PHYS 516 (Statical Mechanics) Notes

Rio Weil

This document was typeset on February 2, 2023

Introduction:
This is a set of lecture notes taken from UBC's PHYS 516 (Graduate Statistical Mechanics) course, taught by Dr. Gordon Semenoff. The course covers fundamentals of statistical mechanics, phase transitions and critical exponents, $D = 1, 2, 3$ Ising models, mean field theory, quantum field theory, universality, renormalization, and elementary conformal field theory. If any errors are found in the notes, feel free to email me at ryoheiweil@phas.ubc.ca.

Contents

1 Introduction, Statistical Mechanics Review

1.1 Overview

This course centers around critical phenomena and phase transitions (primarily in magnetic systems/the Ising model) - PHYS 403 is a more comprehensive overview of the field, this is more specialized. We will discuss models that are analytically solvable (or almost), some renormalization group methods, some conformal field theory and the conformal bootstrap method.

We will begin with a review of some basic statistical mechanics. In a nutshell, statistical mechanics is the application of probability theory to a physical system - typically, with a large number of degrees of freedom as this is the limit where the application is useful. Perhaps saying probability theory is a bit reaching, though - the probability involved is pretty minimal (lots of counting, not a ton of measure theory). It is worth noting that (much like other fields of physics) there are very few systems that are analytically solvable; most systems require the application of approximate techniques.

1.2 Canonical Ensemble

There are various places to begin this discussion; let's start by discussing the canonical ensemble. Let us consider a physical system, which has an array of possible states. Let us assume that it is characterized by energies E_a, and the energy can take up one out of a list of possible values $E_1, E_2, E_3, \ldots$. Given the conservation of energy for a closed system, this is a reasonable way to characterize a state (and given one of our goals of doing thermodynamics with our system, this is a useful quantity). Let us not say too much more about the system - other than perhaps the fact that the energy has a lower bound (but not necessarily an upper bound), and that the energies are ordered. Further, for now let us assume that the energies are discrete - this is of course not true in general (there exist systems for which energy is a continuum, and there we will have to use some kind of binning procedure), but let us assume this simplification for now.

So, how do we make the canonical ensemble? We take $\mathcal{N}$ copies of the system, with various energies, so that $\mathcal{E} = \sum_{i=1}^{\mathcal{N}} E_i$ is the total energy. The $\mathcal{N}$ copies of the system are weakly coupled to each other. This means that energy can flow between the systems, but also that (since the coupling is weak) when we calculate the total energy we can neglect the interaction energies between the systems. In other ensembles, other things that are not the energy can be exchanged (e.g. particles in the grand canonical ensemble).

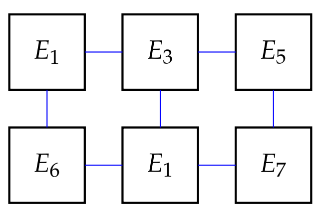

Figure 1.1: Cartoon of the Canonical Ensemble - we consider $\mathcal{N}$ (= 6 here) copies of a system, and weakly couple them such that they can exchange energy. Due to artistic constraints the couplings are only drawn for nearest neighbours, but in reality all the systems are coupled to each other.

In state of the ensemble is specified by the number of systems with a given energy, i.e. there are n_1 systems with energy E_1, n_2 systems with energy E_2, and so on. The total energy is then given by:

$$\mathcal{E} = \sum_a n_a E_a \tag{1.1}$$

and the number of systems in the ensemble is given by:

$$\mathcal{N} = \sum_a n_a \tag{1.2}$$

1.3 Fundamental Postulate - Equal a Priori Probability

To do statistical mechanics, we require a fundamental postulate - namely, an "equal a priori probability". This says that every distinct configuration of the ensemble is equally likely, subject to the total energy and number constraints. Physically, this means that the systems in the ensemble in time are flipping around the possible energy states (in a way that the total energy of the ensemble is conserved). The most probable configuration is the state in which the system spends the most time.

There are other versions of this; we can for example divide a system up in space, and then a spatial average will yield the most probable distribution.

What we look for (since the system visits every possible configuration equally) is the configuration which can be made in the most number of ways. And this is really the only probability theory we have to worry about here; counting up the number of ways to yield a given configuration of the ensemble. So, we ask how many ways are there to make the state $(n_1, n_2, \ldots)$? Let us derive this. Starting with the systems with energy E_1, we have:

$$\mathcal{N}(\mathcal{N}-1)\ldots(\mathcal{N}-n_1+1) \tag{1.3}$$

ways to have n_1 systems with energy E_1 (this is obtained by considering there are $\mathcal{N}$ systems to choose to have energy E_1, then $\mathcal{N}-1$ systems, and so on until all n_1 systems have been chosen). But this is overcounting because we don't care about the order, so really we require to divide this by $n_1!$:

$$\frac{\mathcal{N}(\mathcal{N}-1)\ldots(\mathcal{N}-n_1+1)}{n_1!} \tag{1.4}$$

and we continue with n_2, n_3 and so on until everything is full:

$$\frac{\mathcal{N}(\mathcal{N}-1)\ldots(\mathcal{N}-n_1+1)}{n_1!}\frac{(\mathcal{N}-n_1)\ldots(\mathcal{N}-n_1-n_2+1)}{n_2!}\ldots = \frac{\mathcal{N}!}{n_1!n_2!\ldots} \tag{1.5}$$

So, the most probable state of the system is that for which the above is maximized; in other words, we maximize it subject to $\sum_a n_a = \mathcal{N}$ and $\sum_a n_a E_a = \mathcal{E}$. This is a optimization problem with constraints - this may remind you of Lagrange multipliers which you have seen in classical mechanics. There is an apparent difficulty here in the fact that our numbers are discrete, but we'll get around it. To start, let us take the logarithm of the expression; we can maximize the logarithm of it instead of the original expression, and this is legal as the logarithm is monotonic (this is also a common trick done in machine learning and maximum likelihood estimation). The technique of Lagrange multipliers tells us that the expression of our interest is:

$$\ln\left(\frac{\mathcal{N}!}{n_1!n_2!\ldots}\right) + \beta\left(\sum_a n_a E_a - \mathcal{E}\right) + \gamma\left(\sum_a n_a - \mathcal{N}\right) \tag{1.6}$$

it would be nice to be able to use calculus techniques to solve this problem; to this end let us work in the regime of large $\mathcal{N}$ such that we can make the continuum approximation. At first this might seem like a poor assumption; after all after we saturate $\mathcal{E}$ all of the n_as past that point better not be large, but instead zero! To get around this we could assume some kind of cutoff to the energies. Of course there is still a decaying tail to the n_as, but these turn out to not be a problem.

In any case, let us suppose that we can approximate $\mathcal{N}$ large. Then, we can apply Stirling's formula:

$$\ln \mathcal{N}! \approx \mathcal{N}\ln\mathcal{N} - \mathcal{N}. \tag{1.7}$$

1.4 Interlude - Deriving Stirling's Formula

We start by writing down an integral expression for the factorial:

$$\mathcal{N}! = \int_0^\infty dx x^{\mathcal{N}} e^{-x} \tag{1.8}$$

we solve this via a saddle point technique of replacing the integrand with its maximum value; taking the derivative of the integrand and setting it to zero, we have:

$$\mathcal{N} x^{\mathcal{N}-1} e^{-x} - x^{\mathcal{N}} e^{-x} = 0 \tag{1.9}$$

which is maximized at $x = \mathcal{N}$. so, the approximate value of the factorial is:

$$\mathcal{N}! \approx \mathcal{N}^{\mathcal{N}} e^{-\mathcal{N}} \tag{1.10}$$

and taking logarithms we get Eq. (1.7). This technique also gives us a way of considering corrections to Stirling's formula by considering x near $\mathcal{N}$ (the next order corrections to $\ln \mathcal{N}!$ are $O(\log \mathcal{N})$, for example).

Another quick and dirty way to derive the formula (that doesn't give a nice way to study corrections, but gives us the leading terms that we want). Using the definition of the factorial and laws of logarithms, we have:

$$\ln \mathcal{N}! = \sum_{j=1}^{\mathcal{N}} \ln j \tag{1.11}$$

now approximating the sum as an integral:

$$\ln \mathcal{N}! \approx \int_1^{\mathcal{N}} dj \ln j = \mathcal{N} \ln \mathcal{N} - \mathcal{N} + 1 \tag{1.12}$$

in the large $\mathcal{N}$ limit we may neglect the $+1$, and we (again) obtain Stirling's formula.

1.5 Deriving the Boltzmann Distribution

Applying Stirling's Formula, Eq. (1.6) becomes:

$$\mathcal{N} \ln \mathcal{N} - \mathcal{N} - \sum_a (n_a \ln n_a - n_a) + \beta(\sum_a E_a n_a - \mathcal{E}) + \gamma(\sum_a n_a - \mathcal{N}) = \mathcal{N}\left(\sum_a (-\rho_a \ln \rho_a) + \beta(\sum_a \rho_a E_a - U) + \gamma(\sum_a \rho_a - 1)\right) \tag{1.13}$$

where we define $\rho_a = \frac{n_a}{N}$ and the second expression follows by algebra. Now, since ρ_a varies slowly, we may use techniques of calculus and take a derivative of the above expression and set it to zero. The ρ_a equation reads:

$$-\ln \rho_a - 1 + \beta E_a + \gamma = 0 \tag{1.14}$$

we also take derivatives by β, γ and set them to zero (as we do with the Lagrange multiplier technique):

$$\sum_a \rho_a E_a = U \tag{1.15}$$

$$\sum_a \rho_a = 1 \tag{1.16}$$

Let us rearrange the first equation, which has solution:

$$\rho_a = e^{\beta E_a + \gamma - 1} \tag{1.17}$$

we don't solve the second one, but the third one gives us:

$$\rho_a = \frac{e^{\beta E_a}}{\sum_a e^{\beta E_a}} \tag{1.18}$$

Note that the $e^{\gamma-1}$ goes away when we solve the third equation. It would have been wise to choose β with the other sign to start with. As we have derived things here, things only make sense if $\beta < 0$. We note that we have derived the ever-famous partition function:

$$Z = \sum_a e^{\beta E_a} \tag{1.19}$$

So, we have solved for the most likely distribution ρ_a; this is the known as the "Boltzmann distribution". We have not solved explicitly for β, but the second equation is formally unsolvable, and we will find a nice interpretation for β anyway (as the familiar $\beta = -\frac{1}{k_B T}$).

1.6 Example: Two-level system

We really have not done any physics at all here; but, we have completely generically found the most probable distribution. Let us try applying this to a two-level system and see how good our result is. Particles can have two (spin) states, $\uparrow$ and $\downarrow$. Let us assume we have two particles, and let us assume all orientations of the spins have the same energy. We can just enumerate all the states, and characterize the system by the net magnetization $m = \# \uparrow - \# \downarrow$. We have the four states $\uparrow\uparrow$ with $m = 2$, $\downarrow\downarrow$ with $m = -2$, $\uparrow\downarrow$ and $\downarrow\uparrow$ with $m = 0$. Here, the technique of most probable distribution is quite poor - there is a very good probability that the system is actually ferromagnetic (in fact half of the time) even though the most probable distribution is that the system is unmagnetized. However, we are able to see that unmagnetized is the most probable distribution, and in fact this is true for any number of particles.

So, let's consider generically N particles (note - let us assume that N is even so we can avoid frustration; if N is odd then there exists no configuration with zero magnetization). With N particles, the number of states with n spins up (from which we can obtain the magnetization as $m = n - (N - n) = 2n - N$) is $\frac{N!}{n!(N-n)!}$ of 2^n total possible states.

If we then plot $\ln \frac{N!}{n!(N-n)!}$ (making the continuum approximation), we find:

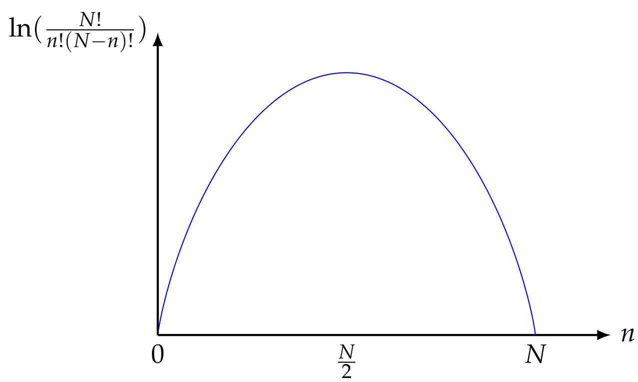

Figure 1.2: Plot of (continuum approximation) $\ln \frac{N!}{n!(N-n)!}$ as a function of number of spin-up spins n. We see the maximum at $n = N/2$ (zero magnetization).

so indeed the state with $n = N/2$ spins up (and $n = N/2$ spins down) - the state with total magnetization $m = 0$ is the most probable state. It is also valuable to ask what fraction of the total number of states is the most likely distribution. This is simply obtained by taking the number of states with $n = N/2$ and dividing by the total 2^n:

$$\frac{N!}{(N/2)!(N/2)!}\frac{1}{2^n} \tag{1.20}$$

We find that (using Stirling's formula that):

$$\ln\left(\frac{N!}{(N/2)!(N/2)!}\frac{1}{2^N}\right) = 0 + \frac{\ln N}{N} \tag{1.21}$$

so in the $N \to \infty$ limit, $\frac{N!}{(N/2)!(N/2)!}\frac{1}{2^N} \approx 1$ to leading order, so the proportion of the most likely distribution to total states of the system is one (with corrections given by the successive terms).

