Statistical mechanics : lecture 1

Introduction to the Gibbs ensemble

A note on quantum vs. classical : statistical mechanics can be constructed for both classical and quantum systems. In this course I will focus on the quantum case and derive the classical case as a particular case. The only reason why I do this is because I think quantum systems are more interesting.

Recomended reading : Feynman Stat Mech, chapter 1
Salinas, chapter 5.

Thermal equilibrium and the Gibbs ensemble

The 1st thing I want to do in this course is to introduce what it means, mathematically, to say a system is in thermal equilibrium. Once we have the math, understanding what this means physically and what are its consequences, will be much easier.

Any physical system can be described by a Hamiltonian operator H. It doesn't matter if the system is a harmonic oscillator, a gas of 10^{23} molecules or a spaceship. You can always, at least in principle, write down a Hamiltonian for it.

Let us introduce the eigenvalues and eigenvectors of such a Hamiltonian as:

$$H|n\rangle = E_n|n\rangle \,, \qquad n = 0, 1, 2, \dots \qquad (1)$$

The number of eigenvalues can be finite of infinite. For small systems these eigenvalues will in general be countable, $E_0, E_1, \dots$. On the other hand, if the system is composed of many particles, they will start to lie very close to each other, forming a quasi-continuum. Notwithstanding, every system can be described by a Hamiltonian and its corresponding eigenvalues and eigenvectors

This is all we need to define thermal equilibrium. Here is the most important result in statistical mechanics:

When a physical system with Hamiltonian (1) is in thermal equilibrium at a certain temperature T, the probability of finding it in a certain energy eigenstate $|n\rangle$ will be given by

$$\boxed{p_n = \frac{e^{-\beta E_n}}{Z}, \qquad Z = \sum_n e^{-\beta E_n}} \tag{2}$$

where $\beta = 1/k_B T$ and k_B is Boltzmann's constant. The quantity Z, called the partition function, is a normalization to ensure that $\sum_n p_n = 1$.

That's it! This is all you need to know about statistical mechanics. ☺

Eq (2) mathematically defines what thermal equilibrium is and says nothing about how the system thermalizes towards it. This, in fact, can be a very hard problem. But it turns out that, whatever the reason, many many systems in Nature have the tendency to thermalize. And once they do, Eq (2) applies.

We could, therefore, simply start applying Eq (2) blindly to a bunch of problems and explain many properties of matter. That is to say, Eq (2) can be used to understand many problems even if you don't understand the Eq itself! Weird Eh? Here we will try to alternate back and forth between the two things. I think this is the best strategy to gain intuition.

About Boltzmann's constant: the value of k_B is

$$k_B = 1.380 \times 10^{-23} \ J/K \tag{3}$$

$$= 8.617 \times 10^{-5} \ eV/K$$

Thus we see that k_B is such that $k_B T$ has units of energy. The Boltzmann constant only exists for historical reasons. We chose to measure temperature in Kelvins. We could've just as well measure temperature in units of energy. This is what $k_B T$ means. It thus makes no sense to keep using k_B. We can simply agree from now on that $k_B T \to T$ or, what is equivalent, set

$$\boxed{k_B = 1} \tag{4}$$

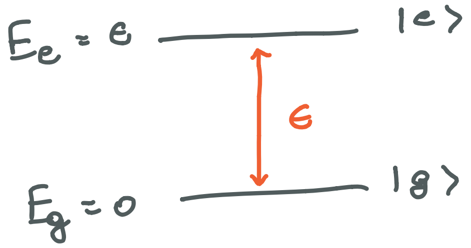
This means that, from now on, temperature is measured in electron-volts. For instance

$$T = 300K \rightsquigarrow k_B T = 0.0258 \ eV \tag{5}$$

So from now on we simply say $T = 0.0258 eV$.

Example: qubit = 2-level system = spin 1/2

Consider a system with two levels, which we label as $|g\rangle$ (for ground-state) and $|e\rangle$ (for excited state). The two states are separated by an energy gap ϵ



$$H = \epsilon\, |e\rangle\langle e| = \begin{pmatrix} 0 & 0 \\ 0 & \epsilon \end{pmatrix} \qquad (6)$$

Reminder: $|e\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\langle e| = \begin{pmatrix} 0 & 1 \end{pmatrix}$

so $|e\rangle\langle e| = \begin{pmatrix} 0 \\ 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$

If such a system is in thermal equilibrium, then the probability of finding the system in the ground and excited states, are

$$P_g = \frac{1}{Z} \qquad P_e = \frac{e^{-\beta\epsilon}}{Z}$$

where

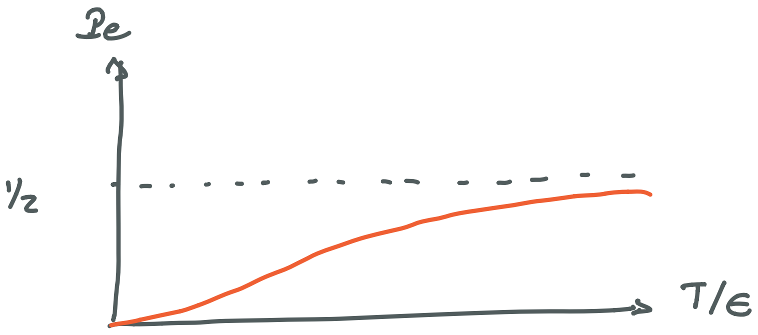
$$Z = \sum_{n=0,1} e^{-\beta E_n} = 1 + e^{-\beta\epsilon} \qquad (7)$$

It is convenient to rearrange P_e as

$$\boxed{P_e = \frac{1}{e^{\beta\epsilon} + 1}} \qquad (8)$$

This represents the probability that we find the system in the excited state. In other contexts, we will find this formula again under the name of Fermi-Dirac distribution.

The distribution P_e as a function of T/ϵ looks like

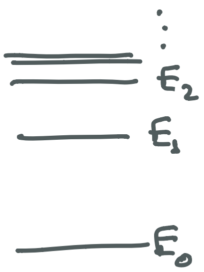


Understanding this simple plot is very very important. Think about T/ϵ as a competition between the thermal fluctuations (T) and the energy gap (ϵ) separating the two levels. When $T/\epsilon \ll 1$ it means the thermal fluctuations are very small, so that the system is with certainty sitting quietly in the ground state.

