The Gibbs state

In most of this course, we will be interested in systems which are in thermal equilibrium. Consider an arbitrary physical system, described by a Hamiltonian H with eigenstuff

$$H|m\rangle = E_m |m\rangle \qquad (1)$$

this can be anything, from a single electron to 10^{23} particles. If this system is in thermal equilibrium, then the probability of finding it in state $|m\rangle$ will be given by

$$P_m = \frac{e^{-\beta E_m}}{z} \qquad (2)$$

which is called the Gibbs state. Here

$$\beta = \frac{1}{k_B T} = \frac{1}{T} \qquad (k_B = 1) \qquad (3)$$

where T is the temperature and k_B is Boltzmann's constant, which I set to $k_B = 1$. Moreover, the quantity z in Eq (2) is a normalization constant called the partition function

$$z = \sum_m e^{-\beta E_m} \qquad (4)$$

Eq (2) is the foundation for all of equilibrium statistical mechanics. So now there are two things left for us to do. We can either try to understand why Eq (2) is true (maybe even try to derive it) or we can simply apply it to a bunch of problems and see what comes out. We shall do both.

!

Once we have the probabilities (2), the next big thing is to compute the expectation value of observables. Given any observable A, its expectation value in the Gibbs state will be given by

$$\langle A \rangle = \sum_m \langle m | A | m \rangle \, P_m \qquad (5)$$

You can see that is a type of double expectation value: it is the expectation value $\langle m | A | m \rangle$ in each energy eigenstate, weighted by the Gibbs probabilities P_m.

The most important observable is the Hamiltonian. The expectation value $\langle H \rangle$ receives a special name, the _internal energy_, and a special symbol U. Since $\langle m | H | m \rangle = E_m$, we thus get

$$U = \langle H \rangle = \sum_m E_m \, P_m \qquad (6)$$
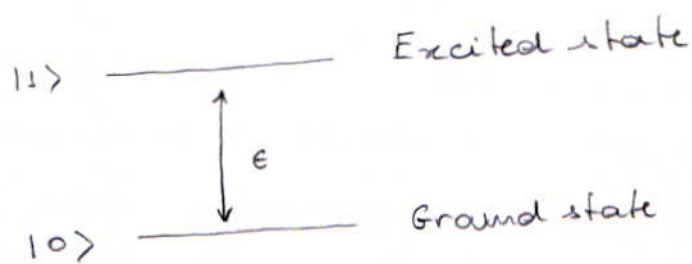
Example: qubit (2-level system)

As our first example, we consider a system that has only two energy levels, which we label as $|0\rangle$ and $|1\rangle$. That is,

$$H|0\rangle = E_0 |0\rangle$$
$$H|1\rangle = E_1 |1\rangle$$

(7)

Since energy is defined only up to a constant, we may, without loss of generality, set $E_0 = 0$ and $E_1 = \epsilon$. Then $|0\rangle$ is the ground state, $|1\rangle$ is the excited state and ϵ is the energy gap between them

In this case the partition function (4) becomes

$$\boxed{z = e^{-\beta E_0} + e^{-\beta E_1} = 1 + e^{-\beta \epsilon}}$$

(8)

whereas the Gibbs probabilities become

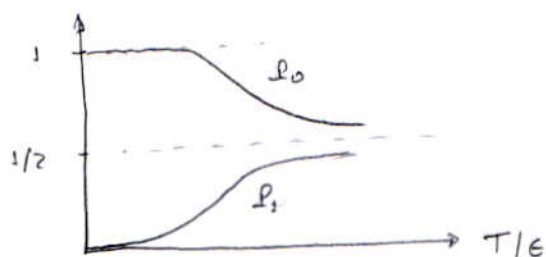$$\boxed{\begin{aligned} P_0 &= \frac{1}{1 + e^{-\beta \epsilon}} \\[2mm] P_1 &= \frac{e^{-\beta \epsilon}}{1 + e^{-\beta \epsilon}} = \frac{1}{e^{\beta \epsilon} + 1} \end{aligned}}$$

(9)

(10)

