Thermodynamics and information

Consider an arbitrary random variable X which can take on d different levels $x_0, x_1, \ldots, x_{d-1}$. We don't know what are the probabilities $P_n = P(X = x_n)$. But we know the expectation value of some function $g(x)$

$$\langle f(x) \rangle = \sum_n P_n f(x_n) = G \tag{1}$$

where G is some number.

Can we figure out what is P_n only from this fact? Of course not! We only know one number, G, so there is definitely not enough information.

But lets ask this from the point of view of statistical inference: what is the best guess we can make about P_n given that all we know is that $\langle g(x) \rangle = G$. The answer to this was given by in parts by Shannon in 1948 and then Jaynes in 1957.

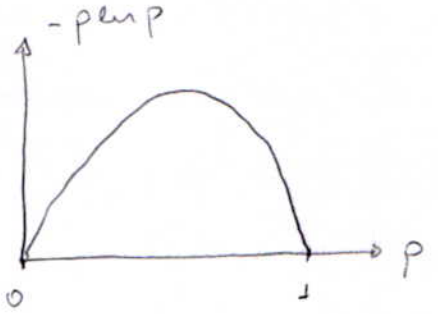
Shannon showed that there is a unique function for quantifying the amount of uncertainty in a distribution P_n, which is given by

$$\boxed{S = -\sum_n P_n \ln P_n} \tag{2}$$

this is called the Shannon entropy and is the starting point of information theory.

The entropy quantifies our lack of information about x.

It is useful to remember the behavior of the function $-p \ln p$



It is zero whenever we have certainty; that is, when $p=0$ or $p=1$. Otherwise, it is always positive.

We thus see that Eq (2) is a sum of positive contributions, which are larger for P_n's far from 0 or 1. The maximum of S occurs for a uniform distribution

$$\boxed{\text{If} \quad P_n = \frac{1}{d} \quad \text{then} \quad S = \ln d} \tag{3}$$

Conversely, if $P_\ell = 1$ for some ℓ, and all other P_n's are zero, then $S=0$. Thus, we have

$$\begin{aligned} &\text{Minimum } S = 0 &&: \text{ deterministic case} \\ &\text{Maximum } S = \ln d &&: \text{ uniform distribution} \\ & && \quad \text{(maximum ignorance)} \end{aligned} \tag{4}$$

We can also define the <u>information</u> about x

$$I = \ln(d) - S \tag{5}$$

It goes from $\ln(d)$, which the largest amount of information we may have, to 0 in the uniformly distributed case.

Let's now go back to our original problem: we have some rv X, for which we don't know P_n. Instead, all we know is that $\langle g(x) \rangle = G$. What is the best (i.e., least biased) distribution P_n that agrees with this information?

The answer is given by the principle of maximum entropy, or MaxEnt: the distribution P_n is that which maximizes S subject to the constraint $\langle g(x) \rangle = G$

To carry out this constrained maximization we introduce Lagrange multipliers, by defining

$$\tilde{S} = - \sum_n P_n \ln P_n + \alpha \left(1 - \sum_n P_n \right) + \beta \left(G - \sum_n g(x_n) P_n \right) \qquad (6)$$

the Lagrange multiplier α ensures that the probabilities are normalized

$$\frac{\partial \tilde{S}}{\partial \alpha} = 1 - \sum_n P_n = 0 \rightsquigarrow \sum_n P_n = 1 \qquad (7)$$

Similarly, β ensures that $\langle g(X) \rangle = G$

$$\frac{\partial \tilde{S}}{\partial \beta} = G - \sum_n g(x_n) P_n = 0 \rightsquigarrow \sum_n g(x_n) P_n = G \qquad (8)$$

then we maximize $\tilde{S}$ with respect to each P_n:

$$\frac{\partial \tilde{S}}{\partial P_n} = - \ln P_n - \frac{P_n}{P_n} - \alpha - \beta g(x_n) = 0$$

thus

$$\ln P_n = -1 - \alpha - \beta g(x_n)$$

or

$$P_n = e^{-\alpha-1} e^{-\beta g(x_n)} \tag{9}$$

We can call $e^{-\alpha-1} = 1/Z$, since this is just a constant. Thus we conclude the distribution which maximizes the entropy subject to $\langle g(x) \rangle = G$ is given by

$$P_n = \frac{e^{-\beta g(x_n)}}{Z} \tag{10}$$

—||—

Side comment: I forgot to say, but the MaxEnt also applies to the case where we have no extra information. Then there is only one Lagrange multiplier

$$\tilde{S} = -\sum_n P_n \ln P_n + \alpha \left(1 - \sum_n P_n\right)$$

and the maximization gives

$$\frac{\partial \tilde{S}}{\partial P_n} = -\ln P_n - 1 - \alpha = 0$$

Thus, all P_n must be equal. By normalization this then implies

$$P_n = \frac{1}{d}$$

which is the uniform distribution. The entropy in this case achieves its largest value possible

$$S = \ln(d)$$

Thus, the conclusion is that if we know nothing about X, then the best possible guess we can make is a uniform distribution (maximum ignorance)