Conversely, when $T/\epsilon \gg 1$ the thermal fluctuations become significant and the probability of finding the system in the excited state starts to become significant.

But, most importantly, we see that $P_e < 1/2$. In fact, even at infinite temperature, it is always more likely to find the system in the ground state than in the excited state. This is thermodynamic stability. Notice the essential role played by the ground state. Whenever you are studying a problem in statistical mechanics, always identify the ground state!

Consider now the electronic levels of an atom (e.g. Hydrogen)



If we forget about all other levels for a second, our previous result shows that the prob. of finding the system in state E_1 will be

$$P_1 \simeq \frac{1}{e^{\beta(E_1-E_0)}+1}$$

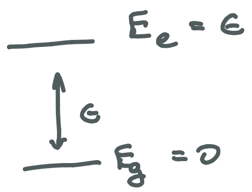
This is an approximate expression because I'm neglecting all other energy levels.

If $T=300\,K$ and $E_1 - E_0 = 1\,eV$ we get

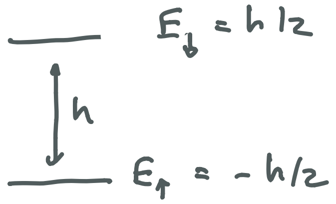
$$P_1 \sim 10^{-17}$$

The probabilities for $E_2, E_3, \ldots$ would be even insanely smaller. We thus see that the probability of finding the system away from the ground state is overwhelmingly small! That's why we don't need to worry about temperature when we constructed the periodic table in high school. We simply "put" the electrons at their pre-assigned seats $1s^1, 2s^2, 2p^6, \ldots.$

Energy is defined up to a constant. I chose to define the two levels as $E_g = 0$ and $E_e = \epsilon$. We can also shift them to be more symmetric



or



ε and h is the same thing. I only switch to h to emphasize it's like a magnetic field

This is what we usually get when the system is a spin 1/2 coupled to a magnetic field h. The Hamiltonian reads.

$$H = \begin{pmatrix} -h/2 & 0 \\ 0 & h/2 \end{pmatrix} = -\frac{h}{2}\sigma_z$$

where $\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ is a Pauli matrix. The ground state is spin up (i.e., pointing in the direction of h).

The connection between the "ge" notation and the "up-down" notation is

$$|g\rangle = |\uparrow\rangle, \qquad |e\rangle = |\downarrow\rangle$$

The partition function in this case becomes

$$Z = e^{\beta h/2} + e^{-\beta h/2} = 2\cosh(\beta h/2) \tag{9}$$

But notice how the probabilities remain the same

$$P_{\uparrow} = \frac{e^{\beta h/2}}{e^{\beta h/2} + e^{-\beta h/2}} = \frac{1}{1 + e^{-\beta h}} \equiv P_g$$

$$P_{\downarrow} = \frac{e^{-\beta h/2}}{e^{\beta h/2} + e^{-\beta h/2}} = \frac{1}{e^{\beta h} + 1} \equiv P_e$$

Energy is only defined up to a constant, so this energy shift cannot have physical consequences.

Example : quantum harmonic oscillator (QHO)

The energy levels of a quantum harmonic oscillator are given by

$$E_n = \hbar\omega\,(n + 1/2), \quad n = 0, 1, 2, \ldots \tag{10}$$

The partition function is thus

$$Z = \sum_{n=0}^{\infty} e^{-\beta E_n} = e^{-\beta\hbar\omega/2} \sum_{n=0}^{\infty} e^{-\beta\hbar\omega n}$$

If we define $x = e^{-\beta\hbar\omega}$, this sum becomes the famous geometric series

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x} \tag{11}$$

Thus, we find that

$$Z = \frac{e^{-\beta\hbar\omega/2}}{1 - e^{-\beta\hbar\omega}} \tag{12}$$

The Gibbs probabilities will then be given by

$$p_n = \frac{e^{-\beta E_n}}{Z} = \frac{(1 - e^{-\beta\hbar\omega})}{e^{-\beta\hbar\omega/2}} \, e^{-\beta\hbar\omega(n + 1/2)}$$

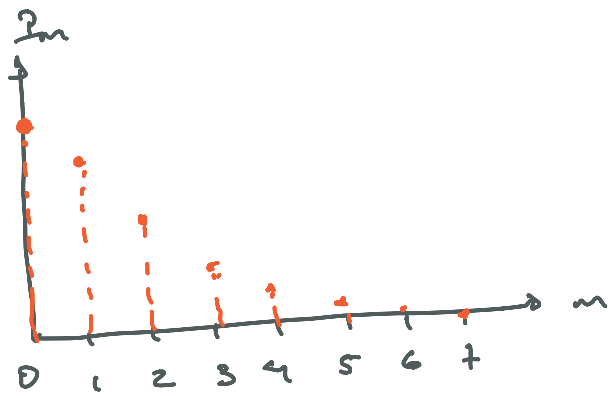
Notice how the factors $e^{-\beta\hbar\omega/2}$ will cancel out: energy is always defined up to a constant, so it is reasonable that the probabilities should not depend on what *we* choose for our zero of energy

As a result, we then get

$$\boxed{P_n = (1-e^{-\beta\hbar\omega})\, e^{-\beta\hbar\omega n}} \qquad (13)$$

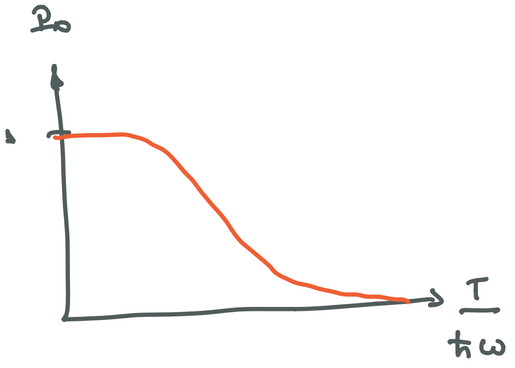
In mathematics this is known as the geometric distribution.

this distribution looks like this



Again, thermodynamic stability means it is always more likely to find the system in a lower energy state than in a high energy state.

the prob. of finding the QHO in the ground state looks like



$$P_0 = 1 - e^{-\beta\hbar\omega}$$

$$= \begin{cases} 1 & T = 0 \\ 0 & T \to \infty \end{cases}$$

Expectation values of observables

Once we have the Gibbs probabilities P_m, what do we do with them? The most common thing is to compute averages (expectation values) of observables. Let A be an arbitrary quantum mechanical observable. If the system is in $|m\rangle$, then the average of A will be $\langle m|A|m\rangle$. But since the system can be found in each $|m\rangle$ with probability P_m, the average of A will be given by the weighted sum

$$\boxed{\langle A\rangle = \sum_m \langle m|A|m\rangle P_m} \tag{14}$$

This is an average of averages: we take the quantum mechanical average, $\langle m|A|m\rangle$, and average it over the thermal distribution P_m.

