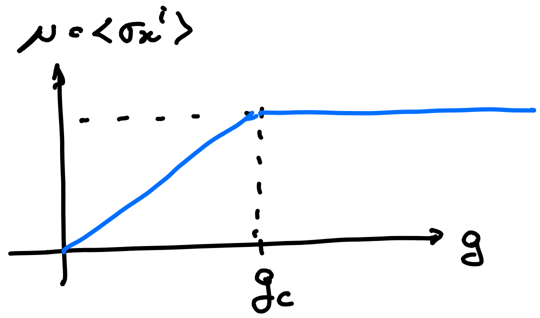
At $T=0$ this reduces to

$$\nu = \frac{g}{\sqrt{T_c^2 m^2 + g^2}} \tag{67}$$

If $g > g_c$ then $m=0$ and we get $\nu = 1$. If $g < g_c$ then m is given by (64) so

$$\nu = \frac{g}{T_c} = \frac{g}{g_c} \quad \text{since} \quad T_c = g_c. \tag{68}$$

Thus, the x magnetization behaves as



Clearly, as m goes down, ν goes up.

Going back to (64), it is also nice to note how it scales as

$$m \sim (g_c^2 - g^2)^{1/2} \tag{69}$$

It depends on g^2 and not g. This is a consequence of the fact that the transition is symmetric with respect to applying $+g\sigma_x^i$ or $-g\sigma_x^i$. But we can simply factor

$$m = \frac{\sqrt{(g_c - g)(g_c + g)}}{g} \tag{70}$$

then if $g > 0$ the relevant part of the scaling is

$$m \sim |g_c - g|^{1/2} \tag{71}$$

—//—

Let us now go back to $T \neq 0$ and the solutions of (62). We know that if $T = 0$ there is a solution at $g = g_c = T_c$. And we know that if $g = 0$ there is a transition at $T = T_c$. But what happens when both $T \neq 0$ and $g \neq 0$?

We know that the characteristic feature of the vicinities of the critical point is that $m \neq 0$ but very tiny. Consider then

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

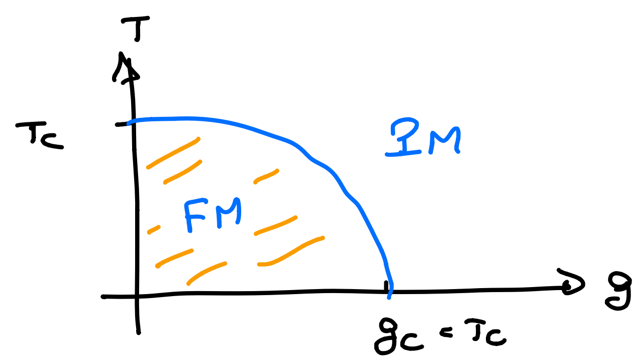
which already assumes $m \neq 0$. We expand both sides assuming $m \ll 1$. The expansion here is easy because we can take only the zeroth order term:

$$\frac{|g|}{T_c} = \tanh\left(\frac{|g|}{T}\right) \tag{72}$$

We can also write this as

$$\boxed{T(g) = \frac{|g|}{\tanh^{-1}(|g|/T_c)}} \tag{73}$$

This gives the phase diagram of the model. It describes the curve in the (T, g) plane where the phase transition occurs:



Finally, we can look at the Landau free energy close to criticality. Expanding (60) for $m \ll 1$ we get

$$f \simeq f_0 + \frac{a}{2} m^2 + \frac{b}{4} m^4 \tag{74}$$

where f_0 is a constant

$$a = \frac{T_c}{2|g|} \left(|g| - T_c \tanh(|g|/T) \right) \tag{75}$$

$$b = \frac{T_c^4}{16 |g|^3} \operatorname{sech}^2(|g|/T) \left[\sinh(2|g|/T) - 2|g|/T \right] \tag{76}$$

Since $\sinh(x) - x > 0$ for $x > 0$, we see that as before, $b > 0$. Also as before, a may change sign when

$$|g| - T_c \tanh(|g|/T) = 0 \tag{77}$$

which is nothing but Eq (72). This gives an alternative way of understanding the phase diagram in the (T, g) plane.