Entropy and free energy

The connection with statistical mechanics should now be a bit evident. We identify the Gibbs state of a system as the state which maximizes the entropy subject to the constraint that

$$U = \langle H \rangle = \sum_n E_n P_n = \mathcal{E} \qquad (11)$$

for some value of $\mathcal{E}$

$$\tilde{S} = -\sum_n P_n \ln P_n + \alpha\left(1 - \sum_n P_n\right) + \beta\left(\mathcal{E} - \sum_n E_n P_n\right) \qquad (12)$$

which will result in

$$P_n = \frac{e^{-\beta E_n}}{Z} \qquad \beta = \frac{1}{T} \qquad (13)$$

We thus see that the structure of the Gibbs state may be found solely from statistical inference. If you are interested in learning more about this, check out E. Jaynes, Physical Review, 106, 620-630 (1957).

A comment, by the way: the Shannon entropy, Eq (2), also appeared (way before Shannon) in the works of Gibbs. So the fact that the same quantity appears in both statistical mechanics and information theory reinforces the deep connection between the two approaches.

Going back to (12), let us now define a new quantity, which is called free energy (the origin of the name "free" will be clarified later)

$$\boxed{F = U - TS} \tag{14}$$

or

$$F = \sum_m \left(E_m P_m + T P_m \ln P_m \right) \tag{15}$$

Then (12) may be written as

$$\tilde{\zeta} = -\beta F + \alpha \left(1 - \sum_m P_m \right)$$

Thus we see that the following two statements are equivalent:

$$\boxed{\begin{array}{l} \text{1) Equilibrium is the state which maximizes} \\ \quad \text{the entropy subject to } \langle H \rangle = U \\ \\ \text{2) Equilibrium is the state which minimizes} \\ \quad \text{the free energy} \end{array}} \tag{16}$$

this renders a really cool interpretation to the free energy. At zero temperature, equilibrium would be the state of smallest energy (the ground state). But for $T \neq 0$ we also have some disorder. So equilibrium is not the state which minimizes $\langle H \rangle$. Instead, there is a competition between $\langle H \rangle$ and $-TS$ and so equilibrium is instead the state which minimizes the free energy

Thermodynamics

In equilibrium the P_n are given by the Gibbs distribution (13). Then writting

$$\ln P_n = -\beta E_n - \ln Z \tag{17}$$

and substituting in (15), we find

$$\boxed{F = -T \ln Z} \qquad \boxed{\left(Z = e^{-\beta F} \right)} \tag{18}$$

Thus we see that the partition function is more than just a normalization constant, but actually has a physical meaning. In fact, it turns out, all thermodynamic quantities are actually encoded in Z. For instance, start with

$$Z = \sum_n e^{-\beta E_n} \tag{19}$$

and differentiate with respect to β

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

we thus see that

$$\frac{1}{Z} \frac{\partial Z}{\partial \beta} = - \frac{\sum_n E_n e^{-\beta E_n}}{\sum_n e^{-\beta E_n}} = -U$$

Hence, from Z we can compute the internal energy as

$$\boxed{U = \langle H \rangle = -\frac{\partial}{\partial \beta} \ln Z} \tag{20}$$

Comparing now (18) and (20) we get

$$U = \frac{\partial}{\partial \beta}\left(\frac{F}{T}\right)$$

using

$$\frac{\partial}{\partial \beta} = \frac{\partial T}{\partial \beta}\frac{\partial}{\partial T} = -T^2\frac{\partial}{\partial T} \qquad (21)$$

we get

$$\boxed{U = -T^2 \frac{\partial}{\partial T}\left(\frac{F}{T}\right)} \qquad (22)$$