The probabilities only depend on the quantity

$$\beta \epsilon = \frac{\epsilon}{T}$$

(1...)

this makes intuitive sense: the probabilities of occupation depend only on the ratio of the energy gap to the thermal energy. In terms of T/ϵ the probabilities P_0 and P_1 look like this

The probability of being found in the ground state (GS) is always larger. As $T \to 0$ we get $P_0 = 1$ and $P_1 = 0$. So at zero temperature the system "settles down" in $|0\rangle$. Conversely, if $T \to \infty$, then both P_0 and P_1 tend to $1/2$. In this case you are equally likely to find the system in either $|0\rangle$ or $|1\rangle$.

The prob. P_1 in Eq (10) represents the prob. of finding the system in the excited states. For reasons that will become clear later on, it is called the Fermi-Dirac distribution. The internal energy, Eq (6), becomes in this case

$$\boxed{U = E_0 \, P_0 + E_1 \, P_1 = \epsilon \, P_1}$$

(11)

Thus, the plot of P_1 in the figure above is also a plot of U. We therefore see that energy increases monotonically with temperature, which makes sense

Example: the harmonic oscillator

As our second example, we consider the harmonic oscillator, with energy eigenvalues

$$E_m = \hbar\omega\left(m + 1/2\right) \qquad m = 0, 1, 2, \dots \qquad (12)$$

In this course I also set $\hbar = 1$ so, neglecting the irrelevant $\hbar\omega/2$ as well, we can simply write

$$E_m = \omega m \qquad m = 0, 1, 2, \dots \qquad (13)$$

The partition function becomes

$$Z = \sum_{m=0}^{\infty} e^{-\beta\omega m}$$

If we define $x = e^{-\beta\omega}$ then we identify here a geometric series

$$\sum_{m=0}^{\infty} x^m = \frac{1}{1-x} \qquad (14)$$

Moreover, in our case the series is always convergent since $e^{-\beta\omega} < 1$ (because both β and ω must be positive). Hence
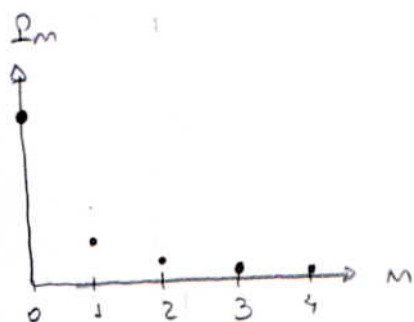
$$\boxed{Z = \frac{1}{1 - e^{-\beta\omega}}} \qquad (15)$$
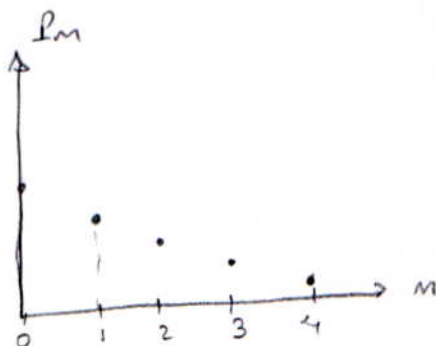
The Gibbs probabilities P_m in (2) then became

$$P_m = (1 - e^{-\beta \omega}) \, e^{-\beta \omega m} \qquad (16)$$

In probability theory this is known as the geometric distribution.

They look something like this:

$$\frac{T}{\omega} \text{ small} \qquad\qquad \frac{T}{\omega} \text{ large}$$

It always goes down with m: lower energy states are always more likely.
For T/ω large it decreases slowly, whereas for T/ω small it is highly
concentrated around $|0\rangle$ and goes down fast with $|m\rangle$. Indeed,
from (16) we see that the probability of finding the system in
the GS is $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (17)$

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

which looks like this

Thus, as $T \to 0$ we get $P_0 = 1$: the system is found with certainty in the
GS. For $T \to \infty$, on the other hand, $P_0 \to 0$ since in this case the probabilities
become diluted over many states.