The most important observable is the Hamiltonian itself. But since $|m\rangle$ are eigenvectors of H, it follows in this case that

$$\langle m|H|m\rangle = E_m \tag{15}$$

The average energy thus simplifies to

$$\boxed{U = \langle H\rangle = \sum_m E_m P_m} \tag{16}$$

For historical reasons the average energy receives the symbol U and is usually called the **internal energy**.

Example: qubit

In the case of a qubit

$$\boxed{U = \epsilon P_e = \frac{\epsilon}{e^{\beta\epsilon}+1} \qquad \text{(qubit)}} \qquad (17)$$

The plot of U vs. T/ϵ will therefore look exactly like the one of P_e.

In the case of spin 1/2 we have $H = -\frac{h}{2}\sigma_z$. Instead of computing U, let us first compute $\langle\sigma_z\rangle$. Using Eq (14) we get

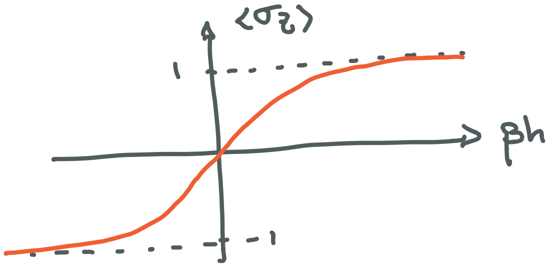
$$\langle\sigma_z\rangle = \langle\uparrow|\sigma_z|\uparrow\rangle P_\uparrow + \langle\downarrow|\sigma_z|\downarrow\rangle P_\downarrow$$

$$= (+1)\frac{e^{\beta h/2}}{2\cosh(\beta h/2)} + (-1)\frac{e^{-\beta h/2}}{2\cosh(\beta h/2)}$$

which simplifies to

$$\boxed{\langle\sigma_z\rangle = \tanh\left(\frac{\beta h}{2}\right)} \qquad (18)$$

This is a famous result: the paramagnetic response of a spin 1/2 system to an external field. It looks like this



when $\beta h \ll 1$ we may expand

$$\tanh(x) \simeq x \tag{19}$$

which leads to

$$\langle \sigma_z \rangle \simeq \frac{h}{2T} \tag{20}$$

this result is known as Curie's law: the response of a paramagnetic system to a weak magnetic field is linear in the field, with a coefficient proportional to $1/T$. This law can indeed be verified experimentally for a large number of systems.

Oh, and of course, I forgot to say, but once we have $\langle \sigma_z \rangle$ it's trivial to compute U since $H = -\frac{h}{2}\sigma_z$:

$$U = \langle H \rangle = -\frac{h}{2}\tanh\left(\frac{\beta h}{2}\right) \tag{21}$$

Example: QHO

The internal energy for the QHO reads

$$U = \sum_{n=0}^{\infty} \hbar\omega(n + 1/2)\, P_n = \hbar\omega \left(\sum_{n=0}^{\infty} n P_n + 1/2 \right)$$

where I used the fact that $\sum_{n=0}^{\infty} P_n = 1$ in the last term. The first term can be identified as the mean number of excitations

$$\langle n \rangle = \sum_{n=0}^{\infty} n P_n = (1 - e^{-\beta\hbar\omega}) \sum_{n=0}^{\infty} n\, e^{-\beta\hbar\omega n} \qquad (22)$$

This is a kind of modified geometric sum, so we can use the following result (see appendix)

$$\sum_{n=0}^{\infty} n x^n = \frac{x}{(1-x)^2}$$

Thus

$$\langle n \rangle = (1 - e^{-\beta\hbar\omega}) \frac{e^{-\beta\hbar\omega}}{(1 - e^{-\beta\hbar\omega})^2}$$

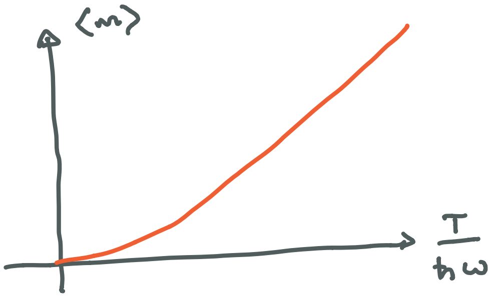
$$= \frac{e^{-\beta\hbar\omega}}{1 - e^{-\beta\hbar\omega}}$$

or

$$\boxed{\langle n \rangle = \frac{1}{e^{\beta\hbar\omega} - 1}} \qquad (23)$$

We will also encounter the same formula later on under the name of **Bose-Einstein distribution**.

Eq (19) looks like



$\langle n \rangle = 0$ when $T = 0$

$\langle n \rangle \sim$ linear when T is high

From (19) the average energy becomes

$$U = \hbar\omega\left(\langle n \rangle + 1/2\right) = \hbar\omega\left(\frac{1}{e^{\beta\hbar\omega}-1} + \frac{1}{2}\right)$$

which can also be written as

$$\boxed{U = \frac{\hbar\omega}{2}\coth\left(\frac{\hbar\omega}{2T}\right)} \tag{20}$$

For high temperatures, $\beta\hbar\omega \ll 1$, we can expand

$$\coth(x) \simeq \frac{1}{x} \tag{21}$$

yielding

$$U \simeq T \tag{22}$$

This is what one gets for a classical harmonic oscillator. It is common that in systems with infinite dimensional Hilbert space one recovers classical results as a particular case.