And finally, once we have F and U, then it's easy to compute the entropy since

$$\boxed{F = U - TS} \qquad (23)$$

so $S = (U-F)/T$. Or, expanding (22),

$$\boxed{U = F - T\frac{\partial F}{\partial T}} \qquad (24)$$

Thus

$$\boxed{S = \frac{U-F}{T} = -\frac{\partial F}{\partial T}} \qquad (25)$$

Heat capacity (specific heat)

Another important quantity, which is one of the most measured quantities in the lab, is the heat capacity or specific heat (there is actually a difference between these two names, which I will explain in a second). It is define as the sensitivity of the system's internal energy, to changes in temperature

$$C = \frac{\partial U}{\partial T} \tag{26}$$

Since we set $k_B = 1$, for us C is dimensionless. Otherwise C would have units of k_B. Experimental data is usually given in J/mol K. To convert (26) to these units, multiply it by the gas constant

$$R = 8.314 \text{ J/mol K} \tag{27}$$

Using Eq (24) we may also write C as

$$C = \frac{\partial F}{\partial T} - \frac{\partial}{\partial T}\left[T \frac{\partial F}{\partial T} \right] = \frac{\partial F}{\partial T} - \frac{\partial F}{\partial T} - T\frac{\partial^2 F}{\partial T^2}$$

thus

$$\boxed{C = \frac{\partial U}{\partial T} = - T\frac{\partial^2 F}{\partial T^2}} \tag{28}$$

A key property of the heat capacity is that it is always non-negative. In the previous lecture notes we showed that

$$\frac{\partial U}{\partial \beta} = - \left[\langle H^2 \rangle - \langle H \rangle^2 \right]$$

Thus, using (21) we conclude that

$$\boxed{C = \frac{1}{T^2} \left[\langle H^2 \rangle - \langle H \rangle^2 \right] \geqslant 0} \tag{29}$$

This is really important: the heat capacity represents the fluctuations of the energy in equilibrium.

From this and (28) we then conclude that the following 3 statements are equivalent

- The heat capacity is non-negative
- The energy is a monotonically increasing function of T (30)
- The free energy is a concave function of T

(31)

Extensivity

So far we have said nothing about the size of the system, or the number of particles in it. But now let's see what happens when the system is composed of N non-interacting parts.

We have seen in the previous lecture notes that in this case the Gibbs probabilities will be statistically independent. So if we let E_m denote the energy of each sub-system (which we assume are all identical) then

$$P_{m_1 \dots m_N} = \frac{e^{-\beta E_{m_1}}}{Z_1} \cdots \frac{e^{-\beta E_{m_N}}}{Z_1} \tag{31}$$

where Z_1 is the partition function of one subsystem

$$Z_1 = \sum_m e^{-\beta E_m} \tag{32}$$

(which is the same for all of them). The partition function of the full system will then be

$$\boxed{Z(N) = Z_1^N} \tag{33}$$

Consequently, the free energy (18) will be given by

$$\boxed{F = -NT \ln Z_1} \tag{34}$$

We see that the free energy is an extensive quantity: it scales with the number of particles.

The same is also true for the internal energy (20), since

$$\boxed{U = -N \frac{\partial}{\partial \beta} \ln Z_1 = N U_1} \tag{35}$$

And then, by (23), it will be true for the entropy

$$S = N S_1 \tag{36}$$

and by (26), for the heat capacity as well

$$C = N C_1 \tag{37}$$

This is now why people distinguish between heat capacity and specific heat

$$\text{specific heat} = \text{heat capacity per particle} \tag{38}$$

Sometimes people also talk about the mass specific heat, which would be the heat capacity per unit mass. Or the molar specific heat, and so on.

— // —

The thermodynamic recipe book

To summarize, when studying equilibrium systems, I recommend starting with the following order

$$\boxed{\begin{array}{lll} 1)\ \text{Find } Z & 3)\ U = -\frac{\partial}{\partial \beta} \ln Z & 5)\ C = \frac{\partial U}{\partial T} \\ 2)\ F = -T \ln Z & 4)\ S = (U - F)/T & \end{array}} \tag{39}$$

In this way, the only real effort is in finding Z. After that, you get F, U, S, C only by differentiating quantities.