2 Free Energy, Ideal Gas, and the Grand Canonical Ensemble

2.1 Thermodynamic Interpretation, Energy, and Free Energy

Last time, we looked at the canonical ensemble. We derived the most probable distribution:

$$\rho_a = \frac{e^{-\beta E_a}}{\sum_a e^{-\beta E_a}} \tag{2.1}$$

and found the partition function:

$$Z = \sum_a e^{-\beta E_a}. \tag{2.2}$$

We argue that this already has a nice thermodynamic interpretation. This comes about if we look at the logarithm of the partition function:

$$F = -\frac{1}{\beta}\ln Z = -\frac{1}{\beta}\ln \sum_a e^{-\beta E_a} \tag{2.3}$$

Note that if there was only one energy level, then this would immediately just be the energy - in general the energy we calculate as the expectation value:

$$U = \frac{\sum_a E_a e^{-\beta E_a}}{\sum_a e^{-\beta E_a}} \tag{2.4}$$

How does F relate to U? Let us write:

$$F = U + \left(-\frac{1}{\beta}\ln \sum_a e^{-\beta E_a} - \frac{\sum_a E_a e^{-\beta E_a}}{\sum_a e^{-\beta E_a}}\right) \tag{2.5}$$

Let us call $e^{-\beta E_a} = Z\rho_a$ and write $E_a = -\frac{1}{\beta}\ln Z - \frac{1}{\beta}\ln \rho_a$. Then, rewriting the above expression, we find:

$$F = U - \frac{1}{\beta}\sum_a \rho_a \ln \rho_a. \tag{2.6}$$

The second term should be familiar to anyone with an information theory background - $S_{VN} = \sum_a \rho_a \ln \rho_a$ is known as the von Neumann entropy. It is the entropy of the distribution - a measure of how little we know about the system when we have the distribution ρ_a. It is minimized if one of the ρs is one and the others are zero, as the entropy is zero (then we know exactly what the system is). It is maximized if all of the ρs are constant (because then we know nothing about the system). If we are willing to accept that the von Neumann entropy is equal to the thermal entropy up to a constant:

$$S = k_B S_{VN} \tag{2.7}$$

where k_B is Boltzmann's constant. Then, we obtain:

$$F = U - \frac{1}{\beta}\frac{S}{k_B} \tag{2.8}$$

which closely resembles:

$$F = U - TS. \tag{2.9}$$

where T is the temperature (if we interpret $\beta = \frac{1}{k_B T}$). This is the familiar thermodynamic expression for the Helmholtz free energy.

This is not the historical order in which things are done - historically the microcanonical viewpoint (due to Boltzmann) came first, but this requires the system to be thermodynamic.

2.2 Example - System of weakly interacting non-relativistic particles

Let us assume we have a collection of N weakly interacting non-relativistic particles of mass m, which obey the laws of classical mechanics. A state of such a system will just be the specifications of the positions and velocity (or momenta) of all the particles (mathematically, this is because Newton's second law is a second-order ODE so we require two boundary conditions to specify the state). We can write the state as a collection of these values $\{\mathbf{q}_1, \mathbf{p}_1, \ldots, \mathbf{q}_n, \mathbf{p}_n\}$ The energy is then given by the Hamiltonian:

$$H = \sum_k \frac{\mathbf{p}_k^2}{2m} \tag{2.10}$$

Note we assume that the masses of the particles are the same and all attributes of the particles (other than position or momentum) are identical - note that in the context of classical mechanics this does not make the particles indistinguishable - we can keep track of them. This is in contrast to quantum statistical mechanics, where particles are truly indistinguishable and are either fermions or bosons.

We can construct the partition function for this system:

$$Z = \int d\mathbf{q}_1 d\mathbf{p}_1 \ldots d\mathbf{q}_n d\mathbf{p}_n e^{-\beta H} \tag{2.11}$$

this looks reasonable, but there are a couple things wrong with this. One problem - Z has dimensions; this is problematic if we want to take functions of it (e.g. logarithms to get the free energy). To deal with this problem, we just divide it by a number that gets rid of the dimensions:

$$Z = \frac{1}{(2\pi\hbar)^{3N}} \int d\mathbf{q}_1 d\mathbf{p}_1 \ldots d\mathbf{q}_n d\mathbf{p}_n e^{-\beta H} \tag{2.12}$$

$\hbar$ we pretty much pulled out of a hat here, but we require something with the dimensions of angular momentum to place there. Let's now do the integral. Let's assume that our particles move in infinite 3-D Euclidean space; we can then write $\int d\mathbf{q}_i = V$ (the volume) as H does not depend on the positions. Further, all momentum integrals are equivalent, so let us write it as the product of momentum integrals:

$$Z = \frac{V^N}{(2\pi\hbar)^{3N}} \left(\int dp e^{-\frac{\beta}{2m}\mathbf{p}^2} \right)^{3N} \tag{2.13}$$

We go into polar coordinates to solve this Gaussian integral:

$$\left(\int dp \right)^{3N} \to \left(\int d^2 p \right)^{\frac{3N}{2}} \to \left(\int \frac{d\varphi p dp}{2} \right)^{\frac{3N}{2}} \tag{2.14}$$

which yields:

$$Z = \frac{V^N}{(2\pi\hbar)^{3N}} \left(\frac{2\pi m}{\beta} \right)^{\frac{3N}{2}} = V^N \left(\frac{m k_B T}{2\pi\hbar^2} \right)^{\frac{3N}{2}} \tag{2.15}$$

The Helmholtz free energy is then:

$$F[T, V, N] = -k_B T N \ln \left[V \left(\frac{m k_B T}{2\pi\hbar^2} \right)^{3/2} \right] \tag{2.16}$$

This system should be truly thermodynamic (as we can take the system to be large), so this should work - we will see in a moment that unfortunately, it does not!

Recall thermodynamic differential relation:

$$dF = -SdT + \mu dN - PdV \tag{2.17}$$

So the entropy is:

$$S = -\frac{\partial F}{\partial T}\bigg|_{N,V} \tag{2.18}$$

the chemical potential is:

$$\mu = \frac{\partial F}{\partial N}\bigg|_{T,V} \tag{2.19}$$

and the pressure is:

$$P = -\frac{\partial F}{\partial V}\bigg|_{T,N} \tag{2.20}$$

so we can go to town and calculate some quantities. For example the pressure we can calculate to be:

$$P = \frac{Nk_B T}{V} \tag{2.21}$$

which is the ideal gas law! Big success (the other quantities will not be as successful...)! The chemical potential we can calculate to be:

$$\mu = -k_B T \ln\left(V\left(\frac{mk_B T}{2\pi\hbar^2}\right)^{3/2}\right) \tag{2.22}$$

The entropy we calculate to be:

$$S = \frac{3}{2}k_B N + k_B N \ln\left(V\left(\frac{mk_B T}{2\pi\hbar^2}\right)^{3/2}\right) \tag{2.23}$$

We can calculate the energy to be:

$$U = F + TS = \frac{3}{2}Nk_B T \tag{2.24}$$

which is again a beautiful formula (and the correct result). However, we should talk about why the formulas for μ, S are wrong. They do not have the correct extensivity properties. Concretely, if one considers two identical volumes of ideal gas separated by a partition, removing and re-inserting the partition should be reversible. However, a calculation of the entropy change shows that removing the partition leads to an increase in entropy of $2k_B N \ln 2$; contradiction. So we're a failure. But we're also clever, and can try to fix it. We introduce a factor of $\frac{1}{N!}$ into the partition function. This introduces a factor of $\frac{1}{N}$ into the logarithm in the free energy expression, which ends up correcting things. This was originally a fudge factor fix, but it turns out to be quite deep - namely, we have overcounted the states in the system somehow, and the $\frac{1}{N!}$ corrects for this. This hints to classical mechanics being problematic (quantum statistics fixes this with indistinguishability).

But let's try taking $Z \to \frac{1}{N!}Z$ (we are exploring what is known as Maxwell-Boltzmann statistics). Then with Stirling's formula:

$$\ln N! \approx N \ln N - N = \ln\left(\frac{N}{e}\right)^N \tag{2.25}$$

Then the free energy becomes:

$$F[T, N, V] = -k_B T \ln\left[\frac{eV}{N}\left(\frac{mk_B T}{2\pi\hbar^2}\right)^{3/2}\right] \tag{2.26}$$

and now we see that F has the correct scaling properties so things have been fixed. If we now recalculate quantities, P, U stay the same (as beautiful as they were):

$$PV = Nk_B T \tag{2.27}$$

$$\frac{U}{N} = \frac{3}{2} k_B T \tag{2.28}$$

And now the entropy is fixed up as well:

$$S = \frac{3}{2} k_B N + k_B N \ln \left(\frac{eV}{N} \left(\frac{m k_B T}{2\pi\hbar^2} \right)^{3/2} \right) \tag{2.29}$$

and this is known as the *Sackur-Tetrode equation*.

Before we go on - there are other ensembles we could have used, e.g. the grand canonical ensembles where the subsystems are allowed to exchange particles as well as energy. We need this as this one is the easier one to use when we consider quantum statistics. So in a way, Maxwell-Boltzmann statistics assumes the distribution is completely symmetric in the particles, but it is blind to how the wavefunction changes - if we take the distribution to be $\rho = \psi^\dagger \psi$, it assumes that each permutation comes out to be symmetric. But is not actually completely correct as in QM we impose the proper statistics on the wavefunctions, rather than the density.

2.3 Grand Canonical Ensemble

The logic follows exactly the same as the Canonical ensemble, with the only difference that we allow for the particle number to change. Suppose a system has N_a particles in state a, and energy E_a. We define n_a to be the number of systems in the ensemble in state a. We have some constraints:

$$\sum_a n_a = \mathcal{N} \tag{2.30}$$

$$\sum_a n_a E_a = \mathcal{E} = \mathcal{N} U \tag{2.31}$$

$$\sum_a n_a N_a = \mathcal{N} N \tag{2.32}$$

in the above, U is the average energy and N is the average number of particles. This differs from what we had before by one equation. We want to find the most probable distribution; for this the mathematics is exactly the same, just with one more Lagrange multiplier. We maximize:

$$\ln \frac{\mathcal{N}!}{\prod_a n_a!} + \beta(U\mathcal{N} - \sum_a n_a E_a) + \alpha \left(N\mathcal{N} - \sum_a n_a N_a \right) + \gamma \left(\mathcal{N} - \sum_a n_a \right) \tag{2.33}$$

The argument is basically identical to what we did to calculate ρ_a for the canonical ensemble. We solve the first and last equations for ρ_a and leave the other two unsolved (they will be thermodynamic quantities we can interpret). After the dust settles, we end up with:

$$\rho_a = \frac{e^{-\beta E_a - \alpha N_a}}{\sum_a e^{-\beta E_a - \alpha N_a}} \tag{2.34}$$

Our grand canonical partition function is:

$$\mathcal{Z} = \sum_a e^{-\beta E_a - \alpha N_a} \tag{2.35}$$

If we identify:

$$\Phi = -k_B T \ln \mathcal{Z} \tag{2.36}$$

with the grand canonical free energy, and go through a similar procedure of identifying the Von Neumann entropy with the thermodynamic entropy (and an identification to relate α with the chemical potential), we obtain:

$$\beta = -\frac{1}{k_B T}, \quad \alpha = -\frac{\mu}{k_B T}. \tag{2.37}$$

The grand canonical free energy now no longer depends on the number of particles, but on the chemical potential. It is however related to the Helmholtz free energy via a Legendre transform:

$$\Phi[T, \mu, V] = F - \mu \mathcal{N} \tag{2.38}$$

If we did things correctly, working with the grand canonical free energy vs. the Helmholtz free energy should yield the same answers. If we recall the canonical partition function for the weakly interacting gas, we had a dependence on the particle number N (note - NOT the average number of particles in the systems of the ensemble, but here really the number of particles in the ideal gas. Sorry for the overload of notation). We can sum over N to get the grand canonical partition function. We can then go through and see if we obtain the same results (and we will). We had:

$$Z[T, N, V] = \frac{1}{N!} V^N \left(\frac{m k_B T}{2\pi\hbar^2} \right)^{\frac{3N}{2}} \tag{2.39}$$

note the inclusion of the $\frac{1}{N!}$ factor so things end up correct. The grand canonical partition function is then:

$$\mathcal{Z}[T, \mu, V] = \sum_N e^{\frac{\mu}{k_B T} N} Z[T, N, V] \tag{2.40}$$

this is an easy sum because its just $\sum_N \frac{1}{N!} x^N$; hopefully this is familiar as just an exponential:

$$\mathcal{Z}[T, \mu, V] = e^{V \left(\frac{m k_B T}{2\pi\hbar^2} \right)^{3/2} e^{\frac{\mu}{k_B T}}} \tag{2.41}$$

so our prediction for the grand canonical free energy is:

$$\Phi = -k_B T \ln \mathcal{Z} = -k_B T V \left(\frac{m k_B T}{2\pi\hbar^2} \right)^{3/2} e^{\frac{\mu}{k_B T}} \tag{2.42}$$

and we can go through the song and dance to obtain the quantities that we solved for using the canonical ensemble.

