Density matrix theory

Before we can start having fun with more complicated problems, we need to develop some more sophisticated tools for dealing with thermal states. Recall that the expectation value of an observable A in equilibrium was given by

$$\langle A \rangle = \sum_n \langle n | A | n \rangle P_n \qquad (1)$$

where $P_n = e^{-\beta E_n}/Z$. This is a weird kind of average because it mixes quantum expectation values, $\langle n|A|n\rangle$, with classical expectation values $\sum_n (\dots) P_n$. Moreover, as we discussed in the previous lecture, it is in general not possible to find a ket $|\psi\rangle$ that can be associated to a Gibbs state.

This motivates the introduction of a new object, called the density matrix, which generalizes the notion of ket. The density matrix of the Gibbs state is defined as

$$\boxed{\rho = \sum_n \frac{e^{-\beta E_n}}{Z} |n\rangle\langle n|} \qquad (2)$$

It is therefore a sum of projection operators $|n\rangle\langle n|$, with statistical weights given by the probabilities P_n. Before I explain why it makes sense to define an object such as this, let us first see how to rewrite the average (1) in terms of ρ.

This is done using the concept of a trace

$$\mathrm{tr}(M) := \sum_{m} \langle m|M|m\rangle \tag{3}$$

Some quick properties of the trace that you should definitely know

- Independent of basis choice: $\sum_{m} \langle m|M|m\rangle = \sum_{i} \langle i|M|i\rangle$

- Sum of eigenvalues: $\mathrm{tr}(M)$ = sum of eigs of M. (4)

- Cyclic: $\mathrm{tr}(AB) = \mathrm{tr}(BA)$
$$\mathrm{tr}(ABC) = \mathrm{tr}(CAB) \neq \mathrm{tr}(ACB) \tag{5}$$

- $\mathrm{tr}(|\psi\rangle\langle\phi|) = \langle\phi|\psi\rangle$ (6)

Using (6) in particular, we see that (1) may be written as

$$\boxed{\langle A\rangle = \mathrm{tr}(A\rho)} \tag{7}$$

or $\mathrm{tr}(\rho A)$ since the trace is cyclic. This equation represents the generalization of expectation values to density matrices, with the Gibbs state (2) being just one particular case.

Meaning of the density matrix

Density matrices are designed so as to combine quantum states with ignorance. If the quantum state of a system is $|\psi\rangle$ then its density matrix is defined as

$$\rho = |\psi\rangle\langle\psi| \tag{8}$$

which is called a pure state. In this case Eq (7) becomes

$$\langle A \rangle = \mathrm{tr}\left(A|\psi\rangle\langle\psi|\right) = \langle\psi|A|\psi\rangle \tag{9}$$

which is just the familiar quantum mechanical expression.

However, it could be that we don't know exactly what the state of the system is. That is, we have some degree of ignorance about it. For instance maybe the state is $|\psi_1\rangle$ with probability q_1 or maybe it is $|\psi_2\rangle$ with probability $q_2 = 1 - q_1$. The corresponding density matrix would then be

$$\rho = q_1 |\psi_1\rangle\langle\psi_1| + q_2|\psi_2\rangle\langle\psi_2| \tag{10}$$

which is what we call a mixed state. Then (7) would become

$$\langle A \rangle = q_1 \langle\psi_1|A|\psi_1\rangle + q_2\langle\psi_2|A|\psi_2\rangle \tag{11}$$

which, just like the Gibbs state, is a mixture of quantum averages $\langle\psi|A|\psi\rangle$ with classical averages $\sum_i (\cdots) p_i$.

In this way, the density matrix offers a generalization to the concept of ket, by treating together quantum features and classical probability theory.

For instance, suppose we have a qubit which is prepared 99% of the time in the north pole of the Bloch sphere, $|\psi_1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. However, 1% of the time the machine preparating it commits an error and spits out a state

$$|\psi_2\rangle = \begin{pmatrix} \cos\theta/2 \\ \sin\theta/2 \end{pmatrix}$$

where $\theta = 5°$ (just to have some concrete numbers). The state of the system in this case will not be a pure state, but rather a mixed state of the form

$$\rho = q_1 |\psi_1\rangle\langle\psi_1| + q_2 |\psi_2\rangle\langle\psi_2| \tag{12}$$

where $q_1 = 0.99$ and $q_2 = 0.01$. We can also write it more explicitly:

$$|\psi_1\rangle\langle\psi_1| = \begin{pmatrix} 1 \\ 0 \end{pmatrix} (1 \;\; 0) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

$$|\psi_2\rangle\langle\psi_2| = \begin{pmatrix} \cos\theta/2 \\ \sin\theta/2 \end{pmatrix} \left(\cos\frac{\theta}{2} \;\; \sin\frac{\theta}{2} \right) = \begin{pmatrix} \cos^2\theta/2 & 2\sin\theta \\ 2\sin\theta & \sin^2\theta/2 \end{pmatrix}$$

Thus

$$\rho = \begin{pmatrix} q_1 + q_2 \cos^2\theta/2 & 2q_2 \sin\theta \\ 2q_2 \sin\theta & q_2 \sin^2\theta/2 \end{pmatrix} \tag{13}$$

I know it looks ugly, but that's what it is

Of course, in (10) I considered only two possibilities. But the definition is more general. Given any set of states $|\psi_i\rangle$ and a corresponding set of probabilities q_i, the density matrix will be given by

$$\boxed{\begin{gathered} \rho = \sum_i q_i |\psi_i\rangle\langle\psi_i| \\ q_i \in [0,1], \quad \sum_i q_i = 1. \end{gathered}} \tag{14}$$

Now suppose the machine is actually quite crappy so that half of the time it prepares the state $|z_+\rangle = |\sigma=+1\rangle = \begin{pmatrix}1\\0\end{pmatrix}$ and half of the time the state $|z_-\rangle = \begin{pmatrix}0\\1\end{pmatrix}$. Then (12) becomes

$$\rho = \frac{1}{2}\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} + \frac{1}{2}\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \tag{15}$$

This is called the <u>maximally</u> <u>mixed</u> <u>state</u>: it cannot get more ignorant than this. It means we know absolutely nothing about the system.