Example: qubit vs. harmonic oscillator

In the previous lecture notes we found the following results for a qubit and a harmonic oscillator

qubit | QHO

$$E_0 = 0, \quad E_1 = \epsilon \qquad \Big| \qquad E_n = \omega n \tag{40}$$

$$Z = 1 + e^{-\beta\epsilon} \qquad \Big| \qquad Z = \frac{1}{1 - e^{-\beta\omega}}$$

Thus, applying the recipe (39) we get

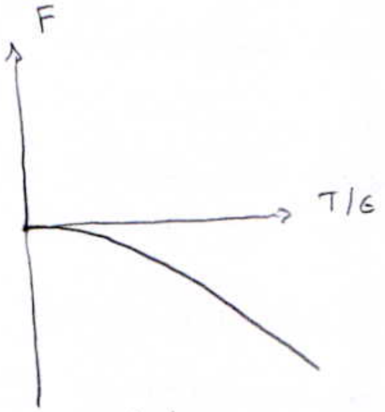
$$F = -T\ln(1 + e^{-\beta\epsilon}) \qquad \Big| \qquad F = T\ln(1 - e^{-\beta\omega}) \tag{41}$$

$$U = \frac{\epsilon}{e^{\beta\epsilon} + 1} \qquad \Big| \qquad U = \frac{\omega}{e^{\beta\omega} - 1} \tag{42}$$

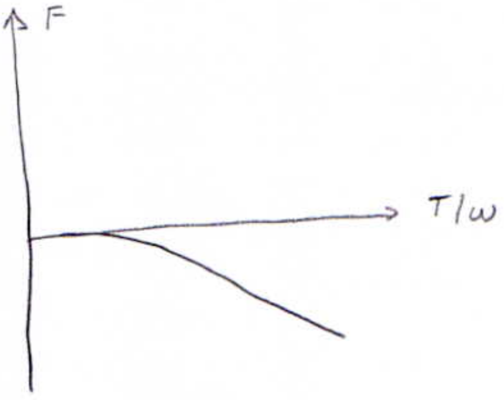
$$S = \frac{\beta\epsilon}{e^{\beta\epsilon} + 1} + \ln(1 + e^{-\beta\epsilon}) \qquad \Big| \qquad S = \frac{\beta\omega}{e^{\beta\omega} - 1} - \ln(1 - e^{-\beta\omega}) \tag{43}$$

$$C = \frac{(\beta\epsilon)^2 e^{\beta\epsilon}}{(e^{\beta\epsilon} + 1)^2} \qquad \Big| \qquad C = \frac{(\beta\omega)^2 e^{\beta\omega}}{(e^{\beta\omega} - 1)^2} \tag{44}$$

these quantities look like this

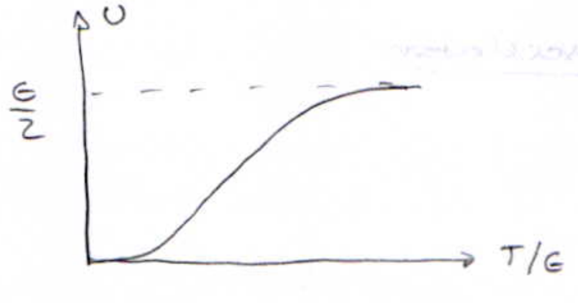


Qubit

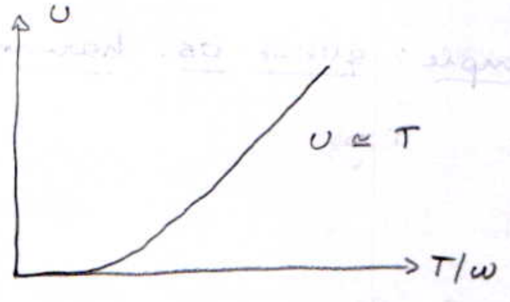


QHO

Not very interesting



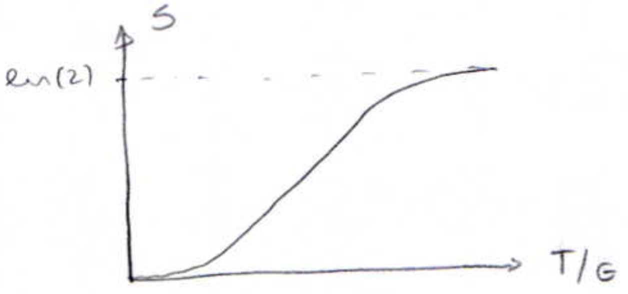
Qubit

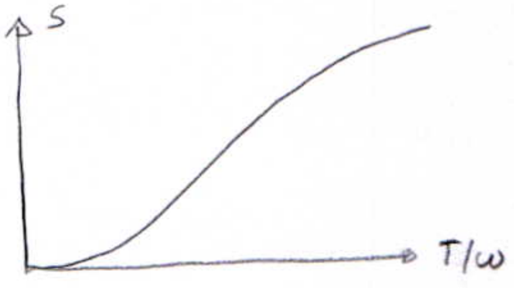


QHO

For the qubit $0 < U < \epsilon/2$.