3 Microcanonical Ensemble, Quantum Statistical Mechanics

We have spent two lectures discussing the canonical ensemble and grand canonical ensemble - in this setting we have discussed the classical perfect gas. We have one more to discuss; the microcanonical ensemble.

3.1 Review of the Canonical Ensemble

The canonical ensemble is a way to find the likelihood that a system is in a particular state; we found the most likely distribution:

$$\rho_a = \frac{e^{-\beta E_a}}{\sum_a e^{-\beta E_a}} \tag{3.1}$$

with $\beta = \frac{1}{k_B T}$. In deriving this we made some assumptions; for example the fact that the system was able to visit different energy states (this may not be true, e.g., if the system obeys some conservation law). In the setting of the perfect gas, we assumed that the particles were weakly interacting so they could exchange energy and visit different energy states (but weakly so we did not have to consider the interaction very carefully) - if the particles were completely free, they could not change their state (the momenta of all the particles would be fixed). We explored this distribution in the context of an ensemble of two-level systems; for two spins, we found that this was only a good description $\sim 50\%$ of the time, but as we increase N this estimate gets better and better.

3.2 Microcanonical Ensemble

The accuracy of the canonical ensemble only depended on the size of the ensemble - if something is not accurate enough, just look at a larger system. Essentially, a given subsystem sees the rest of the system as a heat bath/reservoir, and the estimate will get better as we increase the size of the heat bath. But we might ask - how well is the system described if we assume the system takes the most likely state? Roughly, the system takes the state such that $E^\nu e^{-\beta E}$ is maximized (those with a condensed matter background will be familiar with this sort of expression; the Boltzmann distribution times the density of states).

Concretely, we say:

$$\rho_a = \begin{cases} 1 & a : E_a = U \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

This looks extremely crude; it doesn't always work (e.g. the system needs to be highly degenerate) but for most normal things, it does. Let's reason out why. We consider the heat capacity:

$$C = \left.\frac{\partial U}{\partial T}\right|_{\ldots} \sim Vc \tag{3.3}$$

where C (the total heat capacity) scales with volume (times the per-volume heat capacity/specific heat capacity c).

In the canonical approach, U is the average of the energy:

$$U = \sum_a \frac{E_a e^{-\beta E_a}}{\sum_a e^{-\beta E_a}} = \langle E \rangle \tag{3.4}$$

We can write the heat capacity as:

$$C = \frac{\partial U}{\partial T} = \frac{1}{k_B T^2} \Delta U^2 \tag{3.5}$$

where:

$$\Delta U^2 = \sum_a \frac{E_a^2 e^{-\beta E_a}}{\sum_a e^{-\beta E_a}} - \left(\frac{\sum_a E_a e^{-\beta E_a}}{\sum_a e^{-\beta E_a}} \right)^2 = \left\langle E^2 \right\rangle - \langle E \rangle^2 \tag{3.6}$$

i.e. the variance in E. Note that both C (and so ΔU^2) and U scale with the volume of the system. Now, if we consider the variance of the energy:

$$\sqrt{\frac{\Delta U^2}{U^2}} \sim \frac{\sqrt{C}}{V} \sim \frac{1}{\sqrt{V}} \tag{3.7}$$

so as we take $V \to \infty$, the variance in the energy becomes tiny. So the conclusion is that:

$$\rho_a = \begin{cases} \frac{1}{D_a} & E_a = U \\ 0 & E_a \neq U \end{cases} \tag{3.8}$$

(where D_a is the degeneracy of E_a), is a reasonable distribution for (large) systems.

Note that this can be a useful approach for classical statistical mechanics (for quantum statistical mechanics, the grand canonical ensemble is the best approach). We now have Boltzmann's formula[1] for the entropy in terms of the degeneracy of states:

$$S = k_B \ln W \tag{3.9}$$

where $W = D_a$. In Boltzmann's work, W was known as the number of ways of making the system with energy U (QM, and hence degeneracy, was not conceptualized at the time). This pre-dates Gibbs, and ensembles - it was a fantastic guess by Boltzmann, though it was not believed by the time.

3.3 Ideal Gas - Microcanonical Ensemble Lens

What we need to find is W, which in a sense is the partition function in this context. It depends on U (and other quantities), and we obtain it by integrating $\delta(U - \sum_a \frac{\mathbf{p}_a^2}{2m})$ over all possible positions and momenta:

$$W[U] = \int d\mathbf{q}_1 d\mathbf{p}_1 \ldots d\mathbf{q}_N d\mathbf{p}_N \delta(U - \sum_a \frac{\mathbf{p}_a^2}{2m}). \tag{3.10}$$

There are a couple problems with this formula, one is namely the dimensionality (we want to plug this into a logarithm, so it must be dimensionless). Part of this is the same story as before; we add a phase space volume (fudge factor) of $\frac{1}{(2\pi\hbar)^{3N}}$ to cancel out the dimensions from the integral:

$$W[U] = \int \frac{d\mathbf{q}_1 d\mathbf{p}_1 \ldots d\mathbf{q}_N d\mathbf{p}_N}{(2\pi\hbar)^{3N}} \delta(U - \sum_a \frac{\mathbf{p}_a^2}{2m}) \tag{3.11}$$

Further, we need to cancel out the inverse energy that comes from the dirac delta function, so we add a δU factor:

$$W[U] = \int \frac{d\mathbf{q}_1 d\mathbf{p}_1 \ldots d\mathbf{q}_N d\mathbf{p}_N}{(2\pi\hbar)^{3N}} \delta U \delta(U - \sum_a \frac{\mathbf{p}_a^2}{2m}) \tag{3.12}$$

visually, one can imagine adding a sort of "thickness" of a shell of energy in phase space to the integral. This should go away whenever we start to do thermodynamics, though. Furthermore, we are overcounting states here again; so let us add a factor of $\frac{1}{N!}$ to compensate:

$$W[U] = \frac{1}{N!} \int \frac{d\mathbf{q}_1 d\mathbf{p}_1 \ldots d\mathbf{q}_N d\mathbf{p}_N}{(2\pi\hbar)^{3N}} \delta U \delta(U - \sum_a \frac{\mathbf{p}_a^2}{2m}) \tag{3.13}$$

The $\frac{1}{N!} \frac{1}{(2\pi\hbar)^{3N}}$ part of this is quantum, the rest is classical. Now, let's evaluate this. The position integrals are trivial; nothing in the integrand depends on volume, so we just get a factor of V^N. For the momentum

[1] Carved on his headstone!

integrals, we do a rescaling of the dirac delta function and then evaluate it, and also include the unit sphere factor in $3N$-dimensional Euclidean space:

$$W[U] = \frac{1}{N!} V^N \frac{1}{(2\pi\hbar)^{3N}} \frac{\delta U}{U} (2mU)^{\frac{3N}{2}} \frac{2\pi^{3N/2}}{\Gamma(\frac{3N}{2})} \tag{3.14}$$

Now, recall Stirling's formula $\ln N! = N \ln N - N$. The Gamma function is basically the factorial function, so $\ln \Gamma(x+1) = x \ln x - x + \dots$ (up to terms of order $O(\ln x)$). With this, we can clean up our expression for W:

$$W[U, V, N] = \left(\frac{eV}{N} \left(\frac{2e}{3} \frac{mU/N}{2\pi\hbar^2} \right)^{3/2} \right)^N \frac{\delta U}{U} \tag{3.15}$$

To get the entropy, a la Boltzmann we take the logarithm of this and then multiply by the Boltzmann constant:

$$S[U, V, N] = k_B N \ln \left(\frac{V}{N} \left(\frac{\frac{2}{3} mU/N}{2\pi\hbar^2} \right)^{3/2} \right) + \frac{5}{2} k_B N \tag{3.16}$$

for N large, $\ln \frac{\delta U}{U}$ is of order $\ln N$ (not order N) so we can neglect it. This is now our thermodynamic entropy, and from it we can derive the thermodynamic quantities that we are interested in. For example the (inverse) temperature is given by:

$$\frac{1}{T} = \left. \frac{\partial S}{\partial U} \right|_{V,N} \tag{3.17}$$

from which we can derive:

$$\frac{1}{T} = \frac{3}{2} N k_B \frac{1}{U} \tag{3.18}$$

and so:

$$U = \frac{3}{2} N k_B T \tag{3.19}$$

which is precisely the equipartition of energy formula for an ideal gas. We can proceed similarly with the pressure, and find the ideal gas law:

$$PV = N k_B T \tag{3.20}$$

and all the other nice formulas we derived before. So, this appears to be an equally as valid approach.

So, we've now covered three approaches; if we were to study ideal gases forever, we'd be pretty happy with this results. Particularly at higher temperatures/higher energy, this classical treatment works very well. Note that (after the last 30 minutes of lecture here) we will just use the canonical ensemble to look at Ising systems, and work in the classical regime where we set $\hbar = 0$ (though we will see that quantum field theories emerge in this setting... stay tuned). But for fun, let us quantize this theory, here.

3.4 Quantum Ideal Gas

Why is Maxwell-Boltzmann statistics not quite right? It is because in QM, states are described by the wavefunction $\psi(\mathbf{q}_1, \dots \mathbf{q}_n, t)$, but probabilities are described by $|\psi|^2$. A discussion of identical particles tells us that if we permute the position/particle labels, we get a state with the same energy. Nature tells us that there are only two[2] types of statistics possible; the wavefunction is totally symmetric in its labels for bosons and anti-symmetric for fermions. But Maxwell-Boltzmann statistics tells us that the probability is symmetric in the labels, which is not quite right (not really describing either bosons or fermions correctly).

Let us now count a system of (bosonic or fermionic) particles assuming we have a set of single-particle states that the particles can occupy. For bosons an arbitrary number of particles can fill a given state, for fermions this is forbidden (Pauli exclusion principle). Let's say a particle has energy ϵ_i. We now count

[2]well, there are exotic quasiparticles like anyons... but we leave that to another course

like physicists; count how many states we have for a few particles, and extrapolate. We take the grand canonical point of view where the system is open and can have any number of particles. We assign to each energy a Boltzmann weight. This sounds complicated, but it will turn out to be pretty easy.

In the case where we have no particles (the vaccum) we have only one possible state, so $\mathcal{Z}_B$ and $\mathcal{Z}_F$ both start with a 1:

$$\mathcal{Z}_B = 1 + \dots, \quad \mathcal{Z}_F = 1 + \dots \tag{3.21}$$

if we now have a one particle system, we obtain the single particle term $\sum_i e^{-\beta\mu}e^{-\beta\epsilon_i}$ (with the sum taking care of all possible energies that the particle could have). For a single particle, the bosons and fermion partition functions look the same:

$$\mathcal{Z}_B = 1 + \sum_i e^{\beta\mu}e^{-\beta\epsilon_i} + \dots, \quad \mathcal{Z}_F = 1 + \sum_i e^{\beta\mu}e^{-\beta\epsilon_i} + \dots \tag{3.22}$$

Now for two particles, the two partition functions start to look different. For fermions, we have $e^{2\beta\mu}\sum_{i<j} e^{-\beta(\epsilon_i+\epsilon_j)}$ (noting that the two particles cannot occupy the same energy):

$$\mathcal{Z}_F = 1 + \sum_i e^{\beta\mu}e^{-\beta\epsilon_i} + e^{\beta\mu 2}\sum_{i<j} e^{-\beta(\epsilon_i+\epsilon_j)} + \dots \tag{3.23}$$

for bosons, the two particles can be in the same state, so:

$$\mathcal{Z}_B = 1 + \sum_i e^{\beta\mu}e^{-\beta\epsilon_i} + e^{2\beta\mu}\left(\sum_{i<j} e^{-\beta(\epsilon_i+\epsilon_j)} + \sum_i e^{-2\beta\epsilon_i}\right) + \dots \tag{3.24}$$

If we continue in this fashion, we obtain:

$$\mathcal{Z}_F = \prod_i \left(1 + e^{\beta\mu}e^{-\beta\epsilon_i}\right) \tag{3.25}$$

$$\mathcal{Z}_B = \prod_i \frac{1}{1 - e^{\beta\mu}e^{-\beta\epsilon_i}} \tag{3.26}$$

Recalling the grand canonical free energy formula:

$$\Phi[T,\mu] = -k_B T \ln \mathcal{Z} \tag{3.27}$$

we find for bosons:

$$\Phi_B[T,\mu] = k_B T \sum_i \ln(1 - e^{\beta\mu}e^{\beta\epsilon_i}) \tag{3.28}$$

and for fermions:

$$\Phi_F[T,\mu] = k_B T \sum_i \ln(1 + e^{\beta\mu}e^{-\beta\epsilon_i}) \tag{3.29}$$

having obtained these, we can go ahead and do thermodynamics.