Heat Capacity

The internal energy U can actually be computed directly from the partition function Z, without having to do another sum. Start with

$$Z = \sum_m e^{-\beta E_m}$$

and differentiate with respect to β:

$$\frac{\partial Z}{\partial \beta} = -\sum_m E_m e^{-\beta E_m}$$

Dividing by Z we see that

$$\frac{1}{Z}\frac{\partial Z}{\partial \beta} = -\sum_m E_m \frac{e^{-\beta E_m}}{Z} = -\sum E_m P_m = -U$$

Thus

$$\boxed{U = -\frac{1}{Z}\frac{\partial Z}{\partial \beta} = -\frac{\partial}{\partial \beta}\ln Z} \tag{23}$$

This formula is very useful: when working with hard problems, computing Z can be a very difficult task. This formula shows that once we have Z, U comes for free. It also shows that Z is more than just a "normalization constant". In fact, there is an enormous amount of information contained in Z.

An important quantity, related to U, is the heat capacity, defined as

$$\boxed{C = \frac{\partial U}{\partial T}} \tag{24}$$

It measures how sensitive is the energy U to changes in temperature. The name "heat capacity" is thus quite fitting: it is the capacity of the system to store thermal energy.

I like to read Eq (24) as

$$\Delta U = C \Delta T \tag{24'}$$

In words: if you change the temperature by ΔT, how much thermal energy can the system absorb.

Since we work with $k_B = 1$, T has units of energy, so C is dimensionless. In general, the units of C are the units of k_B.

In systems with many particles, it is natural to talk about the heat capacity per particle, or per gram, or per cm^3, or per mole. When this is the case we use the term specific heat. Thus, the molar specific heat is the heat capacity per mole, and so on. Heat capacity and specific heat are thus quite similar, except that the latter is normalized in some specific way.

Examples: qubit and QHO

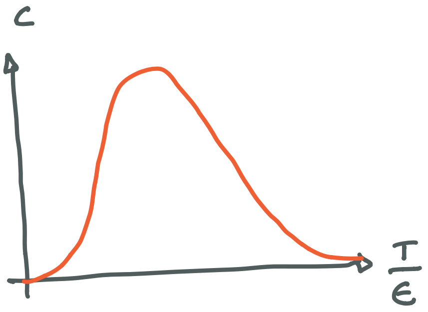
For a qubit we get, from (17),

$$C = \frac{\beta\epsilon\, e^{\beta\epsilon}}{(e^{\beta\epsilon}+1)^2} \qquad \text{(qubit)} \qquad (25)$$

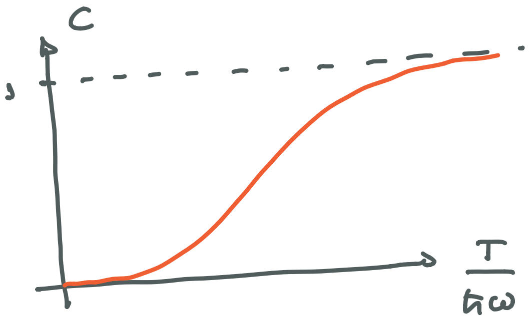
whereas for the QHO, from (20),

$$C = \frac{\beta\hbar\omega\, e^{\beta\hbar\omega}}{(e^{\beta\hbar\omega}-1)^2} \qquad \text{(QHO)} \qquad (26)$$

These two quantities look like this:



(qubit)



(QHO)

The physics behind these results is very nice. Remember that the heat capacity measures how much energy ΔU the system can absorb if we change the temperature by ΔT.

As can be seen, as the temperature increases c also goes up in both cases. For the qubit, however, c eventually reaches a maximum and then start to fall. For the QHO, on the other hand, c keeps on growing and eventually saturate.

The maximum in the qubit case is called Schottky effect and is related to the fact that, if the temperature starts to increase too much, there are no more states left to put more energy, since the dimension of the Hilbert space is finite. In fact, any system with finite Hilbert space will eventually present a maximum for some T.

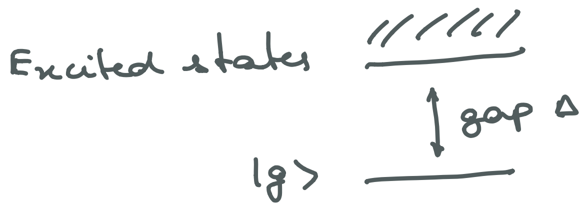
In the case of the QHO, on the other hand, the number of states is infinite so there is always more room to put more energy. But if the temperature is too high, the sensitivity stops depending on T, so c tends to a constant.

Ex: superconductivity

At low temperatures, $\beta\hbar\omega << 1$, we get from (25) and (26)

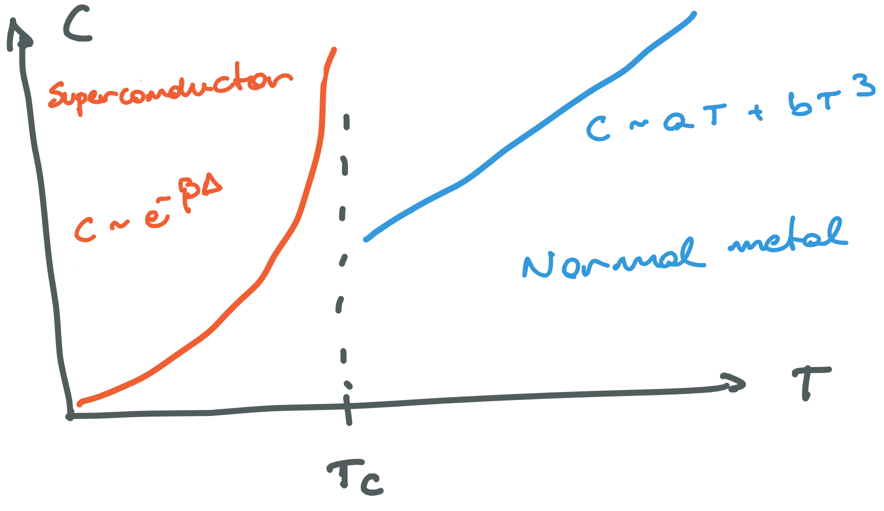
$$c \simeq \beta\epsilon\, e^{-\beta\epsilon} \qquad \text{and} \qquad c \simeq \beta\hbar\omega\, e^{-\beta\hbar\omega} \tag{27}$$

In both cases the heat capacity vanishes exponentially with the energy gap between the ground state and the first excited state



$$c \sim e^{-\beta\Delta} \tag{28}$$

This is actually a very general feature. A really cool example of this is superconductivity. The specific heat of a superconducting material (like Nb) looks somewhat like this:



(specific heat of a type-I superconductor)

Above the critical temperature T_c, a superconductor behaves like a normal metal. For metals the specific heat usually behaves as a power law in T, usually like $C = aT + bT^3$ (where a and b are constants). We will see where this comes from later on.