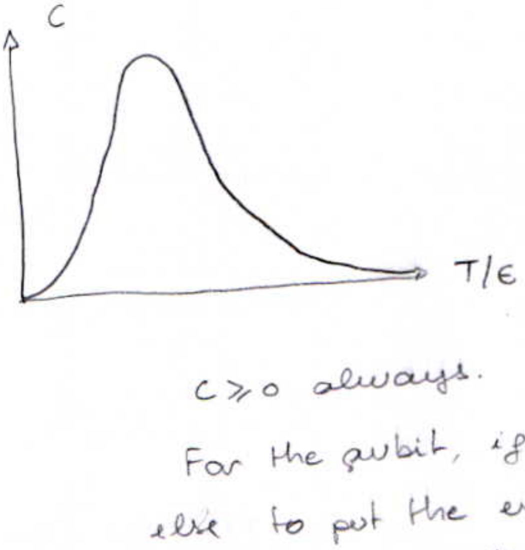
For the QHO U is unbounded

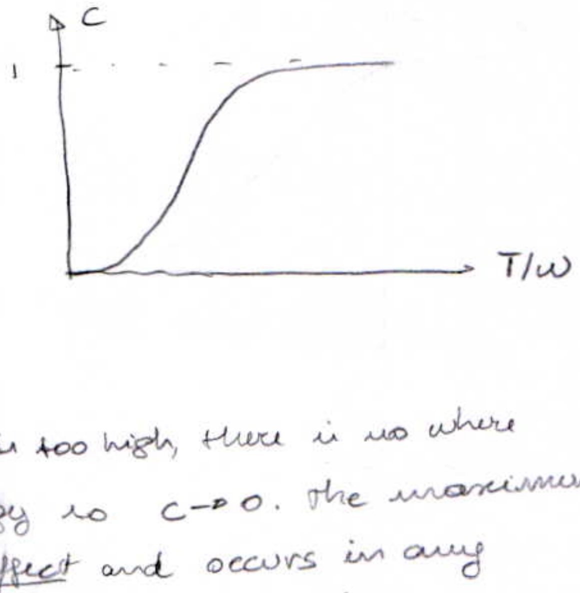




For the qubit $0 < S < \ln(2)$

In both cases $S \to 0$ when $T \to 0$.





$C \geqslant 0$ always.

For the qubit, if T is too high, there is no where else to put the energy so $C \to 0$. The maximum is called Schottky effect and occurs in any system with a finite number of levels.

For the QHO, if T is high the specific heat becomes independent of temperature. This is the law of Dulong and Petit

Spin systems and response functions

Spins system play a huge role in statistical mechanics, as they offer a clean playground for studying phase transitions and many-body effects (the name "Ising model" is a good representative of these problems)

Let us consider a spin 1/2 particle. It can be described by the Pauli matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{45}$$

The thing is constructed in a basis where σ_z is diagonal. We shall denote its eigenvalues by a number $\sigma = \pm 1$. That is

$$\sigma_z |\sigma\rangle = \sigma |\sigma\rangle, \qquad \sigma = \pm 1 \tag{46}$$

where

$$|\sigma = 1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad |\sigma = -1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{47}$$

Of course "spin 1/2" and "qubit" are just the same thing: they are both 2-level systems. The association between $|\sigma\rangle$ and the computational basis, $|0\rangle$ and $|1\rangle$, is a matter of convention. I usually like to choose

$$\begin{aligned} |0\rangle &= |\sigma = +1\rangle \qquad \text{(spin up)} \\ |1\rangle &= |\sigma = -1\rangle \qquad \text{(spin down)} \end{aligned} \tag{48}$$