For free fermions/bosons, we have $\epsilon = \frac{\mathbf{p}^2}{2m} = \frac{\hbar^2\mathbf{k}^2}{2m}$. We can then replace the sums in the continuum approximation $\sum_i = V\int\frac{d^3k}{(2\pi)^3}$ and although these will not be solvable analytically we can obtain an approximate result by expanding. It is of interest to note that we can compare the results we get out of doing this with the previous derivation with Maxwell-Boltzmann statistics, and we find that for either 1 particle or in the high temperature limit things will agree. You will get to explore some of these limits in the first assignment.

4 Introduction to the Ising Model

We have now done a quick (and likely wholly inadequate) review of basic statistical mechanics. From here, we will begin to study phase transitions. We will study on a simple type of phase transition (with some side comments about others) - this is because there is a lot to be learned just by considering this simple example of magnetic/Ising systems. Most of the action here happened in the 1980s (so we aren't learning anything extremely new here) but this is just how physics classes go (but we will discuss the conformal bootstrap at the end of the course, which is something invented this century).

The Ising model was initially studied and solved in one dimension, where there are no phase transitions; it was then initially erroneously concluded that there are no phase transitions in any dimension. Later on it was solved exactly in two dimensions, where it was shown that there was a second order phase transition. This then became the standard model system to consider in the study of phase transitions. Through our study of this system, there is a huge amount of theoretical physics that are seeded by it (e.g. QFT). Even more than that, the language of these fields has largely inspired by the Ising model.

4.1 The Hamiltonian

We will use the canonical ensemble to study the Ising model. This is a physical model, so there is a Hamiltonian H which describes the energy of a configuration, and a density:

$$\rho = e^{-H/k_B T} Z \tag{4.1}$$

with Z the partition function:

$$Z = \sum_{\text{states}} e^{-H/k_B T}. \tag{4.2}$$

We have the thermodynamic quantity of the Helmholtz free energy:

$$F = -k_B T \ln Z. \tag{4.3}$$

These are the very general statistical mechanical ingredients. Let's talk about the model itself. It is a model of a magnetic system in one dimension, composed of a number of spins arranged on a lattice. If the spins like to align with each other, then we have a ferromagnetic material, if the spins like to anti-align per-site then we have a antiferromagnet.

How do we describe these spins? We can describe a spin at point x by:

$$\sigma_x = \begin{cases} +1 \\ -1 \end{cases} \tag{4.4}$$

i.e. a spin at position x which either points up or down. x lives on a lattice; a 2-D lattice is sketched as an example below:

Typically we work in the thermodynamic limit where the lattice is very large (in this limit the boundary conditions should not be important to understand the physics, but this is something we can check). Writing down the Hamiltonian, we have:

$$H_{interaction} = -J \sum_{x,\mu} \sigma_x \sigma_{x+\mu} \tag{4.5}$$

Where μ runs over the vectors that point to the nearest neighbours of x (so $x + \mu$ runs over the nearest neighbours of x). We will not concern ourselves too much with the lattice constant (we can just set it to one). Note that $\sigma_x \sigma_{x+\mu}$ is minimized if the two spins align, (as then $\sigma_x \sigma_{x+\mu}$ is positive and hence the negative sign out front gives the term a negative contribution). We can change the nearest neighbours assumption to include next nearest neighbours, infinite range etc. (and we can also change the behaviour

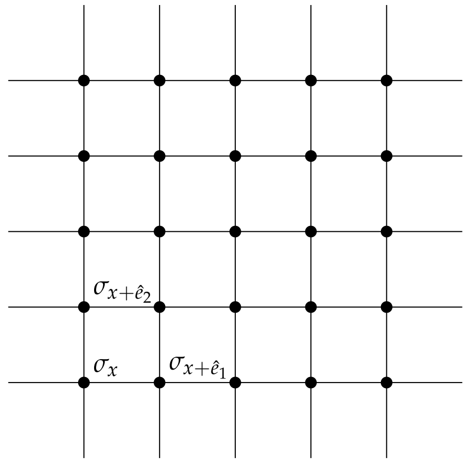

Figure 4.1: 2D Ising model on a square lattice. The spins sit on the nodes/crossings of the lattice, with interaction between neighbouring spins ($\mu = \pm\hat{e}_1, \pm\hat{e}_2$).

of the interactions) but realistic interactions are short range so for now nearest neighbours is good to consider.

This is almost the Ising model; we can also add a coupling term to an external magnetic field:

$$H_B = -B \sum_x \sigma_x \tag{4.6}$$

where the spins will want to align with B so as to lower the energy. So the Ising Hamiltonian is:

$$H = H_{interaction} + H_B = -J \sum_{x,\mu} \sigma_x \sigma_{x+\mu} - B \sum_x \sigma_x. \tag{4.7}$$

This simple two-parameter (J, B) model is already quite interesting. From it we will learn about universality - phase transitions tend to be the same for a large number of microsystems, e.g. the critical behavior for the Ising model and for liquid gas phase transition is the same. So this is why we can get away with studying something so simple.

4.2 Analyzing the Ising Model - Phase Diagram Boundaries

What we would like to compute is the partition function:

$$Z = \sum_{\text{spins}} e^{-H/k_B T} \tag{4.8}$$

and we would then like to calculate the free energy F, which is a function of the temperature T, the field B, and the number N:

$$F[T, B, N] = -k_B T \ln Z \tag{4.9}$$

if we can find F, we can get all sorts of physical quantities, e.g. magnetization, susceptibility etc. For example we have the magnetization (average sum over all spins):

$$M = -\frac{\partial F}{\partial B}\bigg|_{N,T} = \frac{\sum_{\text{spins}} (\sum_x \sigma_x) e^{-H/k_B T}}{Z} \tag{4.10}$$

As mentioned previously, we will assume that $N \to \infty$ (thermodynamic limit). What is left in our free energy is T and B. So, we can draw a phase diagram:

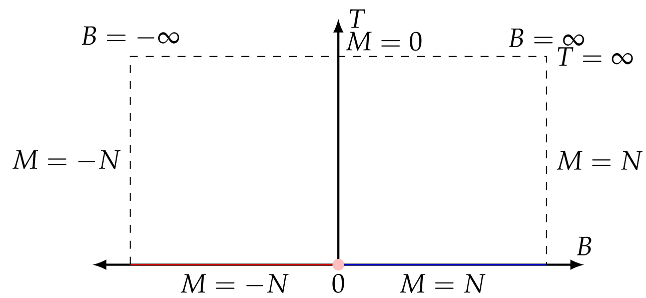

Figure 4.2: Phase Diagram Boundaries for the Ising Model, plotting temperature T and magnetization M. As $B = \pm\infty$, the spins want to align all upwards ($M = N$)/downwards ($M = -N$) respectively. As $T \to \infty$, all configurations are equally likely and so $M = 0$. At $T = 0$, all spins aligning minimizes the energy and so for $B < 0$ we have $M = -N$ and for $B > 0$ we have $M = N$. At $T = 0, B = 0$ there is a first order phase transition.

Note that as $T \to \infty$, we would first notice that our lab had burned down, but before that we would also notice that our material is no longer a ferromagnetic; it is a paramagnet with $M = 0$. We can see this from our expression for M in Eq. (4.10), where if $T \to \infty$ then $e^{-H/k_B T} = 1$ and so the up and down spins have equal weight, so the net summation comes out to zero.

At $B = \pm\infty$, we have $M \neq 0$. Specifically, if $M = \infty$ then all the spins want to align upwards so $M = N$ and if $M = -\infty$ then the spins want to align downwards so $M = -N$. As the temperature goes to zero, we are interested in the lowest energy state of the Hamiltonian. As all the spins are aligned, this minimizes the energy so actually we find $M = -N$ for all $B < 0, T = 0$ and $M = N$ for all $B > 0, T = 0$. What happens then at $B = 0$? We have a phase transitions between the two domains. We have a first order phase transition there where the magnetization completely flips.

What does the order of the phase transition mean? It refers to the number of the derivatives we can take of the free energy before the derivative does not exist. Here (at $T = 0$), $M = \text{sign}(B)N$, and the derivative of this does not exist at $B = 0$; so the first derivative of F exists but the second one does not, so it is a first order phase transition. This comes from the Landau classification of phase transitions. This is not the full story; we have discovered many topological phases in recent times, and it could be argued that the classification of phases is not yet complete. If we integrate, at $T = 0$ we find $F = J|B|$; this is continuous but not differentiable so the phase transition is first order.

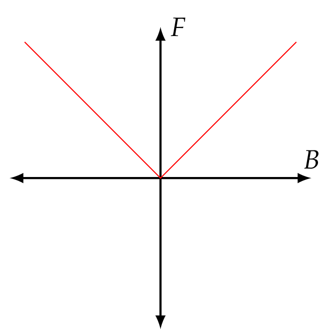

Figure 4.3: Plot of F at $T = 0$ as a function of B. The free energy is continuous but not differentiable, and hence there is a first-order phase transition at $B = 0$. The slope of the free energy is $-J/J$.

4.3 Analyzing the Ising Model - $B = 0$ and Spontaneous Symmetry Breaking

The line of $B = 0$ is interest. This is because the Hamiltonian has more symmetry in this case; namely the symmetry of flipping every $\sigma_x \leftrightarrow -\sigma_x$, or a $\mathbb{Z}_2$ symmetry (which can be represented, e.g. as $\{1, -1\}$ with multiplication). The magnetic field breaks this symmetry.

Note that if we are at some point in the right half of the phase diagram and we do a $\mathbb{Z}_2$ transformation, we map to the point on the other (left) side of the phase diagram (so on the $B = 0$ axis, we do not move at all!) Note however we cannot use symmetry to conclude that $M = 0$ at $B, T = 0$; as the magnetic moment here is dependent on the prior history of the system (how was this critical point reached). So we have some strange behaviour here; at this point, the Hamiltonian has $\mathbb{Z}_2$ symmetry but the actual state of the system does not inherit this symmetry; the system chooses one of the totally magnetized states depending on the prior history. This is known as *spontaneous symmetry breaking*.

Now we can ask; if we go up the $B = 0$ axis (by turning on the temperature) from the $B, T = 0$ point, what happens? Does the symmetry remain broken? In the 1-D Ising model, there is a phase transition where the symmetry is unbroken; the magnetization goes back to $M = 0$. We will show that this comes about through solving the model. In higher dimensions, we cannot exclude the possibility that $M \neq 0$ persists for a while even as we turn up T; for example it might be possible that M decays up to some point after hitting zero. We then have a discontinuity in the phase diagram partially along the $T = 0$ axis. We have a line of first order phase transitions, and it ends at some point in a second order phase transition. Unlike the boundaries of the diagrams, we have not really justified this statement; we will have to return to its proof later.

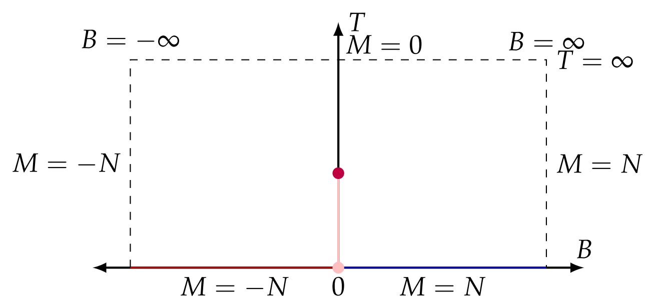

Figure 4.4: For dimensions larger than 1, $M \neq 0$ persists at $B = 0$ even as the temperature is tuned up. This results in a "first order phase transition line" partway up the $B = 0$ axis, culminating in a second order phase transition point.

It is difficult to study the phase transitions in generality, so we will (later) focus our attention to the point around the second order phase transition. Specifically, things are complicated by the magnetic field, so we will study a neighbourhood of the second order phase transition that lies on the $B = 0$ axis.