What is peculiar about superconductors is that below T_c the behavior of $C(T)$ is exponential (which you can see by doing a linear-log plot).

From this simple experimental observation, people already knew that in the superconducting phase the system had to open a gap. In metals the low energy states form a quasi-continuum (vanishing gap). But in the SC phase a gap must open up.

When the BCS theory was formulated many years later, it was in fact seen that this was correct: the key feature of a superconductor is the opening of a gap.

I find it absolutely incredible that such a seemingly bulky measurement (that reminds us of chemistry class) can yield such an important piece of information.

Heat capacity and fluctuations of energy

Let us combine (23) and (24):

$$C = \frac{\partial U}{\partial T} = \frac{\partial U}{\partial \beta}\frac{\partial \beta}{\partial T} = -\frac{1}{T^2}\frac{\partial U}{\partial \beta} \tag{29}$$

Then

$$\frac{\partial U}{\partial \beta} = -\frac{\partial}{\partial \beta}\left(\frac{1}{Z}\frac{\partial Z}{\partial \beta}\right) = -\frac{1}{Z}\frac{\partial^2 Z}{\partial \beta^2} + \left(\frac{1}{Z}\frac{\partial Z}{\partial \beta}\right)^2$$

The last term is just U^2, or $\langle H\rangle^2$. As for the first term, we have

$$\frac{1}{Z}\frac{\partial^2 Z}{\partial \beta^2} = \frac{1}{Z}\frac{\partial^2}{\partial \beta^2}\sum_n e^{-\beta E_n} = \frac{1}{Z}\sum_n (-E_n)^2 e^{-\beta E_n}$$

with a little thought, we recognize this as the 2nd moment of the Hamiltonian [c.f. Eq. (14)]:

$$\langle H^2\rangle = \sum_n E_n^2 P_n = \frac{1}{Z}\frac{\partial^2 Z}{\partial \beta^2} \tag{30}$$

Thus, plugging this in (29), we arrive at

$$\boxed{C = \frac{1}{T^2}\left[\langle H^2\rangle - \langle H\rangle^2\right]} \tag{31}$$

This result is very important, as it gives an entirely new interpretation to the heat capacity, as being related to the variance of H, which measures the fluctuations of the system's energy.

The variance is always non-negative, as it may be written as

$$\langle H^2\rangle - \langle H\rangle^2 = \langle (H - \langle H\rangle)^2\rangle \geq 0 \qquad (32)$$

Thus, Eq (31) also implies that the specific heat is always non-negative

$$\boxed{C = \frac{\partial U}{\partial T} \geq 0} \qquad (33)$$

In words, the energy is a monotonically increasing function of T.

Why $e^{-\beta E}$?

Now that we are getting more familiar with the Gibbs formula (2), we can start to ask where it comes from or why it works. Here I want to discuss (in a non-rigorous way) why $P_n \propto e^{-\beta E_n}$; i.e. why an exponential.

To see why this is so, we only need to postulate one (very reasonable) thing:

Postulate of equal probabilities a priori: the prob. of occupation of a given state $|n\rangle$ only depend on the energy E_n of that state.

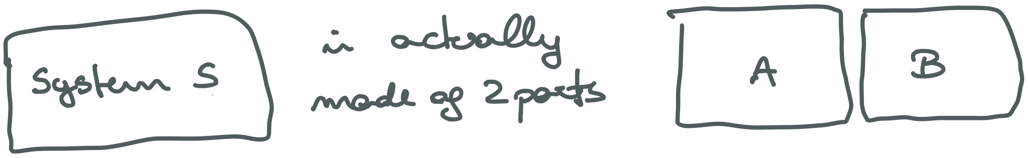
This postulate puts energy on a pedestal. It says that, in equilibrium, all that matters is the energy. In the end of the day, this is a postulate and there is no way of deriving it from some more fundamental principle. But at least in my head, it makes sense.

The fact that $P_n \propto e^{-\beta E_n}$ follows from the postulate. All we need to realize is that

In equilibrium, if energy is a sum, the probabilities should be a product.

Let me explain what this means

The above postulate implies that $P_n = f(E_n)$ for some function f. What we want to know is precisely how the function f looks like. To do that suppose our system is actually composed of two parts, A and B, which do not interact with each other:



Then since energy is an additive quantity, the total energy will be given by

$$E^{AB}_{n,m} = E^A_n + E^B_m \tag{34}$$

where n, m are quantum numbers for A and B respectively.

If our AB system is in equilibrium, then the corresponding probabilities will be given by

$$P^{AB}_{n,m} = f(E^{AB}_{n,m}) = f(E^A_n + E^B_m) \tag{35}$$

However, A and B do not interact (energy is a sum) so that their probabilities should be statistically independent (probabilities should be a product):

$$P^{AB}_{nm} = P^A_n P^B_m \tag{36}$$

But since A and B are in equilibrium, $P^A_n = f(E^A_n)$ and $P^B_m = f(E^B_m)$. Thus we reach the conclusion that the function f must satisfy

$$f(E^A_n + E^B_m) = f(E^A_n)\, f(E^B_m) \tag{37}$$

which kind of function satisfies $f(x+y) = f(x)f(y)$? Answer: the exponential,

$$f(x) = c e^{\alpha x} \tag{38}$$

where c and α are constants. Any other function can be reparametrized in this way by redefining c and α (try it!).

Thus, we conclude from this that equilibrium probabilities must have the form

$$P_n = c e^{\alpha E_n} \tag{39}$$

What can now be said about c and α? Well, c is easy: since the probabilities are normalized

$$1 = \sum_n P_n = c \sum_n e^{\alpha E_n} \rightsquigarrow c = \frac{1}{\sum_n e^{\alpha E_n}} \tag{40}$$

This then leads us to the partition function $Z := 1/c$.