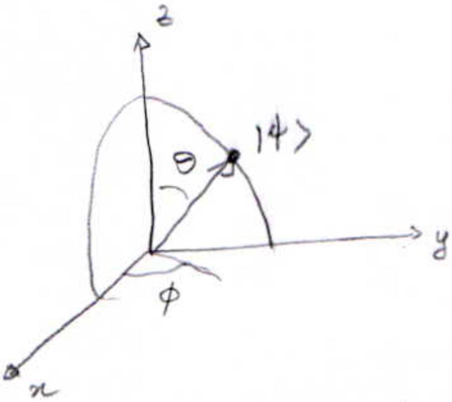
The most general quantum mechanical state of a qubit/spin 1/2 may be parametrized as

$$|\psi\rangle = \begin{pmatrix} \cos\theta/2 \\ e^{i\phi}\sin\theta/2 \end{pmatrix} \qquad (49)$$

where θ and ϕ are just parameters. Since states are invariant under a global phase, it turns out that it suffices to take θ and ϕ in the range

$$0 \leq \theta < \pi \qquad (50)$$
$$0 < \phi \leq 2\pi$$

Interestingly, this is exactly like the way we parametrize the unit sphere in terms of polar and azimuthal angles. This therefore introduces an interesting idea: to represent the state (49), which is a complex two-dimensional vector, as a point in a real three-dimensional sphere



This is called the Bloch sphere. The states $|\sigma=+1\rangle$ is the north pole and $|\sigma=-1\rangle$ is the south pole

To better understand the meaning of the sphere it is very useful to know that

$$\langle \sigma_x \rangle = \langle \psi | \sigma_x | \psi \rangle = \sin\theta \cos\phi$$

$$\langle \sigma_y \rangle = \langle \psi | \sigma_y | \psi \rangle = \sin\theta \cos\phi \qquad (51)$$

$$\langle \sigma_z \rangle = \langle \psi | \sigma_z | \psi \rangle = \cos\theta$$

which are precisely the spherical coordinates of a unit vector in $\mathbb{R}^3$. We will have a lot more to say about Bloch's sphere as we move along

— // —

If a spin 1/2 is subject to a magnetic field in the z direction, the Hamiltonian will be given by

$$\boxed{H = - h\, \sigma_z} \qquad (52)$$

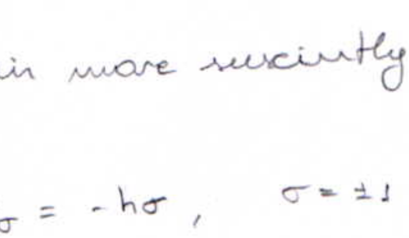
where I choose units so that h has dimensions of energy. The Hamiltonian (52) is already diagonal, because σ_z is diagonal. Thus, the energies are

$$E_+ = -h \qquad E_- = +h$$

or we can write this more succinctly as

$$E_\sigma = -h\sigma, \quad \sigma = \pm 1 \qquad (53)$$

The quantum number is therefore $\sigma = \pm 1$, the ground state is spin up (aligned with h) and the gap is 2h.



The partition function is then

$$Z = \sum_\sigma e^{-\beta E_\sigma} = e^{\beta h/2} + e^{-\beta h/2} \tag{54}$$

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

The Gibbs probabilities are then

$$P_\sigma = \frac{e^{\beta h \sigma}}{2\cosh(\beta h)} \tag{55}$$

Recall that from this we can now compute the expectation value of any observable A as

$$\langle A \rangle = \sum_\sigma \langle \sigma | A | \sigma \rangle P_\sigma \tag{56}$$

the obvious choice, in this case, are the Pauli matrices. However, σ_x and σ_y [Eq (45)] have no diagonal matrix elements in the $|\sigma\rangle$ basis. Thus $\langle \sigma | \sigma_x | \sigma \rangle = \langle \sigma | \sigma_y | \sigma \rangle = 0$ and

$$\langle \sigma_x \rangle = \langle \sigma_y \rangle = 0 \tag{57}$$

As for σ_z, we have $\langle \sigma | \sigma_z | \sigma \rangle = \sigma$ so

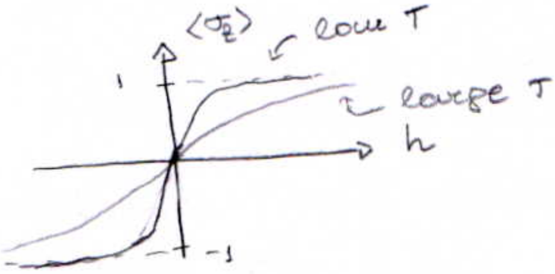
$$\langle \sigma_z \rangle = \sum_\sigma \sigma P_\sigma = \frac{e^{\beta h} - e^{-\beta h}}{2\cosh(\beta h)}$$

or, simplifying,

$$\boxed{m := \langle \sigma_z \rangle = \tanh(\beta h)} \tag{58}$$

which is the <u>magnetization</u>

This is a beautiful and important result. It looks something like this



the spin always tends to polarize in the same direction as h. When the temperature is low, then a small field already pushes $\langle\sigma_z\rangle$ close to ± 1. But for high T, this polarization is less pronounced and higher fields are required. Indeed, if we expand (58) close to $h=0$ we get <u>Curie's law</u>

$$\langle\sigma_z\rangle \simeq \frac{h}{T} \qquad (59)$$

so near $h=0$ the behavior is linear and the slope is $1/T$.