We will be able to study the 1-D model in full generality, but this is not very interesting as the "phase transition line" we see above will get shrunk to a point. The 2-D model has been solved but only at $B = 0$.

4.4 Solving the 1-D model

One of the slightly unsatisfying things about the study of this system is that the models in different dimensions have different ad-hoc ways of solving them analytically. When we start to look at approximations, we will see a unifying formalism (e.g. renormalization group).

We consider the 1-D Ising model, which is just a chain of classical spins.

The Hamiltonian takes the form:

$$H = -H \sum_{n=1}^{N-1} \sigma_n \sigma_{n+1} - B \sum_n \sigma_n \tag{4.11}$$

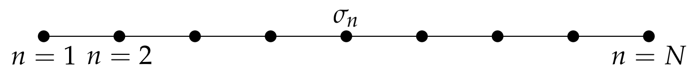

Figure 4.5: Cartoon of 1D Ising model (pictured with $N=9$ spins). The spin at site n is denoted σ_n.

The route to the solution here is just a brute-force calculation of the partition function:

$$Z = \sum_{\sigma_1=\pm1,\ldots,\sigma_n=\pm1} e^{\frac{J}{k_BT}\sum_n \sigma_n\sigma_{n+1} - \frac{B}{k_BT}\sum_n \sigma_n} \tag{4.12}$$

we rewrite this as an extended multiplication:

$$Z = \sum_{\sigma_1=\pm1}\sum_{\sigma_2=\pm1}\cdots\sum_{\sigma_n=\pm1} e^{-\frac{B}{2k_BT}\sigma_1}\left[e^{\frac{J}{k_BT}\sigma_1\sigma_2-\frac{B}{2k_BT}(\sigma_1+\sigma_2)}\right]\left[e^{\frac{J}{k_BT}\sigma_2\sigma_3-\frac{B}{2k_BT}(\sigma_2+\sigma_3)}\right]\cdots\left[e^{\frac{J}{k_BT}\sigma_{N-1}\sigma_N-\frac{B}{2k_BT}(\sigma_{N-1}+\sigma_N)}\right]e^{-\frac{\beta}{2k_BT}\sigma_N} \tag{4.13}$$

Why would we do this? Because we can take the objects inside the square brackets and call it a 2x2 matrix T_{ab} (with elements that depend on whether the spins are up or down):

$$T_{ab} = e^{\frac{J}{k_BT}\sigma_a\sigma_b - \frac{B}{2k_BT}(\sigma_a+\sigma_b)} \tag{4.14}$$

where T_{11} has $\sigma_a=\sigma_b=1$, T_{22} has $\sigma_a=\sigma_b=-1$ and so on. Explicitly:

$$T_{ab} = \begin{pmatrix} e^{\frac{J-B}{k_BT}} & e^{-\frac{J}{k_BT}} \\ e^{-\frac{J}{k_BT}} & e^{\frac{J+B}{k_BT}} \end{pmatrix} \tag{4.15}$$

We can then write the partition function as:

$$Z = \sum_{\sigma_1=\pm1}\sum \sigma_N = \pm 1 e^{-\frac{B}{2k_BT}\sigma_1}(T^{N-1})_{\sigma_1\sigma_N}e^{-\frac{B}{2k_BT}\sigma_N} \tag{4.16}$$

what does this buy us? Because T_{ab} is real and symmetric, it can be diagonalized by a similarity transform. So, T can be written as:

$$T = R\Lambda R^T = \begin{pmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix} t_+ & 0 \\ 0 & t_-\end{pmatrix}\begin{pmatrix}\cos\theta & \sin\theta \\ -\sin\theta & \cos\theta\end{pmatrix} \tag{4.17}$$

with R rotation matrices. The partition function then becomes:

$$Z = \sum_{\sigma_1=\pm1,\sigma_N=\pm1} e^{-\frac{B}{2k_BT}\sigma_1} R\begin{pmatrix} t_+^{N-1} & 0 \\ 0 & t_-^{N-1}\end{pmatrix} R^T e^{-\frac{B}{2k_BT}\sigma_N} \tag{4.18}$$

where we have used that $RR^T = R^TR = \mathbb{I}$ and so all but the first and last rotation matrices cancel.

Now, what happens here? Since N we take to be large, one of the eigenvalues will grow much more and will dominate. If we are interested in taking a logarithm of Z and choosing the part that grows like N, we can just look at the larger eigenvalue. We don't really have to calculate the details if we just figure out what the larger of the two eigenvalues of the transfer matrix is; we can just consider:

$$F = -k_BTN\ln t_+. \tag{4.19}$$

But we are out of time, so we leave this to next class...

5 The 1-D Ising Model, Continued

Recall we were studying the 1D Ising Hamiltonian:

$$H = -\sum_{n=1}^{N-1} J\sigma_n\sigma_{n+1} - B\sum_{n=1}^{N}\sigma_n. \tag{5.1}$$

This is one of the few exactly solvable cases with a nonzero external B field. All of the other famous solvable Ising models have the B-fields turned off. Note that solving the 1-D case is really easy (left as an exercise) if $B = 0$. With $B \neq 0$, we use a transfer matrix technique that we finish today.

Last class we also studied the phase diagram of the Ising model; we considered the boundaries of the phase diagram, though the central areas were left undetermined. In 1D the center of this phase diagram is not very interesting at all (and no interesting phase transition is really observed) but it is still valuable to consider to get used to the model and to explore the analytical techniques to solve it.

5.1 Large-field case

Note that in the case where $B \gg J$, the partition function is very simple as we can neglect the exchange/J term in the energy:

$$Z = \sum_{\sigma_n}\prod_n e^{\frac{B}{k_B T}\sigma_n} = \left(2\cosh\frac{B}{k_B T}\right)^N \tag{5.2}$$

so:

$$F = -k_B T\ln Z = -k_B T N\ln(2\cosh\frac{B}{k_B T}) \tag{5.3}$$

with magnetization per spin:

$$m = \frac{M}{N} = \frac{-\left.\frac{\partial F}{\partial B}\right|_{N,T}}{N} = \tanh(\frac{B}{k_B T}) \tag{5.4}$$

which is ± 1 (i.e. full magnetization upwards/downwards) as $B \to \pm\infty$.

5.2 The Transfer Matrix and Solving 1-D Ising Model

We recall we found the transfer matrix:

$$T_{ab} = \begin{pmatrix} e^{\frac{J+B}{k_B T}} & e^{-\frac{J}{k_B T}} \\ e^{-\frac{J}{k_B T}} & e^{\frac{J+B}{k_B T}} \end{pmatrix} \tag{5.5}$$

and with this we were able to express the partition function for the (general) 1-D Ising model partition function as:

$$Z = \begin{pmatrix} e^{\frac{B}{2k_B T}} & e^{-\frac{B}{2k_B T}} \end{pmatrix} T^{N-1} \begin{pmatrix} e^{\frac{B}{2k_B T}} \\ e^{-\frac{B}{2k_B T}} \end{pmatrix} \tag{5.6}$$

we used open boundary conditions here; but it would also be possible to fix the orientations of spins at the edges of the chain, or to use periodic boundary conditions where we identify the first and last spin of the chain. In the limit of large N, the boundary conditions are not relevant and the boundary conditions do not matter.

So how do we go about analyzing this partition function? Well, T_{ab} is a real symmetric matrix. A real symmetric matrix can be diagonalized via a similarity transformation:

$$T = R\Lambda R^T = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} t_+ & 0 \\ 0 & t_- \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{5.7}$$

Then, since $RR^T = \mathbb{I}$ it follows that:

$$T^{N-1} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} t_+^{N-1} & 0 \\ 0 & t_-^{N-1} \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{5.8}$$

Now we study this when N is large; then the partition function is dominated by the larger eigenvalue t_+; we can keep terms of $O(N)$ and throw away terms of $O(1)$. We thus come to the conclusion:

$$F = -k_B T N \ln t_+. \tag{5.9}$$

Now if we wanted to find this free energy explicitly and confirm this assertion that the larger eigenvalue dominates, we could also explicitly solve this system. For this, the periodic boundary conditions may be easiest, as we can just take the trace of T^{N-1} (and we don't get any boundary terms).

Solving for the eigenvalues of T, we find:

$$t_\pm = e^{\frac{J}{k_B T}} \cosh \frac{B}{k_B T} \pm \sqrt{e^{\frac{2J}{k_B T}} \sinh^2 \frac{B}{k_B T} + e^{-\frac{-2J}{k_B T}}} \tag{5.10}$$

which we note are both real and positive. The free energy is then:

$$F = -k_B T \ln t_+^N = -NJ - k_B T N \ln \left[\cosh \frac{B}{k_B T} + \sqrt{\sinh^2 \frac{B}{k_B T} + e^{-\frac{4J}{k_B T}}} \right] \tag{5.11}$$

here $-NJ$ is very much like a "zero point energy". If we look at the limits of this expression, this matches up with the boundaries of the phase diagram that we analyzed previously! (we can find the magnetization $M = \frac{\partial F}{\partial B}\Big|_{N,T}$ and then see that it reproduces our predictions in the $T \to 0/\infty$ and $B \to \pm\infty$ limits). We also note that F is completely non-singular in the inside of the phase diagram, and with the exception at $F = 0, B = 0$ it is also completely non-singular on the boundaries of the phase diagram.

At that point, we cannot use symmetry to conclude that the magnetization is $M = 0$; instead it is history dependent. It is highly ambiguous, and the symmetry is spontaneously broken there.

5.3 Correlation Functions

So, we've solved for the thermodynamics of the system! But another line of interest in studying spin systems are correlation functions.

To start, consider $\langle \sigma_x \rangle$ (one-point correlation function). If we have periodic BCs, then this is site independent and furthermore just gives the magnetization per spin and so:

$$\langle \sigma_x \rangle = m \tag{5.12}$$

but if there are other BCs, then we need to take into account corrections:

$$\langle \sigma_x \rangle = m + \text{boundary corrections} \tag{5.13}$$

we can explicitly evaluate $\langle \sigma_x \rangle$ are using the transfer matrix formalism. E.g. for the case of open boundary conditions:

$$\langle \sigma_x \rangle = \frac{\begin{pmatrix} e^{\frac{B}{2k_B T}} & e^{-\frac{B}{2k_B T}} \end{pmatrix} T^{x-1} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} T^{N-x} \begin{pmatrix} e^{\frac{B}{2k_B T}} \\ e^{-\frac{B}{2k_B T}} \end{pmatrix}}{Z} \tag{5.14}$$

Now in principle we could evaluate this. We skip this as this is not particularly illuminating. With periodic BCs, we get rid of the boundary terms and just take the trace:

$$m = \frac{\mathrm{Tr}\left[T^{x-1}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}T^{N-x}\right]}{\mathrm{Tr}T^{N-1}} = \frac{\mathrm{Tr}\left[T^{N-1}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\right]}{\mathrm{Tr}T^{N-1}} = \frac{c_+t_+^{N-1} - c_-t_-^{N-1}}{c_+t_+^{N-1} + c_-t_-^{N-1}} \tag{5.15}$$

where we have used the cyclicity of the Trace; $\mathrm{Tr}(ABC) = \mathrm{Tr}(CAB) = \mathrm{Tr}(BCA)$.

There are then higher-point correlation functions. We consider the two-point correlation function (which already give us some interesting information):

$$\left\langle \sigma_x \sigma_y \right\rangle \tag{5.16}$$

but actually the connected correlation function will be a little more interesting to study:

$$\left\langle \sigma_x \sigma_y \right\rangle - \left\langle \sigma_x \right\rangle \left\langle \sigma_y \right\rangle \tag{5.17}$$

this in principle again we can evaluate for any boundary condition, but again let's just look at the simplest case of the periodic BCs. Then:

$$\left\langle \sigma_x \sigma_y \right\rangle = \frac{\mathrm{Tr}\left[T^{x-1}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}T^{y-x}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}T^{N-y}\right]}{\mathrm{Tr}T^{N-1}} \tag{5.18}$$

and $\left\langle \sigma_x \right\rangle$ is independent of x/the site, so $\left\langle \sigma_x \right\rangle \left\langle \sigma_y \right\rangle = \left\langle \sigma_x \right\rangle^2$ and so:

$$\left\langle \sigma_x \sigma_y \right\rangle - \left\langle \sigma_x \right\rangle \left\langle \sigma_y \right\rangle = \frac{\mathrm{Tr}\left[T^{x-1}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}T^{y-x}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}T^{N-y}\right]}{\mathrm{Tr}T^{N-1}} - \left(\frac{\mathrm{Tr}\left[T^{x-1}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}T^{N-x}\right]}{\mathrm{Tr}T^{N-1}}\right)^2 \tag{5.19}$$

Now we can plug in the diagonal form of T and evaluate all of this; at the end of the day[3] we end up with:

$$\left\langle \sigma_x \sigma_y \right\rangle - \left\langle \sigma_x \right\rangle \left\langle \sigma_y \right\rangle = \frac{t_+^{y-x}t_-^{N-1-(y-x)} + t_-^{y-x}t_+^{N-1-(y-x)}}{t_+^{N-1} + t_-^{N-1}} \tag{5.20}$$

Now we notice a symmetry; the connected correlation function only depends on the distance from y to x, but not the actual sites themselves! Now, if we assume that the sites y, x are sufficiently close, i.e. we work in the limit:

$$N \gg (y - x) \geq 1 \tag{5.21}$$

the above simplifies to:

$$\left\langle \sigma_x \sigma_y \right\rangle - \left\langle \sigma_x \right\rangle \left\langle \sigma_y \right\rangle \approx \left(\frac{t_-}{t_+}\right)^{y-x} = e^{-|x-y|/\xi} \tag{5.22}$$

where ξ is the *correlation length*. To get this correlation length, we can take a logarithm of both sides:

$$\xi = \frac{1}{\ln(t_+/t_-)} \tag{5.23}$$

which we can then find explicitly by plugging in the form of the eigenvalues.