But what about α? First of all, we can prove it has to be negative. The reason is simple. Suppose we order the energy levels of a system in ascending order

$$E_0 \geqslant E_1 \geqslant E_2 \geqslant E_3 \geqslant \dots \tag{41}$$

The ratio of two probabilities is

$$\frac{P_n}{P_m} = e^{\alpha (E_n - E_m)} \tag{42}$$

Suppose that $E_n > E_m$. Then if $\alpha > 0$, $e^{\alpha(E_n - E_m)} > 1$ and thus $P_n > P_m$. If α were positive, higher energy states would always be more likely. An ionized electron would always be more probable than a bound electron. Matter would be unstable and we would all be dead. Since we are not dead (I think, #sanjunipero), we conclude that $\alpha < 0$. We therefore redefine it as

$$\alpha := -\beta, \quad \beta > 0. \tag{43}$$

The probabilities now acquire the form

$$P_n = \frac{e^{-\beta E_n}}{Z}. \tag{44}$$

Now, besides being positive, what else can we say about β? Can we relate it to temperature somehow? Well, we have two pieces of evidence we can use:

- β is the only parameter which characterizes equilibrium (i.e., which enters into P_n).
- U is monotonically decreasing in β [because of (33)].

It thus makes some sense to say that β should be proportional to the inverse of the temperature.

But is it $\beta \propto 1/T$ or $\beta \propto 1/T^2$ or some other weird function? This is where human input becomes necessary. The reason why $\beta = 1/T$ is because we chose to define temperature using a certain scale. We could have chosen some other scale, in which case the relation $\beta(T)$ would be completely different. For instance changing $T \to T^2$ would still give a valid scale of temperature.

I find it very interesting that the postulate of equal a priori probabilities can say so much about the physical structure of the equilibrium state.

Summary of useful formulas:

Gibbs ensemble:

$$p_n = \frac{e^{-\beta E_n}}{Z}$$

$$Z = \sum_n e^{-\beta E_n}$$

= sum over states

$p_n > p_m$ if $E_n < E_m$

Expectation value of observables:

$$\langle A \rangle = \sum_n \langle n|A|n\rangle p_n$$

$$U = \langle H \rangle = \sum_n E_n p_n = -\frac{\partial}{\partial \beta} \ln Z$$

$$C = \frac{\partial U}{\partial T} = \frac{1}{T^2}\left[\langle H^2\rangle - \langle H\rangle^2\right] \geq 0$$

U always grows with T.

Examples:

Qubit

$$H = \epsilon |e\rangle\langle e|$$

$$E_g = 0, \; E_e = \epsilon$$

$$Z = 1 + e^{-\beta\epsilon}$$

$$p_e = \frac{1}{e^{\beta\epsilon} + 1}$$

= Fermi-Dirac

$$U = \epsilon p_e$$

$$C = \beta\epsilon \frac{e^{\beta\epsilon}}{(e^{\beta\epsilon}+1)^2}$$

Spin 1/2

$$H = -\frac{h}{2}\sigma_z$$

$$E_\uparrow = -h/2, \; E_\downarrow = h/2$$

$$Z = 2\cosh(\beta h/2)$$

$$\langle \sigma_z \rangle = \tanh\left(\frac{\beta h}{2}\right)$$

QHO

$$H = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 x^2$$

$$E_n = \hbar\omega(n + 1/2)$$

$$Z = \frac{e^{-\beta\hbar\omega}}{1 - e^{-\beta\hbar\omega}}$$

$$\langle n \rangle = \frac{1}{e^{\beta\hbar\omega} - 1}$$

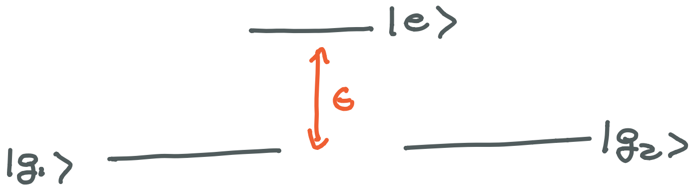
= Bose - Einstein

$$U = \frac{\hbar\omega}{2} \coth\left(\frac{\hbar\omega}{2T}\right)$$

$$C = \beta\hbar\omega \frac{e^{\beta\hbar\omega}}{(e^{\beta\hbar\omega} - 1)^2}$$

Example: a Λ atom

Consider a system with 3 levels $|g_1\rangle$, $|g_2\rangle$ and $|e\rangle$, displaced according to



$$H = \epsilon |e\rangle\langle e| = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \epsilon \end{pmatrix} \qquad (45)$$

In this case the ground-state is doubly degenerate. This is important because it means there are more configurations available for the system. This is visible in the partition function. It is useful to remember that the letter Z is used due to the German word Zustandssumme, which means a sum over states. It is not a sum over energies, but over the available states. Thus, in this example

$$Z = \sum_n e^{-\beta E_n} = 1 + 1 + e^{-\beta\epsilon} = 2 + e^{-\beta\epsilon} \qquad (46)$$

This is different from the previous example due to the factor of 2. We then have that

$$P_{g_1} = P_{g_2} = \frac{1}{Z} \qquad (47)$$

If you want to ask about the prob. of finding the system with energy 0, then it's a different question and we have to add P_{g_1} and P_{g_2}

$$P(\text{ground-state}) = P_{g_1} + P_{g_2} = \frac{2}{Z} \qquad (48)$$

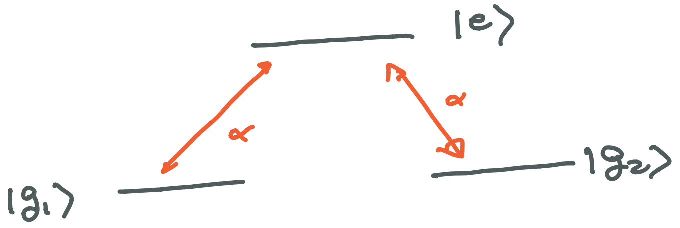
Example: a driven Λ-atom

Λ systems, like the one in the previous example, appear often quantum optics and atomic physics. When such a system is also subject to an external laser drive, the Hamiltonian is modified to

$$H = \epsilon |e\rangle\langle e| + \alpha\left(|e\rangle\langle g_1| + |g_1\rangle\langle e|\right) + \alpha\left(|e\rangle\langle g_2| + |g_2\rangle\langle e|\right)$$

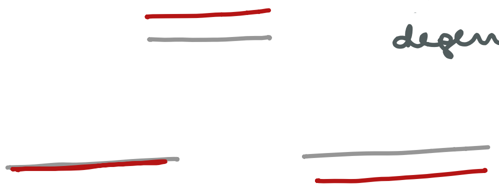
$$= \begin{pmatrix} 0 & 0 & \alpha \\ 0 & 0 & \alpha \\ \alpha & \alpha & \epsilon \end{pmatrix} \tag{49}$$

where $\alpha \in \mathbb{R}$ represents the strength of the drive. A term like $|e\rangle\langle g_1|$ represents a transition amplitude from $|g_1\rangle$ to $|e\rangle$. Thus, this Hamiltonian diagrammatically looks like



Note that since Hamiltonians are Hermitian, if we can jump from $|g_1\rangle \to |e\rangle$, then we must also be able to jump from $|e\rangle \to |g_1\rangle$, with the same probability.