Now let's go back to Bloch's sphere. We can ask, what is the state $|\psi\rangle$ (or the angles θ, ϕ) which correspond to the Gibbs state? The answer is that there is none! There is no quantum mechanical ket which represents thermal equilibrium.

The easiest way of seeing this is by noting that for any pure state $|\psi\rangle$, if we take the expectation values (51), they will always satisfy

$$\langle\sigma_x\rangle_\psi^2 + \langle\sigma_y\rangle_\psi^2 + \langle\sigma_z\rangle_\psi^2 = 1 \qquad (60)$$

However, the same is definitely not true for the Gibbs state since

$$\langle \sigma_x \rangle^2 + \langle \sigma_y \rangle^2 + \langle \sigma_z \rangle^2 = \tanh^2(\beta h) < 1 \tag{61}$$

Thus, we can say that a Gibbs state lies inside the Bloch sphere, whereas $|\psi\rangle$ always lies on the surface.

As we will learn soon, a Gibbs state is what is known as a mixed state, which is not described by kets, like $|\psi\rangle$, but by a more general object called a density matrix. Don't worry, we will get there. I promise.

—//—

Going back now to (58), it is interesting to show that it can also be written as a derivative of the partition function, just like the energy in (20). But now the derivative is with respect to the magnetic field:

$$\frac{\partial Z}{\partial h} = \frac{\partial}{\partial h} \sum_\sigma e^{\beta h \sigma} = \sum_\sigma \beta \sigma e^{\beta h \sigma}$$

Thus

$$\frac{1}{Z} \frac{\partial Z}{\partial h} = \beta \sum_\sigma \sigma \frac{e^{\beta h \sigma}}{Z} = \beta \langle \sigma_z \rangle$$

Hence

$$\langle \sigma_z \rangle = T \frac{\partial}{\partial h} \ln Z \tag{62}$$

But $-T \ln Z$ is the free energy F so

$$\boxed{m = \langle \sigma_z \rangle = -\frac{\partial F}{\partial h}} \qquad (63)$$

To find the magnetization, simply differentiate the free energy with respect to the field. the magnetization is what we call a *response function*. You apply a field h and $\langle \sigma_z \rangle$ is how the system respond. Later we will consider systems composed of many *interacting* spins. But the idea will remain the same: we apply a field and the system will respond. By understanding *how* it responds we learn something about the system.

Together with the magnetization, another important quantity is the *susceptibility*

$$\boxed{\chi = \frac{\partial m}{\partial h} = -\frac{\partial^2 F}{\partial h^2}} \qquad (64)$$

It describes the sensitivity of m to the magnetic field. Thus, it plays a role analogous to the specific heat (compare with Eq (28)).

In particular, the susceptibility at zero field is obtained from (59) and reads

$$\chi(h=0) = \frac{1}{T} \qquad (65)$$

which is known as *Curie's law*

We may also associate the susceptibility with the *fluctuations of the magnetization*, just like the heat capacity is related to the energy fluctuations. From (62)

$$m = T\left(\frac{Z'}{Z}\right) \qquad Z' = \frac{\partial Z}{\partial h}$$

so

$$\chi = \frac{\partial m}{\partial h} = T\left(\frac{Z''}{Z} - \left(\frac{Z'}{Z}\right)^2\right)$$

But

$$\frac{Z'}{Z} = \frac{\langle \sigma_z \rangle}{T}$$

and

$$\frac{Z''}{Z} = \frac{1}{Z}\frac{\partial^2 Z}{\partial h^2} = \frac{1}{Z}\sum_\sigma (\beta\sigma)^2 e^{\beta h \sigma}$$

$$= \frac{1}{T^2}\sum_\sigma \sigma^2 \frac{e^{\beta h \sigma}}{Z}$$

$$= \frac{1}{T^2}\langle \sigma_z^2 \rangle$$

Of course, I'm being a bit silly because $\sigma_z^2 = 1$. But I leave it like this because the idea generalizes to other spin values or to multiple spins. Anyway, we conclude that

$$\boxed{\chi = \frac{1}{T}\left[\langle \sigma_z^2 \rangle - \langle \sigma_z \rangle^2\right]} \qquad (66)$$

which is analogous to (29). In particular, since this is a variance, this number must be positive, so m must be a monotonically increasing function of the field h.