[3] I take home my hard earned pay all alone...

5.4 The 1-D Ising Model is Disordered at Finite Temperature

Landau gives the following argument for why the Ising model is always disordered in one dimension, that is to say, why is $m = 0$ in 1-D on the $B = 0$ line for all $T > 0$.

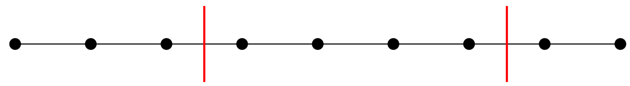

Figure 5.1: Cartoon of the setup for Landau's argument for disorder in the 1D Ising model. We consider flips in the bonds between spins (domain walls) in the chain of spins.

The Hamiltonian is minimized if all spins are aligned. The next lowest energy state is one where there is a single flipped spin. Compared to the lowest energy state, this has energy penalty $-2J$ and so has a boltzmann factor e^{-2J/k_BT} that suppresses the probability of this happening. So up to an overall factor e^{JN/k_BT} which takes into account the zero point energy, the first twoo terms in the partition function (the low-T expansion) is:

$$Z \sim 1 + Ne^{-2J/k_BT} \tag{5.24}$$

with the N appearing as these are the possible bonds (domain walls) between spins that could be flipped. Then we add the next term for 2 flips, which yields:

$$Z \sim 1 + Ne^{-2J/k_BT} + \binom{N}{2} e^{-4J/k_BT} \tag{5.25}$$

and in general the term with n flipped bonds has coefficient $\binom{N}{n}$ and so:

$$Z \sim 1 + Ne^{-2J/k_BT} + \binom{N}{2} e^{-4J/k_BT} + \ldots + \binom{N}{n} e^{-2nJ/k_BT} + \ldots \tag{5.26}$$

now, if we look at the most likely distribution, we see that the $n = N/2$ unmagnetized state is favoured. The entropy grows faster than the number of bonds, vs. the energy grows proportionally than the number of bonds; so the entropy wins out and we see that the unmagnetized state is favoured.

Why does this argument fail in higher dimensions? Because the domain walls scale differently in 2D. The entropy grows as the number of closed paths (rather than as the number of points on the line) which means that the above argument does not apply in the same way (and we do get a phase transition in higher dimensions)!

Lots of cool stuff that Ising models can be tied to... e.g. Majorana fermions in 2D, random surface theories, fermionic string theories in 3D...

5.5 Teaser - Infinite Range Ising Model

We look at another analytically solvable Ising model. Here, we remove the restriction that the spins only interact with their nearest neighbours. Specifically, we consider the case where every spin interacts with every other spin with the same interaction strength. While this is not at all a physically realistic system (because interactions are not infinite range in real life, of course) but we will see that the phase diagram of this model is more interesting, with phase transitions occuring at finite T across $B = 0$.

Our Hamiltonian is:

$$H = -\frac{J}{N} \sum_{n,n'} \sigma_n \sigma_{n'} - B \sum_n \sigma_n \tag{5.27}$$

where the sum $\sum_{n,n'}$ runs over all pairs of spins in the lattice. We are able to evaluate this by summing over one spin at a time (with each spin giving an equal contribution to the overall partition function), because all spins look the same here:

$$Z = \sum_{\sigma_n} e^{\frac{J}{k_B T} \sum_{n,n'} \sigma_n \sigma_{n'} + \frac{B}{k_B T} \sum_n \sigma_n} \tag{5.28}$$

This looks awful, but we can consider an analogy with a Gaussian integral. Consider:

$$\int_{-\infty}^{\infty} dm e^{-\left(m - \frac{1}{N} \sum_x \sigma_x\right)^2 \frac{JN}{k_B T}} \tag{5.29}$$

We can solve this by translating the integral to get rid of $\sum_x \sigma_x$ which is a constant w.r.t. the variable of integration. The integral then evaluates to:

$$\int_{-\infty}^{\infty} dm e^{-\left(m - \frac{1}{N} \sum_x \sigma_x\right)^2 \frac{JN}{k_B T}} = \sqrt{\frac{k_B T}{NJ} \pi} \tag{5.30}$$

So then:

$$1 = \sqrt{\frac{NJ}{\pi k_B T}} \int_{-\infty}^{\infty} dm e^{-\left(m - \frac{1}{N} \sum_x \sigma_x\right)^2 \frac{JN}{k_B T}} \tag{5.31}$$

Now let us put this factor of 1 into the partition function. We will then see that terms cancel, and we get an integral that we can solve - if N is large - using a saddle point integration technique. From this we can obtain the free energy and learn the thermodynamic properties of the system.

6 The Infinite Range Ising Model

6.1 Motivation - Mean Field Theory

The Hamiltonian of the infinite range Ising model is:

$$H = -\frac{J}{N}\sum_{x,y}\sigma_x\sigma_y - B\sum_x \sigma_x \tag{6.1}$$

The sum is taken over all pairs of spin, so every spin interacts with the same interaction strength as every other spin. This is obviously unrealistic physically, but it is of interest to study as the solution gives us insight into mean field theory, which is a approximation technique that can tell us a lot about second-order phase transitions.

Why is this? We can move the sum $\frac{1}{N}\sum_y$ inside of the $\sum_x$ summation, and then this looks like the spin σ_x interacts with the mean/average spin $\frac{1}{N}\sum_y \sigma_y$ (this is analytically true for this model). Note that mean field theory is wrong in general (it is an approximation technique) but it is a widely used tool (e.g. in condensed matter theory).

6.2 Gaussian Integral

Recall from our teaser last time the identity:

$$1 = \left(\frac{NJ}{\pi k_B T}\right)^{1/2}\int_{-\infty}^{\infty} d\varphi e^{-\frac{NJ}{k_B T}\varphi^2} \tag{6.2}$$

How is this integral evaluated? This is a classic integral; let's go through it together. Let's consider the general (1-D) Gaussian integral:

$$\int_{-\infty}^{\infty} dx e^{-ax^2} \tag{6.3}$$

The trick is to write the above as a square root of its square:

$$\int_{-\infty}^{\infty} dx e^{-ax^2} = \sqrt{\left(\int_{-\infty}^{\infty} dx e^{-ax^2}\right)^2} = \sqrt{\left(\int_{-\infty}^{\infty} dx e^{-ax^2}\right)\left(\int_{-\infty}^{\infty} dy e^{-ay^2}\right)} \tag{6.4}$$

Now this looks like a 2-dimensional integral; one we can solve by going into polar coordinates!

$$\int_{-\infty}^{\infty} dx e^{-ax^2} = \sqrt{\int_{-\infty}^{\infty} dx dy e^{-a(x^2+y^2)}} = \sqrt{\int_0^{2\pi} d\varphi \int_0^{\infty} r dr e^{-ar^2}} \tag{6.5}$$

Now making the substitution $u = ar^2$, $du = 2ardr$, the integral is easily solved:

$$\int_{-\infty}^{\infty} dx e^{-ax^2} = \sqrt{2\pi \frac{1}{2}\frac{1}{a}} = \sqrt{\frac{\pi}{a}} \tag{6.6}$$

which is the familiar result.

6.3 Solving the Infinite Range Ising Model

Now, let us take this result, and consider that we can shift the variable of integration $\varphi \to \varphi + \sum_x \sigma_x$ (plus a constant - does not change the result of the integral as the range of integration is $-\infty$ to ∞):

$$1 = \left(\frac{NJ}{\pi k_B T}\right)^{1/2}\int_{-\infty}^{\infty} d\varphi e^{-\frac{NJ}{k_B T}\left(\varphi - \frac{1}{N}\sum_x \sigma_x\right)^2} \tag{6.7}$$

Now, recall back to the partition function for the model we wish to compute:

$$Z[T,B,N] = \sum_{\text{spins}} e^{\frac{J}{Nk_BT}\sum x,y\sigma_x\sigma_y + \frac{B}{k_BT}\sum_x \sigma_x} \tag{6.8}$$

Now let us insert a factor of 1 as in Eq. (6.7); this will yield some useful cancellations:

$$Z[T,B,N] = \left(\frac{NJ}{\pi k_B T}\right)^{1/2} \int_{-\infty}^{\infty} d\varphi e^{-\frac{NJ}{k_BT}\varphi^2} \sum_{\text{spins}} e^{\frac{B+2J\varphi}{k_BT}\sum_x \sigma_x} \tag{6.9}$$

Now, the sum over spins can be carried out explicitly as we can just independently sum over and take the product:

$$\sum_{\text{spins}} e^{\frac{B+2J\varphi}{k_BT}\sum_x \sigma_x} = \left(e^{\frac{B+2J\varphi}{k_BT}} + e^{-\frac{B+2J\varphi}{k_BT}}\right)^N \tag{6.10}$$

Recognizing the identity $e^x + e^{-x} = 2\cosh(x)$, our partition function becomes:

$$Z[T,B,N] = \left(\frac{NJ}{\pi k_B T}\right)^{1/2} \int_{-\infty}^{\infty} d\varphi e^{-N\left[\frac{J}{k_BT}\varphi^2 - \ln(2\cosh(\frac{B+2J\varphi}{k_BT}))\right]} \tag{6.11}$$

So we are left with an integral that is hard. What do we do? Remember that we are doing stat mech/thermo here, so $N \to \infty$. In this case, we can evaluate this integral by the saddle point technique. We replace the integral with where the integrand is maximized[4]:

$$Z[T,B,N] \approx e^{-N\inf_\varphi\left[\frac{J}{k_BT}\varphi^2 - \ln(\cosh(\frac{B+2J\varphi}{k_BT}))\right] + N\ln 2} \tag{6.12}$$

where inf (infimum) denotes the greatest lower bound. So then the Helmholtz free energy is:

$$F[T,B,N] = k_BTN\inf_\varphi\left(\frac{J\varphi^2}{k_BT} - \ln\cosh\frac{B+2J\varphi}{k_BT}\right) - k_BTN\ln 2 \tag{6.13}$$

So, we should try to find this infimum/minimum; of course this is a calculus problem, which we can solve by taking the derivative and setting the expression to zero:

$$\frac{2J}{k_BT}\varphi - \frac{2J}{k_BT}\tanh\frac{B+2J\varphi}{k_BT} = 0 \tag{6.14}$$

We then look for solutions to the expression:

$$\varphi = \tanh\frac{B+2J\varphi}{k_BT} \tag{6.15}$$

which when graphically plotted, appears as in Fig. 6.1.

It should be clear from the graph and the behavior of tanh at $\pm\infty$ (going to ± 1) that we are guaranteed to have at least one solution to this expression (note that for a smaller range of Bs, it is also possible to have three solutions). If we only have one solution, it is easy to determine which is the minimum. If we have three, we have to decide between the three; but the minimum will still be the furthest to the right (except in the case when $B = 0$, in which case the leftmost and rightmost solutions are degenerate).

[4]Note: we are allowed to search in the complex plane; more formally, we have an integral over the real line, the contour which we then deform off the real axis such that the contour follows the "mountain passes" in the complex landscape. Here this is not so complicated, as we have real solutions.

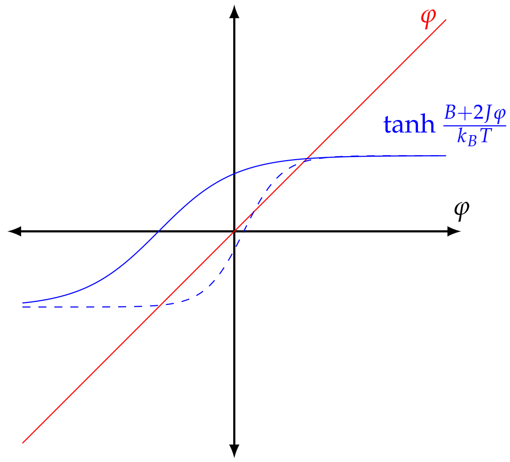

Figure 6.1: Plot of the LHS (red) and RHS (blue - solid line with $\frac{B}{k_BT} = \frac{2J}{k_BT} = 1$) of Eq. (6.15) as a function of φ. There always exists one solution to the transcendental equation, and depending on B, there may exist 3 solutions (blue - dashed line with $\frac{B}{k_BT} = -\frac{1}{4}, \frac{2J}{k_BT} = 2$).

