PHYS 516 (Statical Mechanics) Notes

Rio Weil

This document was typeset on January 19, 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 **2**
1.1 Overview . . . 2
1.2 Canonical Ensemble . . . 2
1.3 Fundamental Postulate - Equal a Priori Probability . . . 3
1.4 Interlude - Deriving Stirling's Formula . . . 3
1.5 Deriving the Boltzmann Distribution . . . 4
1.6 Example: Two-level system . . . 5

2 Free Energy, Ideal Gas, and the Grand Canonical Ensemble **6**
2.1 Thermodynamic Interpretation, Energy, and Free Energy . . . 6
2.2 Example - System of weakly interacting non-relativistic particles . . . 7
2.3 Grand Canonical Ensemble . . . 9

3 Microcanonical Ensemble, Quantum Statistical Mechanics **11**
3.1 Review of the Canonical Ensemble . . . 11
3.2 Microcanonical Ensemble . . . 11
3.3 Ideal Gas - Microcanonical Ensemble Lens . . . 12
3.4 Quantum Ideal Gas . . . 13

4 Introduction to the Ising Model **15**
4.1 The Hamiltonian . . . 15
4.2 Analyzing the Ising Model - Phase Diagram Boundaries . . . 16
4.3 Analyzing the Ising Model - $B = 0$ and Spontaneous Symmetry Breaking . . . 18
4.4 Solving the 1-D model . . . 18

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).

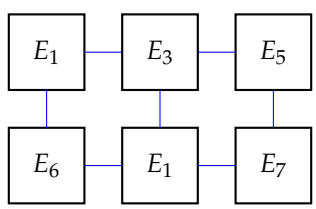

Figure 1.1: Cartoon of the Canonical Ensemble - we consider $\mathcal{N}$ ($= 6$ here) copies of a system, and weakly couple them such that they can exchange energy. Due to artistic constraints the couplings are only drawn for nearest neighbours, but in reality all the systems are coupled to each other.

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:

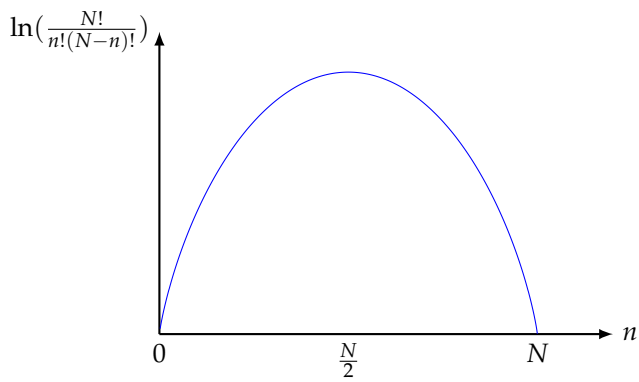

Figure 1.2: Plot of (continuum approximation) $\ln \frac{N!}{n!(N-n)!}$ as a function of number of spin-up spins n. We see the maximum at $n = N/2$ (zero magnetization).

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):

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

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:

$$P = \frac{Nk_BT}{V} \tag{2.21}$$

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

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

The entropy we calculate to be:

$$S = \frac{3}{2}k_BN + k_BN\ln\left(V\left(\frac{mk_BT}{2\pi\hbar^2}\right)^{3/2}\right) \tag{2.23}$$

We can calculate the energy to be:

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

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 $2k_BN\ln 2$; 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:

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

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):

$$PV = Nk_BT \tag{2.27}$$

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

And now the entropy is fixed up as well:

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

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:

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

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:

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

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

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

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:

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

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 $\ln \Gamma(x+1) = x \ln x - x + \ldots$ (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 $\psi(\mathbf{q}_1, \ldots \mathbf{q}_n, t)$, 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:

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

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:

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

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):

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

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

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

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

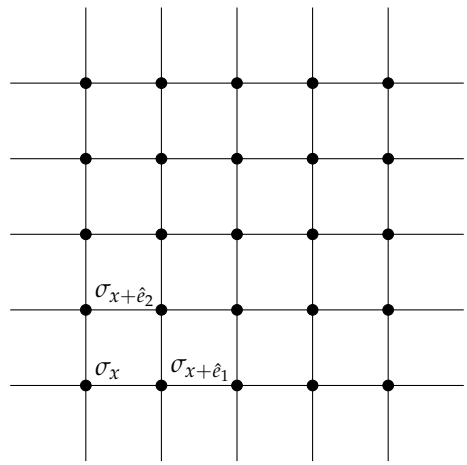

Figure 4.1: 2D Ising model on a square lattice. The spins sit on the nodes/crossings of the lattice, with interaction between neighbouring spins ($\mu = \pm\hat{e}_1, \pm\hat{e}_2$).

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:

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

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:

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

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:

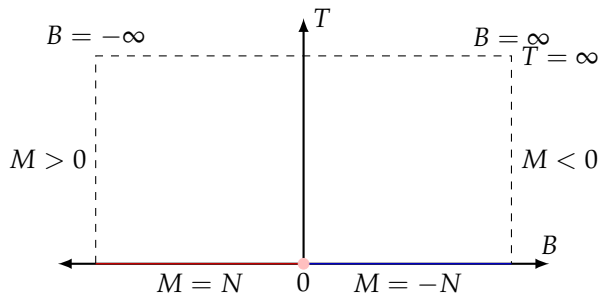

Figure 4.2: Phase Diagram Boundaries for the Ising Model, plotting temperature T and magnetization M. As $B = \pm\infty$, the spins want to align downwards ($M < 0$)/upwards($M > 0$) respectively. As $T \to \infty$, all configurations are equally likely and so $M = 0$. At $T = 0$, all spins aligning 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 the spins want to align downward so $M < 0$ and if $M = -\infty$ then the spins want to align upwards so $M > 0$. 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, $M = -\text{sign}(B)N$, 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.

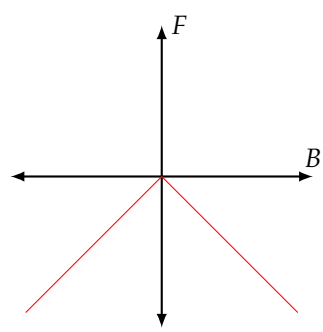

Figure 4.3: Plot of F at $T = 0$ as a function of B. The free energy is continuous but not differentiable, and hence there is a first-order phase transition at $B = 0$. The slope of the free energy is $J/-J$.

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.

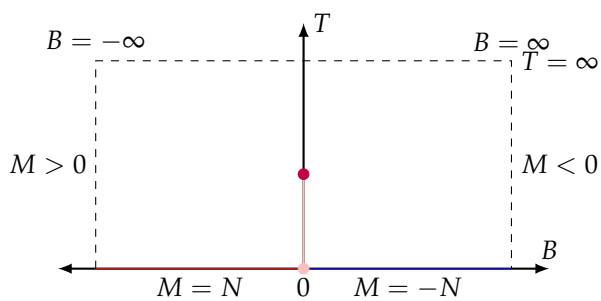

Figure 4.4: For dimensions larger than 1, $M \neq 0$ persists at $B = 0$ even as the temperature is tuned up. This results in a "first order phase transition line" partway up the $B = 0$ axis, culminating in a second order phase transition point.

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:

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

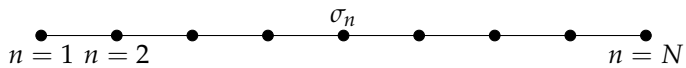

Figure 4.5: Cartoon of 1D Ising model (pictured with $N = 9$ spins). The spin at site n is denoted σ_n.

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:

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

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...