Thermodynamic sanity checks

There are several sanity checks we can make about thermodynamic quantities, which are useful when dealing with more complicated problems. For concreteness, we will suppose here the energy levels are always numbered in increasing order:

$$E_0 \leqslant E_1 \leqslant E_2 \leqslant \dots \tag{67}$$

To start, let us then consider the zero temperature limit. We have

$$\frac{p_n}{p_0} = e^{-\beta(E_n - E_0)} < 1 \tag{68}$$

As $T \to 0$, $e^{-\beta(E_n - E_0)}$ becomes very small, so the probabilities concentrate on the ground state

$$p_0 \to 1 \tag{69}$$

At $T=0$, the system tends to the ground state. But watch out, $p_0 = 1$ only if the ground state is non-degenerate. We could have, for instance

$$E_0 = E_1 = E_2 < E_3 \leqslant E_4 \leqslant \dots \tag{70}$$

this is a 3-fold degeneracy. In this case we would then have $p_0 = p_1 = p_3$, so when $T \to 0$ they would each tend to 1/3

Thus, more generally, for all states with the same ground-state energy, we have

$$\boxed{P_{gs,i} = \frac{1}{g} \quad , \quad \begin{array}{l} g = \text{degeneracy of the GS} \\ i = 1, \dots, g \end{array}} \tag{71}$$

with all other $P_n = 0$.

consequently, U, F and S will tend to

$$\boxed{\begin{aligned} U &= E_{gs} \\ S &= \ln(g) \\ F &= U - TS = U = E_{gs} \end{aligned}} \tag{72}$$

As for the specific heat, it is very useful to know the following fact: close to T=0, the only significantly populated states are the GS and the first excited state. So, if we denote the energy gap by Δ, then this will be effectively a 2-level system, so C will be given by Eq (44; ref)

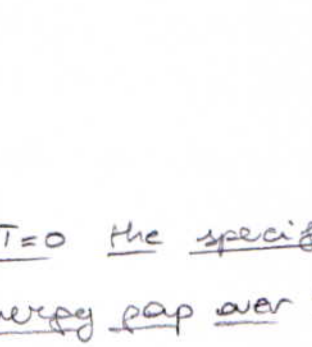
$$C \simeq \frac{(\beta\Delta)^2 e^{\beta\Delta}}{(e^{\beta\Delta}+1)^2} \tag{73}$$

For $T \to 0$ $(\beta \to \infty)$ this can be further approximated to

$$\boxed{C \simeq (\beta\Delta)^2 e^{-\beta\Delta}} \tag{74}$$

what matters here is the exponential: close to T=0 the specific heat tends to zero exponentially with the energy gap over T

this is used, for instance, to know that there is an energy gap in BCS superconductivity.



Next let's talk about T large. In this case it is hard to say much if the number of states is infinite. So let's suppose we have d levels

$$E_0 \leq E_1 \leq \dots \leq E_{d-1} \tag{75}$$

If T is very large then all configurations become equally

$$\boxed{P_n = \frac{1}{d} \quad \text{for all } n} \tag{76}$$

consequently

$$\boxed{\begin{aligned} U &= \frac{E_0 + \dots + E_{d-1}}{d} \\ S &= \ln(d) \end{aligned}} \tag{77}$$

and $F \to -\infty$. Since $C = \partial U / \partial T$, then we also have $C \to 0$.

Summary

$T \to 0$	$T \to \infty$ (finite dimension)
$P_{0i} = \frac{1}{g}$, g = degeneracy of GS $P_n = 0$, $n \neq 0$	$P_n = \frac{1}{d}$ d = dimension of Hilbert space
$U = E_{gs}$	$U = \frac{E_0 + \dots + E_{d-1}}{d}$
$S = \ln(g)$	$S = \ln(d)$
$C \simeq e^{-\beta\Delta}$, Δ = gap	$C \to 0$

On top of that, we also have the monotonicity properties

- U is monotonically increasing with T
- S is monotonically increasing with T
- F is concave with T
- $C \geq 0$, but not necessarily monotonic.