Instead, suppose now that the machine prepares with 50/50 probability the states

$$|x_+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} \qquad |x_-\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix} \tag{16}$$

then

$$\rho = \frac{1}{2}\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix} + \frac{1}{2}\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \tag{17}$$

Thus, in this case we get the same state as (15), even though we began in drastically different situations. This is called the <u>ambiguity of mixtures</u>; when we mix stuff we loose information about what we began with in the first place.

Here are some general properties of the density matrix, which can be derived solely from (14):

1) $\rho^\dagger = \rho$ (18)

2) $\boxed{\mathrm{tr}(\rho) = 1}$ Normalization of the quantum state (19)

check: $\mathrm{tr}(\rho) = \sum_i q_i \, \mathrm{tr}(|\psi_i\rangle\langle\psi_i|) = \sum_i q_i \underbrace{\langle\psi_i|\psi_i\rangle}_{1} = 1.$

3) Since ρ is Hermitian it can be diagonalized as

$$\boxed{\rho = \sum_k p_k \, |k\rangle\langle k|} \tag{20}$$

for some basis $|k\rangle$. From $\mathrm{tr}(\rho) = 1$ we see that

$$\sum_k p_k = 1 \tag{21}$$

4) For any state $|\phi\rangle$,

$$\langle\phi|\rho|\phi\rangle = \sum_i q_i \, |\langle\phi|\psi_i\rangle|^2 \geq 0 \tag{22}$$

= Prob. of finding state $|\phi\rangle$ given the system is at ρ.

5) ρ is positive semi-definite: all $p_k \geq 0$.

Proof: put $|\phi\rangle = |k\rangle$ in (22).

combined with (21) this gives

$$\boxed{p_k = \mathrm{eigs}(\rho): \quad p_k \in [0,1], \quad \sum_k p_k = 1} \tag{23}$$

Thus, the eigenvalues of ρ behave like probabilities

We define the purity of the state ρ as

$$\boxed{\text{Purity} = tr(\rho^2) \leq 1} \tag{24}$$

the reason why the purity is ≤ 1 is because the eigenvalues of ρ^2 are p_k^2. Since the trace is the sum of the eigenvalues, we get $\text{Purity} = \sum_k p_k^2 \leq 1$ due to (23).

The purity will be 1 if and only if one of the p_k is 1 and the others are zero. Thus

$$\text{Purity} = 1 \quad \text{if} \quad \rho = |k_0\rangle\langle k_0| \quad \text{for some } k_0 \tag{25}$$

which is precisely a pure state, like (8). Thus, the smaller the purity the more mixed is the state.

Bloch's sphere for mixed states

Let us consider once again the case of a qubit. The density matrix will be 2×2 so we may parametrize it as

$$\rho = \begin{pmatrix} p & c \\ c^* & 1-p \end{pmatrix} \tag{26}$$

where $p \in [0,1]$ is the population and c is the coherence. To further clarify the meaning of these parameters, it is convenient to note that

$$\begin{aligned} \langle \sigma_z \rangle &= \mathrm{tr}(\sigma_z \rho) = 2p-1 \\ \langle \sigma_x \rangle &= \mathrm{tr}(\sigma_x \rho) = c+c^* \\ \langle \sigma_y \rangle &= \mathrm{tr}(\sigma_y \rho) = i(c-c^*) \end{aligned} \tag{27}$$

Thus, if we define

$$s_\alpha = \langle \sigma_\alpha \rangle \tag{28}$$

then we may also parametrize ρ as

$$\rho = \frac{1}{2}\begin{pmatrix} 1+s_z & s_x - i s_y \\ s_x + i s_y & 1-s_z \end{pmatrix} = \frac{1}{2}(1 + \vec{s}\cdot\vec{\sigma}) \tag{29}$$

I will leave for you as an exercise to compute the purity (24) of this state. It reads

$$\mathrm{tr}(\rho^2) = \frac{1+s_x^2+s_y^2+s_z^2}{2} = \frac{1+s^2}{2} \tag{30}$$

For instance, the maximally mixed state (15) has the smallest possible purity

$$\mathrm{tr}(\rho^2) = \frac{1}{2}, \quad s = 0 \tag{31}$$

This offers us an interesting interpretation: we previously saw that pure states live on the surface of the unit sphere. Now we can say that mixed states live <u>inside</u> <u>Bloch's sphere</u>. Moreover, the radius of the Bloch sphere quantifies the purity

$$s^2 = s_x^2 + s_y^2 + s_z^2 = \begin{cases} 0 & \text{maximally mixed} \\ 1 & \text{pure state} \end{cases} \qquad (32)$$

Graphically this looks like this