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 Ising model.

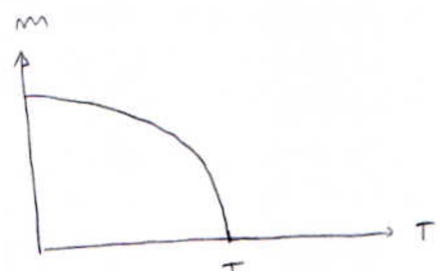
Consider a d-dimensional lattice, where each site is associated with spin 1/2 operators σ_x^i, σ_y^i, σ_z^i. the classical Ising model is then usually written as

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

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 order parameter: it is different from zero when $T < T_c$ (called the ferromagnetic phase) and identically zero for $T > T_c$ (called the paramagnetic phase)

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

$$m \sim |T - T_c|^{\beta} \qquad (3)$$

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

What is remarkable is that β depends only on symmetries and the dimensionality 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 universality 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 transverse field g

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

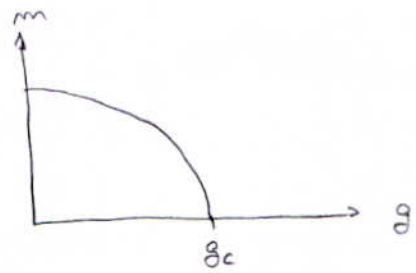
The important point here is that the σ_x guys don't commute with the other terms. Thus, they represent a competition of two effects. On the one hand $\sigma_z^i \sigma_z^j$ tends to align the spins in the z axis. On the other, σ_x^i tends to push them to the x direction.

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

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

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

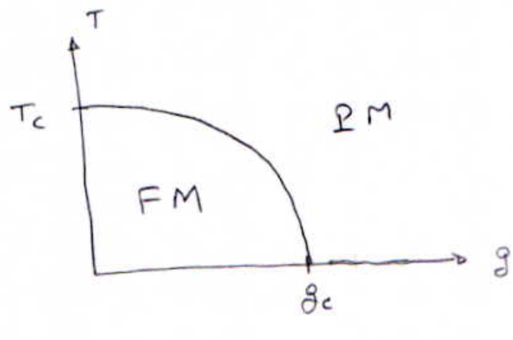
$$m \sim |g - g_c|^{\beta'} \qquad (6)$$

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

$$\beta'(d) = \beta(d+1) \qquad (7)$$

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 <u>phase diagram</u> 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

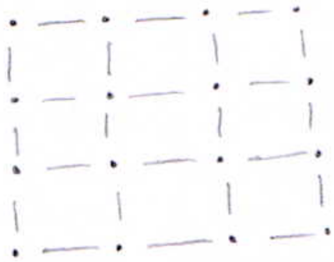
$$\boxed{\sigma_z^i\sigma_z^j \simeq m(\sigma_z^i + \sigma_z^i) - m^2} \tag{10}$$

Next we plug this in Eq (5):

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

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 <u>bonds</u>. 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 then 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