The Hamiltonian (49) is no longer diagonal. To apply the Gibbs formula we thus first need to find its eigenvalues. They read (you can use Mathematica to check)



α breaks the degeneracy.

$$E_0 = 0 \qquad (50)$$

$$E_\pm = \frac{\epsilon}{2} \pm \frac{1}{2}\sqrt{\epsilon^2 + 8\alpha^2}$$

The partition function is thus

$$Z = 1 + 2e^{-\beta\epsilon/2}\cosh\left(\frac{\beta\Omega}{2}\right) \qquad (51)$$

$$\Omega = \sqrt{\epsilon^2 + 8\alpha^2}$$

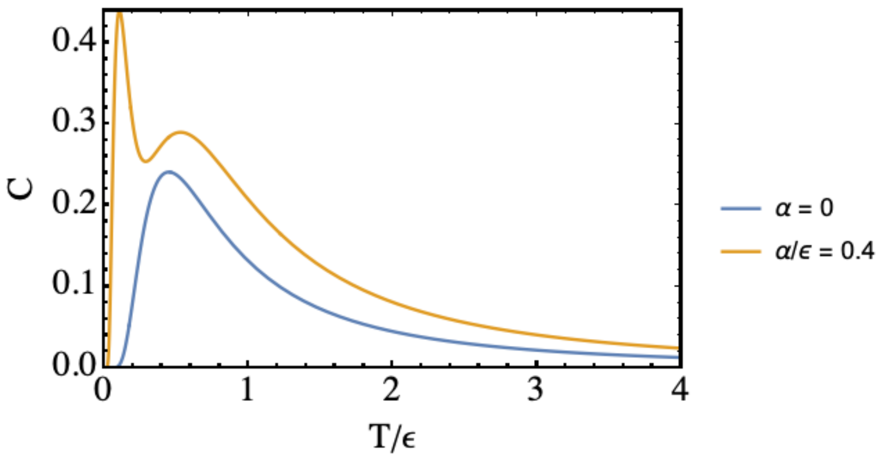
From this we get all thermodynamic quantities:

$$U = -\frac{\partial}{\partial\beta}\ln Z = \frac{\epsilon\cosh(\beta\Omega/2) - \Omega\sinh(\beta\Omega/2)}{e^{\beta\epsilon/2} + 2\cosh(\beta\Omega/2)}$$

$$C = \frac{\partial U}{\partial T} = \frac{e^{\beta\epsilon/2}\beta^2(\Omega^2+\epsilon^2)\cosh\beta\Omega/2 + 2\beta^2\Omega\left[\Omega - \epsilon e^{\beta\epsilon/2}\sinh(\beta\Omega/2)\right]}{2\left(e^{\beta\epsilon/2} + 2\cosh(\beta\Omega/2)\right)^2}$$

The formulas are ugly, I know. But that doesn't matter. They are trivial to compute once we have Z and they are also easy to make plots with.

Here is the specific heat for two values of α.



when $\alpha = 0$ we get a schottky peak but when $\alpha \neq 0$ we can get either 1 or 2 peaks. The reason for this is due to the fact that when $\alpha \neq 0$ the degeneracy is lifted, so that we now have 3 distinct energy levels. the peaks appear depending on the spacing between these levels.

Once we have Z in Eq (51), we can also compute the Gibbs probabilities

$$P_m = \frac{e^{-\beta E_m}}{Z}$$

But what does "m" mean here? It always means the energy eigenstates, which in this case are not the original states $|g_1\rangle$, $|g_2\rangle$ and $|e\rangle$ in Eq (49).

How can we then compute the probability P_{g_1}, P_{g_2} and P_e of finding the atom in these states? Answer: we compute the expectation value of a projection operator.

Recall Eq (14):

$$\langle A \rangle = \sum_m \langle m|A|m\rangle P_m .$$

this holds for any observable A. If we choose $A = |\psi\rangle\langle\psi|$, where $|\psi\rangle$ is an arbitrary state, we then get

$$\boxed{\langle |\psi\rangle\langle\psi| \rangle = \sum_m |\langle\psi|m\rangle|^2 P_m} \qquad (52)$$

this formula makes sense: $|\langle\psi|m\rangle|^2$ is simply the Born rule. It is the prob. that you find the system in $|\psi\rangle$ given that it is in $|m\rangle$. Thus Eq (52) is again an average of averages: you weight each $|\langle\psi|m\rangle|^2$ by P_m.

In order to apply this to our example in (49), we also need the eigenvectors of H. They can be written as (the ordering of the vectors is as in (49), $|g_1\rangle, |g_2\rangle, |e\rangle$).

$$|E_0\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \tag{53a}$$

$$|E_+\rangle = \frac{1}{\sqrt{2\Omega(\Omega+\epsilon)}}\begin{pmatrix} 2\alpha \\ 2\alpha \\ \Omega+\epsilon \end{pmatrix} \tag{53b}$$

$$|E_-\rangle = \frac{1}{\sqrt{2\Omega(\Omega-\epsilon)}}\begin{pmatrix} -2\alpha \\ -2\alpha \\ \Omega-\epsilon \end{pmatrix} \tag{53c}$$

It is interesting to note how the eigenstate with energy $E_0 = 0$ is just a superposition of the two original eigenstates $|g_1\rangle$ and $|g_2\rangle$, which already had energy 0. This combination is thus unaffected by the external drive. For this reason, the state (53a) is called a dark state.