The average energy (6) in this case becomes

$$U = \sum_m \omega m \, P_m \tag{18}$$

It is convenient to get rid of the ω in front and simply speak of the average occupation number

$$\langle m \rangle = \sum_{m=0}^{\infty} m \, P_m = (1 - e^{-\beta \omega}) \sum_{m=0}^{\infty} m \, e^{-\beta \omega m} \tag{19}$$

This sum can be computed using the following neat trick. Differentiate the geometric series (14) with respect to x:

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

Now multiply by x on both sides to get

$$\sum_{m=0}^{\infty} m \, x^m = \frac{x}{(1-x)^2} \tag{20}$$
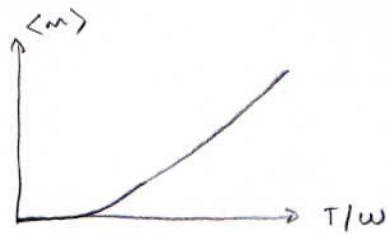
Thus (19) becomes

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

Or, what is more convenient,

$$\boxed{\langle m \rangle = \frac{1}{e^{\beta \omega} - 1}} \tag{21}$$

Again for reasons that will become clear later on, this is called the Bose-Einstein distribution. It looks similar to (10), except for the minus sign.

The function $\langle m \rangle$ looks like this

For $T \to 0$ we get $\langle m \rangle = 0$ (no occupation). then it curves up and for large T it becomes linear.

It is also sometimes convenient to note that we can write (21) as

$$\boxed{\langle m \rangle + 1/2 = \frac{1}{2} \coth\left(\frac{\beta \omega}{2}\right)}$$
(22)

For small x (23)

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

so we get, for large T/ω, (24)

$$\langle m \rangle + 1/2 \simeq \frac{T}{\omega}$$

If we resubstitute the $1/2$ that I threw away in the energy (12), we then get (25)

$$\boxed{U = \frac{\omega}{2} \coth\left(\frac{\beta \omega}{2}\right)}$$

For large T we get $U \simeq T$.

why the Gibbs state?

Now I want to argue with you as to why the Gibbs formula (2) makes sense. I will not try to be rigorous, but simply appeal to our intuition.

I will start by making one fundamental postulate, which is present in all theoretical formulations of statistical mechanics.

<div style="border:1px solid black; padding:10px;">

<u>Postulate</u> of <u>equal</u> <u>a priori</u> <u>probabilities</u>: quantum states with the same energy are equally likely
</div>

(26)

This puts energy on a pedestal: it says that as a consequence of the complicated microscopic interactions in nature, when a system is in equilibrium, the energy is the only things which can tell states apart.

This postulate means that the probabilities P_m of a given state $|m\rangle$ must be a function only of its energy E_m:

$$P_m = f(E_m)$$

(27)

where $f(x)$ is some unknown function. Now I want to convince you that this function must be an exponential.

To do that, let's imagine just for now that we actually have two system, A and B, both in equilibrium. And we assume that they do not interact at all so the total energy will be

$$E_m^A + E_m^B$$

(9)

But if they don't interact, they must be statistically independent so the probabilities must obey

$$P_{m,mn} = P_m \, P_{mn}$$

But in terms of the function $f(x)$ in (27), this means that

$$f(E_m^A + E_{mn}^B) = f(E_m^A) \, f(E_m^B) \qquad (28)$$

we can then ask what kinds of functions satisfy $f(x+y) = f(x) \, f(y)$. Answer: only the exponential. Thus, the function $f(x)$ may be written as

$$f(x) = \frac{e^{-\beta x}}{z} \qquad (29)$$

where β and z are unknown constants. Try to think about other functions that satisfy (28) and you will see that they can always be mapped in the form (29) by suitably renaming the constants.

Ok. Now we don't need two systems any more. The conclusion we reached is that because of the postulate (26) and our argument on statistical independence, the equilibrium probabilities must have the form

$$P_m = \frac{e^{-\beta E_m}}{z} \qquad (30)$$

where z and β are constants to be determined.

10

The constant z is easy since it is just a normalization. But can we say something about β. Could we maybe relate it to temperature somehow. The answer, at this stage is no. The connection with temperature requires a bit more physical input.