Note that if we take the second derivative and plug in our solution at the minimum, we get the curvature at the minimum:

$$\frac{2J}{k_BT} - \left(\frac{2J}{k_BT}\right)^2 + \left(\frac{2J}{k_BT}\right)^2 \tanh^2 \frac{B+2J\varphi}{k_BT} \tag{6.16}$$

So, if the temperature is large enough, that is, $\frac{2J}{k_BT} - \left(\frac{2J}{k_BT}\right)^2$ is always positive, since the $\tanh^2$ term is always positive the second derivative of the function is always positive. So, the function is convex, and there exists one unique minimum. Specifically, this occurs when

$$T > \frac{2J}{k_B} = T_c. \tag{6.17}$$

Conversely, when $T < T_c$ the first term can fluctuate between being positive and negative and we get two minima (Mexican hat).

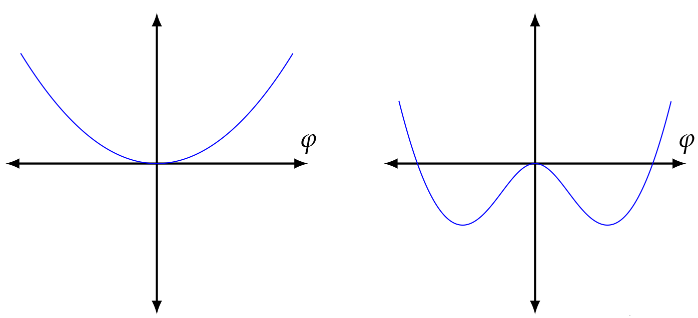

Figure 6.2: Plot of $\frac{J\varphi^2}{k_BT} - \ln\cosh\frac{B+2J\varphi}{k_BT}$ in the cases where $T > T_c$ (left - with $B = 0$ and $\frac{J}{k_BT} = \frac{1}{5}$) and $T < T_c$ (right - with $B = 0$ and $\frac{J}{k_BT} = 1$). I made a desmos graph here where you can play with J, B and see how this changes the function.

This tells us that something "happens" at $T = T_c$ (specifically, a phase transition)! Note if we set $B = 0$, then the Hamiltonian has higher symmetry $\mathbb{Z}_2$; the Hamiltonian is unchanged under spin flips. So in this case the magnetization goes to zero (which we will see shortly can be identified with φ). And we can also show mathematically that if $B = 0$ then the magnetization goes to zero. But if $B \neq 0$ we can have $T < T_c$, in which case $\varphi = 0$ is actually a local maximum. So if $T < T_c$ we see ferromagnetism; spontaneous magnetization occurs due to the turning on of a magnetic field and a decrease in temperature.

Sketching the phase diagram for the infinite range Ising model, we have a significant difference to the standard $D = 1$ Ising model; the first phase transition persists at finite temperature at $B = 0$ (where we have the crossover of magnetization, as we go from one minimum of finite magnetization to the other (with opposite sign)). At $T = T_c$, the two minima solutions merge; they don't dissapear (really), but they go off into the complex plane where we don't have to worry about them anymore, and the $B = 0$ maxima becomes the minimum. We will see in a few moments that this is a second order phase transition!

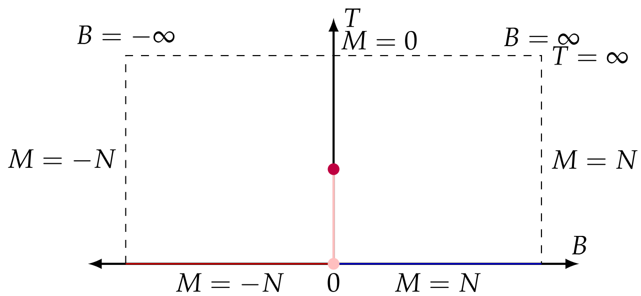

Figure 6.3: Sketch of the phase diagram of the infinite range Ising model. The first order phase transition at $B = 0$ persists at finite temperature, and at the critical temperature $T = T_c$ we have a second order phase transition.

For a second order phase transition, the free energy is smoother at the transition point; the first derivative exists.

6.4 Analysis Near the Critical Temperature

Can we be a little more analytic here? It's not so simple as we have these awful functions we are trying to minimize. If $\frac{B}{k_B T}$ is "small" and $T \sim T_c$ then φ is also small. Then, we are able to approximate these formulas by Taylor expanding in φ and B. An approximate form of the free energy reads:

$$F[T, B, N] = N k_B T \inf_{\varphi} \left[\frac{J\varphi^2}{k_B T} - \frac{1}{2} \left(\frac{2J\varphi}{k_B T} \right)^2 + \frac{1}{12} \left(\frac{2J\varphi + B}{k_B T} \right)^4 + \dots \right] - k_B T N \ln 2 \tag{6.18}$$

where we have used $\ln \cosh x = \frac{x^2}{2} + \frac{x^4}{12} + O(x^6)$. Let's analyze this expression when $B = 0$:

$$F[T, B = 0, N] = -N \inf_{\varphi} \left[J \left(1 - \frac{T_c}{T} \right) \varphi^2 + \frac{1}{12} \left(\frac{T_c}{T} \right)^4 \varphi^4 + \dots \right] \tag{6.19}$$

where we have absorbed $\frac{1}{k_B T}$ into the expression inside of the infinum and used that $T_c = \frac{2J}{k_B}$. When $T > T_c$, the solution is easily read off as $\varphi = 0$. When $T < T_c$, things are a bit trickier; taking a derivative by φ^2 and setting things to zero we get:

$$\frac{J}{k_B T} \left(1 - \frac{T_c}{T} \right) + \frac{1}{6} \left(\frac{T_c}{T} \right)^4 \varphi^2 = 0 \tag{6.20}$$

so then:

$$\varphi = \pm \sqrt{K \left(-1 + \frac{T_c}{T} \right)} \tag{6.21}$$

where K is some number. The free energy is then:

$$F = \begin{cases} C & T > T_c \\ C + K\left(\frac{T_c}{T} - 1\right)^2 & T < T_c \end{cases} \tag{6.22}$$

What we are really interested in here is how many derivatives we can take here before things go awry. It is clearly 2; when we take the first derivative, things are still continuous at T_c where $F' = 0$. Taking two derivatives, we find that this is no longer continuous where at T_c there is a jump from $F'' = 0$ to F'' some positive nonzero constant.

6.5 Magnetization and Critical Exponents

Now, let's try to interpret φ physically. It turns out to be the magnetization. One can argue this from the initial expression, or by calculating the magnetization explicitly:

$$M = -\left.\frac{\partial F}{\partial B}\right|_{T,N} = N \tanh \frac{2J\varphi + B}{k_B T} = N\varphi \tag{6.23}$$

where we have used Eq. (6.15) in the last equality. From this we conclude:

$$m = \frac{M}{N} = \varphi. \tag{6.24}$$

So all of our analysis of φ was indeed an analysis of a physical quantity.

Another thing we can do is to set T_c and approach the phase transition by varying B. We note that:

$$m(T = T_c) \sim B^{1/3} \tag{6.25}$$

$$m(B = 0) \sim (T_c - T)^{1/2} \quad T < T_c \tag{6.26}$$

$$F \sim (T_c - T)^2 \quad T < T_c \tag{6.27}$$

what we are interested in is the scaling behavior of T; these are known as critical exponents. Specifically here we are finding the critical exponents of mean field theory. This is quite profound; if a transition is described by mean field theory, they have the same critical exponents. Phase transitions with the same critical exponents have the same universality class. This was realized by Landau, who formalized the Landau theory of phase transitions. There is not much missing here from the Landau theory; the only thing which was missing is the correlation length of the system (which makes no sense for the infinite range Ising model, anyway). Note that (for systems where it is meaningful to discuss correlation lengths) the correlation lengths diverge at the phase transition, and the way these diverge are of interest indeed.

Note that it was unclear as to whether Landau theory was correct or wrong for a while; however when the 2D Ising model was solved, it was found to not have the same critical exponents as the Landau prediction (note that this does not mean that Landau theory is incorrect, only that 2D Ising lies in a different universality class. Note that for $D > 4$, Ising models are in fact well described by Landau theory/mean field theory).

Let us consider some parameters/quantities that are of interest to study when we consider phase transitions. We call $\boxed{m}$ the *order parameter* in the Landau framework; the expectation value of the magnetization (per spin) tells us about the symmetry of the system (and, here, the fact that the $\mathbb{Z}_2$ symmetry is spontaneously broken at $B = 0$). We also have the (normalized) distance $\boxed{\tau}$ from the critical temperature:

$$\tau = \frac{T - T_c}{T_c} \tag{6.28}$$

We have the magnetic field $\boxed{B}$. We have the susceptibility $\boxed{\chi}$ which is:

$$\chi \sim \frac{\partial m}{\partial B} \tag{6.29}$$

i.e. if we vary the field, how does the magnetization change. At the critical point, this tends to diverge. We also have the correlation length $\boxed{\xi}$ which is given by:

$$\left\langle \sigma_x \sigma_y \right\rangle \sim e^{-\frac{1}{\xi}|x-y|} \tag{6.30}$$

where $\boxed{\left\langle \sigma_x \sigma_y \right\rangle}$ is a correlation function. If we drop down from the infinite range model to the standard Ising model, another parameter to consider is the dimension of the model $\boxed{D}$. We also have the specific heat $\boxed{C}$:

$$C \sim \frac{T}{N}\frac{\partial^2 F}{\partial T^2} \tag{6.31}$$

We also have critical exponents (a whole family of them...). For above the phase transition $\tau > 0$, we have:

$$C \sim \tau^{-\alpha} \tag{6.32}$$

$$\tau \sim \tau^{-\gamma} \tag{6.33}$$

$$\xi \sim \tau^{-\nu} \tag{6.34}$$

below the phase transition $\tau < 0$ we have the same idea, but we put a prime on things to distinguish:

$$C \sim \tau^{-\alpha'} \tag{6.35}$$

$$\tau \sim \tau^{-\gamma'} \tag{6.36}$$

$$\xi \sim \tau^{-\nu'} \tag{6.37}$$

We also have that the order parameter scaling as:

$$m \sim \tau^{\beta} \tag{6.38}$$

for $\tau = 0, B = 0$, we have:

$$B \sim m^{\delta} \tag{6.39}$$

$$\left\langle \sigma_x \sigma_y \right\rangle \sim \frac{1}{|x-y|^{-d+2-\eta}} \tag{6.40}$$

What we have found through our analysis is a bunch of these critical exponents. Mean field theory makes the following predictions:

$$\alpha = \alpha' = 0 \tag{6.41}$$

$$\beta = \frac{1}{2} \tag{6.42}$$

$$\gamma = \nu' = 1 \tag{6.43}$$

$$\delta = 3 \tag{6.44}$$

We don't learn things about the correlation lengths from this model as the interactions are infinite range; but this can be fixed up via Landau-Ginsberg theory, which tells us:

$$\eta = 0 \tag{6.45}$$

$$\nu = \frac{1}{2} \tag{6.46}$$

Experimental values for comparison can be compared with the table posted on the course webpage. These are not that easy to measure because they involve the scaling of something at a critical point. But some can be measured; for example α for the λ-transition (Bose condensation transition) in liquid Helium-4 is experimentally determined to be:

$$\alpha = -0.0127 \pm 0.0003 \tag{6.47}$$

so it is known to be nonzero. So mean field theory does not quite describe this transition, though it comes close. In the theory, $\alpha = 0$ as $F \sim T^2$ and hence specific heat is a constant.

7 Landau Theory

7.1 Introduction and Motivation

Last time, we discussed the infinite-range Ising model. It is a nice model that we can write down and solve analytically. We ended last lecture by calculating the critical exponents of the model. This model does have some shortcomings; of course it is completely physically unrealistic, and it also does not let us calculate correlation functions (and there are some critical exponents that come from these). A lot of the discussion of it was poached from Pairisi's book on statistical field theory; there he has a hand-wavey argument about how to get correlation functions from the model (which Gordon does not understand - so we do something else)!

We note that if we are interested in the behaviour near the singular points (i.e. characterizing the critical exponents) then a lot of the fine-grained details of the model does not matter. We introduced Landau theory, which predicts the same (mean-field) critical exponents as the infinite range Ising model, but also can do more - e.g. providing us with the correlation function for spins. The main attribute which we discussed and which we will continue to discuss is the correlation length, which describes the (exponential) degree to which correlations die off with distance[5] The correlation length has a power law behaviour, and the exponent on the $\frac{1}{|x-y|}$ also is a critical exponent.