Let us now apply (52). For instance, the prob. of finding the atom in the state $|e\rangle$ is

$$P_e = \underbrace{|\langle e|E_0\rangle|^2}_{0} P_0 + |\langle e|E_+\rangle|^2 P_+ + |\langle e|E_-\rangle|^2 P_-$$

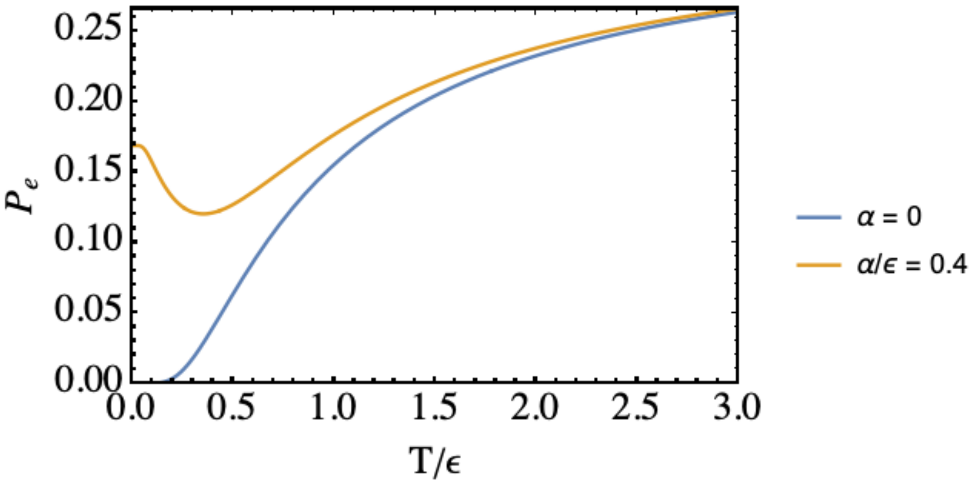
$$= \left(\frac{\Omega+\epsilon}{\sqrt{2\Omega(\Omega+\epsilon)}}\right)^2 \frac{e^{-\beta E_+}}{Z} + \left(\frac{\Omega-\epsilon}{\sqrt{2\Omega(\Omega-\epsilon)}}\right)^2 \frac{e^{-\beta E_-}}{Z}$$

$$= \frac{\Omega+\epsilon}{2\Omega} \frac{e^{-\beta\epsilon/2} e^{-\beta\Omega/2}}{1 + e^{-\beta\epsilon/2}(e^{\beta\Omega/2} + e^{-\beta\Omega/2})} + \frac{\Omega-\epsilon}{2\Omega} \frac{e^{-\beta\epsilon/2} e^{+\beta\Omega/2}}{1 + e^{-\beta\epsilon/2}(e^{\beta\Omega/2} + e^{-\beta\Omega/2})}$$

$$= \frac{\Omega+\epsilon}{2\Omega} \frac{1}{e^{\beta(\Omega+\epsilon)/2} + e^{\beta\Omega} + 1} + \frac{\Omega-\epsilon}{2\Omega} \frac{e^{\beta\Omega}}{e^{\beta(\Omega+\epsilon/2)} + e^{\beta\Omega} + 1}$$

$$\therefore \quad P_e = \frac{1}{2\Omega} \frac{(\Omega+\epsilon) + (\Omega-\epsilon)e^{\beta\Omega}}{1 + e^{\beta\Omega} + e^{\beta(\Omega+\epsilon)/2}}$$

Here is how this looks like:



when $\alpha = 0$, increasing T increases the population of $|e\rangle$ monotonically. And if $T=0$, $P_e = 0$. But for $\alpha \neq 0$ we get a finite P_e even for $T=0$. This is the effect of the drive, which makes the ground state (53b) have a component in the $|e\rangle$ direction. As T increases, the populations redistribute and it may even happen that P_e goes down with T.

We also see from the figure that as $T\to\infty$ P_e always tends to $1/3$, for any value of α. We already know that this is the case for the P_m: $T=\infty$ everyone is equally populated. But it turns out it is also true for any state. This can be shown from (52):

$$\langle|\psi\rangle\langle\psi|\rangle = \sum_m |\langle\psi|m\rangle|^2 \frac{1}{3}$$

$$= \frac{1}{3}\sum_m \langle\psi|m\rangle\langle m|\psi\rangle$$

Completeness Relation

$$\sum_m |m\rangle\langle m| = 1$$

$$= \frac{1}{3}\langle\psi|\psi\rangle$$

$$= \frac{1}{3}$$

for any $|\psi\rangle$. Infinite temperature is very democratic: all states are equally likely.

Appendix: geometric and related series

The geometric series reads

$$\sum_{m=0}^{\infty} x^m = \frac{1}{1-x} \qquad (A.1)$$

To derive this, the easiest way is to simply do a Taylor expansion. You may convince yourself that

$$\frac{d^{(m)}}{dx^m} \frac{1}{1-x} \bigg|_{x=0} = m!$$

Thus, from Taylor's theorem

$$\frac{1}{1-x} = \sum_{m=0}^{\infty} \frac{m!}{m!} x^m = \sum_{m=0}^{\infty} x^m$$

which is (A.1).

Other sums can also be computed from (A.1). For instance, to compute $\sum_{m=0}^{\infty} m x^m$, differentiate both sides of (A.1) with respect to x to get

$$\sum_{m=0}^{\infty} m x^{m-1} = \frac{1}{(1-x)^2}$$

Multiplying by x on both sides yields the sum we wanted

$$\sum_{m=0}^{\infty} m x^m = \frac{x}{(1-x)^2} \qquad (A.2)$$

Another potentially useful series is the finite version of (A.1):

$$\sum_{n=0}^{L} x^n$$

for some integer L. this sum is super easy to compute because we can write it as

$$\begin{aligned}\sum_{n=0}^{L} x^n &= 1 + x + \dots + x^L \\ &= 1 + x + \dots + x^L + x^{L+1} + \dots \quad \leftarrow \text{Geometric series} \\ &\quad - x^{L+1}\left(1 + x + x^2 + \dots\right) \quad \leftarrow \\ &= \left(\sum_{n=0}^{\infty} x^n\right) - x^{L+1}\left(\sum_{n=0}^{\infty} x^n\right)\end{aligned}$$

Using (A.1) we then get

$$\sum_{n=0}^{L} x^n = \frac{1 - x^{L+1}}{1 - x} \tag{A.3}$$