But we can nonetheless say a few things about β. Most importantly, we can see that it must be positive: given two states with $E_m > E_n$, we get from (30)

$$\frac{P_m}{P_n} = e^{-\beta(E_m - E_n)}$$

If β was negative then $P_m > P_n$, so the more excited is a state, the higher is it's occupation probability. The system would then be unstable and the world would explode. Since we are all alive, then we must have $\beta > 0$. This is called <u>thermodynamic stability</u>.

Another cool thing we can say about β is that the <u>energy</u> is monotonically <u>decreasing</u> with β. Since β is a property of the state, it then makes sense that $\beta \propto 1/T$, as we know energy should increase with temperature.

To see all this we differentiate U in Eq (6) with respect to β

$$\frac{\partial U}{\partial \beta} = \frac{\partial}{\partial \beta} \left[\frac{\sum_m E_m e^{-\beta E_m}}{\sum_m e^{-\beta E_m}} \right]$$

$$= \frac{\sum_m (-E_m)^2 e^{-\beta E_m}}{\sum_m e^{-\beta E_m}} - \frac{\left(\sum_m E_m e^{-\beta E_m}\right)\left(-\sum_m E_m e^{-\beta E_m}\right)}{\left(\sum_m e^{-\beta E_m}\right)^2}$$

But this can be written as

$$\frac{\partial U}{\partial \beta} = - \sum_m E_m^2 \, P_m + \left(\sum_m E_m \, P_m \right)^2$$

$$= - \left[\langle H^2 \rangle - \langle H \rangle^2 \right]$$

(31)

The quantity $\langle H^2 \rangle - \langle H \rangle^2$ is a variance, which is always non-negative. Thus

$$\frac{\partial U}{\partial \beta} < 0$$

(32)

which, as stated, means U is monotonically decreasing with β.

Some remarks

Now is a good time to stop and make some remarks about what we have seen so far.

First, let's talk about constants. the Boltzmann constant, which appears in $\beta = 1/k_B T$, has the value

$$k_B = 8.617 \times 10^{-5} \text{ eV}/K \qquad (33)$$

It therefore converts temperature to energy. For historical reasons we ended up measuring temperature in another unit called Kelvin. But temperature is a form of energy and could thus simply be measured in eV. For instance

$$T = 300 \text{ K} \quad \leadsto \quad k_B T = 0.026 \text{ eV} \qquad (34)$$

when we set $k_B = 1$ we are simply measuring temperature in energy units. So there is no need to worry: I guarantee that you will never run into confusion with this.

Setting constants to 1 is not merely a silly theoretical thing. It actually teaches you something. ☺

the same is true for $\hbar = 1$. You don't need to be afraid to do it. It simply sets energies and frequencies to have the same units (e.g. $\hbar \omega$ becomes only ω).
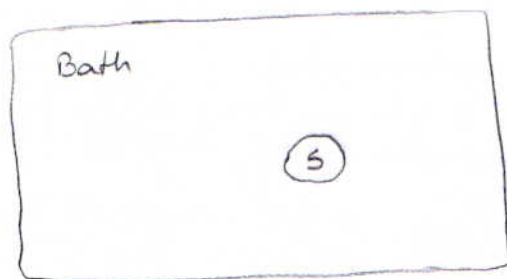
when $k_B = \hbar = 1$,

$$\frac{\hbar \omega}{k_B T} = \frac{\omega}{T} \quad \overleftarrow{\text{Elegant}}$$

clumsy ↗

Next comment: how does a system reach a Gibbs state? thermal equilibrium is a rather universal feature of nature, so there are very many ways through which a system may equilibrate. But the most common situation is when a system is coupled to a heat bath: that is, a very large body which is itself in thermal equilibrium at a temperature T.

this is cool because the system for us can be anything, from 1 electron to 10^{23} atoms. But there is no sense in talking about the temperature of a single electron. Thus, the β that appears in the Gibbs is simply a parameter of the state. But with this system + bath picture, we can think about β as the temperature of the bath which the system is connected to

As another remark, let us talk about what is the meaning of probabilities like p_n. This question has actually bugged quite a lot the founding fathers of statistical mechanics.