7.2 The Landau Potential and Free Energy

To discuss correlation functions, we require some position-dependent behavior. A motivation for how to do this - iron has some microscopic (fine-grained) crystal structure, but macroscopically the iron looks like a continuum and we cannot see the lattice. In this limit, we see the magnetization smoothly varying like a continuous field. So, let us write down the Landau potential:

$$\Gamma[\varphi] = \int dx \left[\tau \frac{\varphi^2(x)}{2} + \lambda \frac{\varphi^4(x)}{4!} - B(x)\varphi(x) + \frac{1}{2}\nabla\varphi \cdot \nabla\varphi\right] \tag{7.1}$$

where $\tau \sim \left(\frac{T}{T_c} - 1\right)$, λ is some (positive) coupling constant, and B is the magnetic field strength. The first three terms already looks like a Hamiltonian, but we want something to "smooth it out" as with just those three each point is independent of one another and there is nothing stopping some very large fluctuations in φ. So the last term penalizes this, by ensuring that large fluctuations are unfavourable. Note that Landau theory is also able to acommodate different kinds of spins, e.g. not just ± 1 at each site but pointing in some arbitrary direction (but for now let us just stick with the simplest case).

Now, the idea is to get the free energy as:

$$F = \inf_{\varphi} \Gamma[\varphi] \tag{7.2}$$

and this gets us into the realm of variational calculus/functional calculus. This is because Γ is a functional - it is a function of the field φ which itself is a function of position. We wish to minimize Γ over all φs, which requires some functional calculus in general; but here this is not necessary.

If $B \leq 0$ and $\tau > 0$, then all the terms are positive and so the easy minimum is taking $\varphi = 0$.

If $B = 0$ and $\tau < 0$, then now things are not so trivial. In this case we write the potential as:

$$\Gamma = \int dx \left[\frac{1}{2}\nabla\varphi \cdot \nabla\varphi + \frac{\lambda}{4!}\left(\varphi^2 + \frac{\tau 4!}{4\lambda}\right)^2 - \frac{\lambda}{4}\left(\frac{\tau 4!}{4\lambda}\right)^2\right] \tag{7.3}$$

[5] We will have the machinery to explain why this decay is exponential by the end of the course. For now, we take it as a result.

Here, the potential is minimized if the first two terms vanish (the last term is just a constant). The first term is minimized if φ is a constant, and the second (positive) term is then minimized by taking it to zero:

$$\varphi = \pm\sqrt{\frac{4!}{4\lambda}(-\tau)} \tag{7.4}$$

and here we already recover the square root scaling of the magnetization (i.e. the critical exponent of $\frac{1}{2}$) that we discussed last class.

Now, let us plug the minimizing φ back into Γ and solve for the free energy F:

$$F = \begin{cases} 0 & B \leq 0, \tau > 0 \\ V\frac{\lambda}{4}\left(\frac{4!}{4\lambda}\right)^2\tau^2 & B = 0, \tau < 0 \end{cases} \tag{7.5}$$

7.3 Finding Critical Exponents; Specific Heat, Equation for φ, Susceptibility

Now, the specific heat can be obtained as:

$$C = -T\frac{\partial^2 F}{\partial T^2} \sim \begin{cases} |\tau|^{-\alpha} & \tau > 0 \\ |\tau|^{-\alpha'} & \tau < 0 \end{cases} \quad \alpha = \alpha' = 0 \tag{7.6}$$

where the critical exponents are found to be zero because there is no singularity! Again we recover the mean-field critical exponents from alst time.

The next critical exponents are not quite as simple to calculate, but they also aren't that bad. We consider the equation for φ; this can be obtained by doing a variational minimization of Γ:

$$(-\boldsymbol{\nabla}^2 + \tau)\varphi + \frac{\lambda}{3!}\varphi^3 - B = 0 \tag{7.7}$$

If we assume φ is a constant and $\tau = 0$ (i.e. $T = T_c$) then we find:

$$\varphi \sim B^{1/3} \tag{7.8}$$

or equivalently:

$$B \sim \varphi^3 \tag{7.9}$$

which is another critical exponent (that we recover last time). Every exponent we derive here will agree with what we found last class, and we can further define more.

Recall the susceptibility, defined as:

$$\chi = \frac{\partial m}{\partial B} \tag{7.10}$$

Note that while φ is more general than m, after we have solved everything, we can identify $m = \varphi$. To this end, we should explore how this solution φ changes as we vary B. Again assuming that φ does not vary in position, we have:

$$\tau\varphi + \frac{\lambda}{3!}\varphi^3 - B = 0 \tag{7.11}$$

Taking the derivative w.r.t. B we find:

$$\left(\tau + \frac{\lambda}{2!}\varphi^2\right)\frac{\partial\varphi}{\partial B} - 1 = 0 \tag{7.12}$$

We then find that:

$$\chi \sim \begin{cases} \frac{1}{\tau} & \tau > 0 \\ -\frac{1}{2\tau} & \tau < 0 \end{cases} \tag{7.13}$$

this is called the *Curie-Weiss Law* (and also gives us two more critical exponents). This is not an exact fit to the experimentally measured susceptibility of Iron, but it was close enough that it did take a little while to notice a discrepancy.

7.4 Calculating Correlation Functions

To actually calculate a correlation function, we need to re-introduce spatial dependence into our field. We are interested in the eqaution:

$$\left(-\nabla^2 + \tau + \frac{\lambda}{2!}\varphi^2(x)\right)\chi(x,y) = \delta(\mathbf{x}-\mathbf{y}) \tag{7.14}$$

where:

$$\chi(x,y) = \langle m(x)m(y)\rangle - \langle m(x)\rangle\langle m(y)\rangle \tag{7.15}$$

is the connected correlation function. Note that our problem has now reduced to finding a Green function. The standard way to approach this is a plane wave ansatz/fourier transform.

Let us study the special case of $\tau = 0$ ($T = T_c$) and $\varphi = 0$. In this case, we have the very simple equation:

$$-\nabla^2\chi = \delta(\mathbf{x}-\mathbf{y}) \tag{7.16}$$

if $D = 3$, then this is just the famous Coloumb potential and so the solution is:

$$\chi = \frac{1}{4\pi|\mathbf{x}-\mathbf{y}|} \tag{7.17}$$

where here the critical exponent is $\eta = 0$ (recalling that $\chi \sim \frac{1}{|\mathbf{x}-\mathbf{y}^{-D+2-\eta}|}$).

A question to test your mathematical intuition - what do solutions to this equation look like in other dimensions? By dimensional analysis, we have:

$$\chi \sim \frac{1}{|\mathbf{x}-\mathbf{y}|^{D-2}} \tag{7.18}$$

and so $\eta = 0$ in any dimension for this special case, actually! A quick review of the dimensional analysis argument; $[\delta] = [X]^{-d}$ (i.e. inverse of the dimension) and so I want something which taking two spatial derivatives (i.e. subtract two dimensions) yields dimensions of $-D$ and so the dimensions of χ must be $-D+2$.

Away from the critical point when $\tau \neq 0$, things look bad; but then it is balanced out by $\frac{\lambda}{2!}\varphi^2(x)$... then we have to be a bit more sophisticated, and consider that the Green function decays exponentially, i.e. know that $\chi(x,y) \sim e^{-|x-y|/\eta}$. The coefficient can be obtained via dimensional analysis, again; the argument of a transcendental function must be dimensionless, and so we conslude that:

$$\eta = \frac{1}{\tau^{1/2}} \tag{7.19}$$

The last thing we have not derived is how this exponential decay comes about. We could do this by taking the fourier transform of the equation, then for large $|\mathbf{x}-\mathbf{y}|$ carry out the integral via saddle point technique which gives the exponential decay.

7.5 Parameters in Landau Theory

What are parameters/inputs into the Landau theory? Note that since the integral goes over the volume, the number of spins is implicitly included. So we won't bother even including this on our list.

(i) $\tau \sim (T - T_c)$.

(ii) Dimension D - it appears in the integral, and in the derivatives, and if we have to solve the differential equation, this of course would highly depend on the dimensionality of the model.

(iii) Symmetry and number of components of magnetization (here just one - φ is a function spits out one real number). Also, the fact that Γ has a zero is encoded by the fact that the theory is symmetric under interchange of $\varphi \leftrightarrow -\varphi$ (if one also puts $B \leftrightarrow -B$).

(iv) $\lambda > 0$ - Required for stability of the model.

This is not very many parameters, but we get a lot of output out of it!

When this is applied to a superconductor, then this becomes Landau-Ginzberg theory (one puts in the electromagnetic field in the natural way).

One more comment about Landau theory - it can be used to describe first order phase transitions, in addition to second order phase transitions. We can add more terms, then we get two (or more) extrema with an energy barrier between them. The first order phase transition (e.g. bubble nucleation) will be when the two extrema have the same energy (the crossover point when they are both the global minima). There is less to study here as there are no associated critical exponents, but it is nevertheless something that the theory is capable of accommodating.

Q - as it is, this looks identical to the massless scalar field theory with $m = \tau$ if $B = 0$. Is there a deeper correspondence? The answer will turn out to be yes, as we will learn as the term progresses (there is identical mathematical structure here, but it is worth noting that the theory here is completely classical). Somehow Landau guessed the classical limit of quantum field theory from this modelling.

7.6 $O(3)$ (Vector) Model

If we want to analyze something else, e.g. a piece of iron where the magnetization can point in some arbitrary direction in $\mathbb{R}^3$, we can modify the theory we wrote above by changing φ to be a three-component function $\boldsymbol{\varphi}$ (and $B \to \mathbf{B}$ to be a three-component field, rather than just a projection along a single axis).

This is called the $O(3)$ model as the spins can point in three directions. Note that we can still place our lattice in an arbitrary dimension. We could also allow our spins to point in more directions ($O(N)$ model). Our field is now a vector:

$$\boldsymbol{\varphi} = \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{pmatrix} \tag{7.20}$$

Our Landau potential genrealizes in the expected way:

$$\Gamma = \int d^3x \left(\frac{1}{2} \boldsymbol{\nabla}_i \varphi_j \boldsymbol{\nabla}_i \varphi_j + \frac{\tau}{2} \boldsymbol{\varphi}^2 + \frac{\lambda}{4!} \left(\boldsymbol{\varphi}^2 \right)^2 - \mathbf{B} \cdot \boldsymbol{\varphi} \right) \tag{7.21}$$

Note that we are allowed to add another term, actually; in three dimensions there is a further rotational symmetry (we can rotate the space and leave the space the same, or now we can rotate the spins and leave the space the same). So, we could add a term of the form $(\boldsymbol{\nabla} \cdot \boldsymbol{\varphi})^2$ - but this would have less symmetry (e.g. we could add this term in the case where we have phonons in the lattice, which breaks this two-fold rotational symmetry; we would have to rotate the space and the lattice. This added term corresponds to this broken two-fold rotational symmetry, as the function is no longer invariant under the two types of rotations).

To go to superconductivity, there is a breakdown of phase symmetry of the wavefunction, so we add a complex phase function (and then this becomes Landau-Ginzeberg theory, but we don't discuss this here).

7.7 Teaser - Spherical Model

Next class we will discuss how Landau theory is actually wrong. One demonstration of this was the exact solution of the 2D Ising model, which does not agree with the mean field Landau theory predictions.

The spherical model will provide us another route to showing how Landau theory is not correct. It is in some sense a simplified Ising model, and one which we will be able to solve analytically (somewhat).

The Hamiltonian is given by:

$$H = -J\sum_{x,i}\sigma_x\sigma_{x+i} - B\sum_x \sigma_X \tag{7.22}$$

where the spins sit on a regular lattice in an arbitrary dimension (a line in 1D, a square lattice in 2D, a cubic lattice in 3D, a hypercube in 4D and so on...) i labels the connections between the lattice sites, and there is a coupling of neighbours (e.g. see Fig. 4.1).

This looks like the Ising model (and has the same Hamiltonian) but it is not the Ising model. This is because we will allow σ_x to be any real number, not just ± 1, rather $\sigma_x \in (-\infty, \infty)$. This simplifies things because continuous math is easier than discrete math.

The obvious problem with this is it seems as though we can have an arbitrary amount of magnetization to the system (and hence the model is unstable); we therefore impose the constraint:

$$\sum_x \sigma_x^2 = N \tag{7.23}$$

where N is the number of sites. While it is not obvious, one is able to show that all spins pointing up or downwards is the degenerate energy minima of the system.

Next time, we will be able to "solve" the system (up to some leftover functions) and we will be able to analyze the critical exponents of the model. For dimensions $D \geq 4$ we will find that the critical exponents match exactly those of Landau/mean field theory. For $D < 4$ this is not the case.