

PHYS 516 (Statistical Mechanics) Notes

Rio Weil

This document was typeset on January 10, 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

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, \dots . 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).

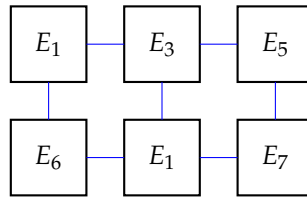


Figure 1.1: Cartoon of the Canonical Ensemble - we consider \mathcal{N} ($= 6$ here) copies of a system, and weakly couple them such that they can exchange energy.

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:

$$\mathcal{E} = \sum_a n_a E_a \quad (1.1)$$

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

$$\mathcal{N} = \sum_a n_a \quad (1.2)$$

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, \dots) ? Let us derive this. Starting with the systems with energy E_1 , we have:

$$\mathcal{N}(\mathcal{N} - 1) \dots (\mathcal{N} - n_1 + 1) \quad (1.3)$$

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 $n_1!$:

$$\frac{\mathcal{N}(\mathcal{N} - 1) \dots (\mathcal{N} - n_1 + 1)}{n_1!} \quad (1.4)$$

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

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

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:

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

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_a s 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_a s, 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:

$$\ln \mathcal{N}! \approx \mathcal{N} \ln \mathcal{N} - \mathcal{N}. \quad (1.7)$$

1.4 Interlude - Deriving Stirling’s Formula

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

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

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:

$$\mathcal{N}x^{\mathcal{N}-1}e^{-x} - x^{\mathcal{N}}e^{-x} = 0 \quad (1.9)$$

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

$$\mathcal{N}! \approx \mathcal{N}^{\mathcal{N}}e^{-\mathcal{N}} \quad (1.10)$$

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:

$$\ln \mathcal{N}! = \sum_{j=1}^{\mathcal{N}} \ln j \quad (1.11)$$

now approximating the sum as an integral:

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

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:

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

where we define $\rho_a = \frac{n_a}{\mathcal{N}}$ 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:

$$-\ln \rho_a - 1 + \beta E_a + \gamma = 0 \quad (1.14)$$

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

$$\sum_a \rho_a E_a = U \quad (1.15)$$

$$\sum_a \rho_a = 1 \quad (1.16)$$

Let us rearrange the first equation, which has solution:

$$\rho_a = e^{\beta E_a + \gamma - 1} \quad (1.17)$$

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

$$\rho_a = \frac{e^{\beta E_a}}{\sum_a e^{\beta E_a}} \quad (1.18)$$

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:

$$Z = \sum_a e^{\beta E_a} \quad (1.19)$$

So, we have solved for the most likely distribution ρ_{a_i} ; this is 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:

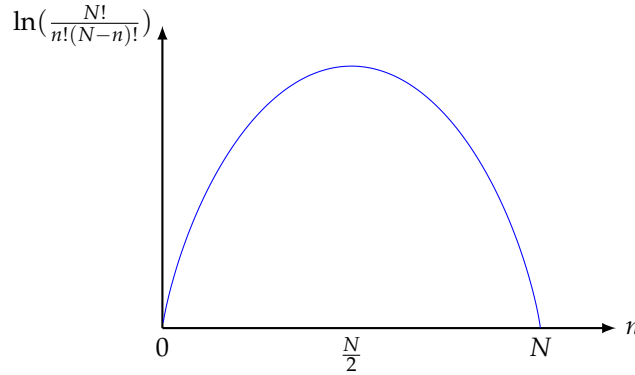


Figure 1.2: Plot of (continuum approximation) $\ln \frac{N!}{n!(N-n)!}$ as a function of number of spin-up spins n . We see the maximum at $n = N/2$ (zero magnetization).

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 :

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

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

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

so in the $N \rightarrow \infty$ 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).