In kinetic theory we have this very mechanical picture of thermodynamic systems such as gases and liquids. A gas then, is a system of 10^{23} particles moving according to Newton's second law of Schrödinger's equation. Their evolution, then, was ultimately deterministic.

But, of course, when you have 10^{23} particles and highly chaotic dynamics, this determinism doesn't quite fit in. So the founding fathers came up with a construct called <u>ensembles</u>. In an ensemble, you imagine there are an infinite number of copies of your system, each starting in a random initial condition. This was they the way they found to cope with the inevitable randomness of 10^{23} particles.

But the ensemble concept is only necessary if one has a <u>frequentist</u> view of probability, like they did. If you adopt a <u>Bayesian</u> viewpoint, then this entire idea of ensembles becomes completely unnecessary.

In the Bayesian approach, we take the <u>probabilities</u> as <u>quantifying our lack of knowledge about the world</u>. Thus, we say "the prob. of finding the qubit in $|1\rangle$ is p_1", we mean a probability <u>conditioned on the information</u> that qubit is in equilibrium.

Of course, if we were to measure the qubit and find it in $|1\rangle$, then it's probability would no longer be given by the Gibbs state, but would be simply $P_1 = 1$, $P_0 = 0$. But that's because we updated our information about the qubit with the outcome of the measurement.

On the other hand, if we knew nothing about the qubit, then our best guess would be $P_0 = P_1 = 1/2$. But if we further update this with the fact that the qubit is in equilibrium, we update the probabilities to the Gibbs state.

Statistical independence of non-interacting systems

As already discussed in pages (9) and (10), one of the most important features of the Gibbs state is that when two systems do not interact, their probabilities will be a product of the individual probabilities. That is, they will be statistically independent.

Let's make this more formal. If we have two systems A and B, with eigenstuff

$$H_A |m_A\rangle = E^A_{m_A} |m_A\rangle \qquad (35)$$

$$H_B |m_B\rangle = E^B_{m_B} |m_B\rangle$$

then the eigenstates and eigenenergies of

$$(36)$$

$$H = H_A + H_B \qquad (37)$$

will be

$$H |m_A, m_B\rangle = E_{m_A, m_B} |m_A, m_B\rangle \qquad (38)$$

where

$$E_{m_A, m_B} = E_{m_A} + E_{m_B}$$

Thus, the Gibbs state in this case will be labeled by a collective index (m_A, m_B):

$$P_{m_A, m_B} = \frac{e^{-\beta E_{m_A m_B}}}{Z} \qquad (39)$$

where

$$Z = \sum_{m_A, m_B} e^{-\beta E_{m_A, m_B}} \qquad (40)$$

But using (38) we get

$$z = \sum_{M_A, M_B} e^{-\beta E_{M_A}^A} e^{-\beta E_{M_B}^B} = \sum_{M_A} e^{-\beta E_{M_A}^A} \sum_{M_B} e^{-\beta E_{M_B}^B}$$

$$= z_A \, z_B \qquad (41)$$

Thus (39) simply factors as

$$P_{M_A, M_B} = \frac{e^{-\beta E_{M_A}^A}}{z_A} \frac{e^{-\beta E_{M_B}^B}}{z_B} = P_{M_A} P_{M_B} \qquad (42)$$

when the probabilities factor as a product, then A and B are statistically independent.

I know this sounds simple and a bit silly, but this turns out to be a huge <u>simplification</u>. the reason is that since A and B are statistically independent, as far as A is concerned it is as if B didn't exist.

For instance, in page 3 we worked out the properties of a single qubit. If we now have a sample of 10^{23} qubits which do not interact, then nothing changes: the properties of each qubit remain exactly as before. Of course, this holds only for <u>non-interacting systems</u>. Interactions make everything more complicated (but also more interesting!). So the message I want you to remember is

$$\boxed{\begin{array}{c} \text{Non-interacting systems} \Leftarrow \text{easy} \\ \text{Just focus on each one individually} \end{array}} \qquad (43)$$