Statistical Mechanics: lecture 4

Systems composed of multiple particles

Disclaimer: so far we have essentially considered only systems composed of a single particle. In these notes we will first develop the theoretical framework for dealing with multiple particles, then explore the physical consequences.

Recommended reading: Nielsen chapter 2.1.
Salinas, chapter 5.

Chapter 3

Composite Systems

3.1 The age of Ulkron

So far we have considered only a single quantum system described by a basis $|i\rangle$. Now we turn to the question of how to describe mathematically a system composed of two or more sub-systems. Suppose we have two sub-systems, which we call A and B. They can be, for instance, two qubits: one on earth and the other on mars. How to write states and operators for this joint system? This is another postulate of quantum theory. But instead of postulating it from the start, I propose we first try to formulate what we intuitively expect to happen. Then we introduce the mathematical framework that does the job.

For me, at least, I would expect the following to be true. First, if $\{|i\rangle_A\}]$ is a set of basis vectors for A and $\{|j\rangle_B\}$ is a basis vector for B, then a joint basis for AB should have the form $|i, j\rangle$. For instance, for two qubits one should have four possibilities:

$$|0,0\rangle, \qquad |0,1\rangle, \qquad |1,0\rangle, \qquad |1,1\rangle.$$

Secondly, again at least in my intuition, one should be able to write down operators that act *locally* as if the other system was not there. For instance, we know that for a single qubit σ_x is the bit flip operator:

$$\sigma_x|0\rangle = |1\rangle, \qquad \sigma_x|1\rangle = |0\rangle.$$

If we have two qubits, I would expect we should be able to define two operators σ_x^A and σ_x^B that act as follows:

$$\sigma_x^A|0,0\rangle = |1,0\rangle, \qquad \sigma_x^B|0,0\rangle = |0,1\rangle.$$

Makes sense, no? Similarly, we expect that if we apply both σ_x^A and σ_x^B the order shouldn't matter:

$$\sigma_x^A\sigma_x^B|0,0\rangle = \sigma_x^B\sigma_x^A|0,0\rangle = |1,1\rangle.$$

This means that operators belonging to different systems should commute:

$$[\sigma_x^A, \sigma_x^B] = 0. \tag{3.1}$$

The tensor/Kronecker product

The mathematical structure that implements these ideas is called the **tensor product** or **Kronecker product**. It is, in essence, a way to glue together two vector spaces to form a larger space. The tensor product between two states $|i\rangle_A$ and $|j\rangle_B$ is written as

$$|i, j\rangle = |i\rangle \otimes |j\rangle. \tag{3.2}$$

The symbol $\otimes$ separates the two universes. We read this as "i tens j" or "i kron j". I like the "kron" since it reminds me of a crappy villain from a Transformers or Marvel movie. Similarly, the operators σ_x^A and σ_x^B are defined as

$$\sigma_x^A = \sigma_x \otimes I, \qquad \sigma_x^B = I \times \sigma_x, \tag{3.3}$$

where I is the identify matrix.

In order for us to make sense of these definitions, we must of course specify the basic rules for how objects behave around the $\otimes$. Lucky for you, there is only one rule that we really need to remeber: *stuff to the left of $\otimes$ only interact with stuff to the left and stuff to the right only interact with stuff to the right*. In symbols:

$$\boxed{(A \otimes B)(C \otimes D) = (AC) \otimes (BD),} \tag{3.4}$$

In this rule A, B, C and D can be *any* mathematical object, as long as the multiplications AC and BD make sense.

Let's see how this works. For instance,

$$\sigma_x^A|0, 0\rangle = (\sigma_x \otimes I)(|0\rangle \otimes |0\rangle) = (\sigma_x|0\rangle) \otimes (I|0\rangle).$$

The only thing I did was apply the rule (3.4) to combine stuff to the left of $\otimes$ with stuff to the left and stuff to the right with stuff to the right. Now that we have $\sigma_x|0\rangle$ we are back to the single qubit business, so we can just write $\sigma_x|0\rangle = |1\rangle$. Then we recombine the result:

$$(\sigma_x|0\rangle) \otimes (I|0\rangle) = |1\rangle \otimes |0\rangle = |1, 0\rangle,$$

which is what we would expect intuitively. As another example, the property (3.1), that operators pertaining to different systems should commute, now follows directly from our definitions:

$$\sigma_x^A\sigma_x^B = (\sigma_x \otimes I)(I \otimes \sigma_x) = (\sigma_x \otimes \sigma_x),$$
$$\sigma_x^B\sigma_x^A = (I \otimes \sigma_x)(\sigma_x \otimes I) = (\sigma_x \otimes \sigma_x),$$

which are definitely the same thing.

Everything we just said also holds for systems composed of 3, 4 or any number of parts, of course. In this case we simply add more and more $\otimes$. For instance, for 3 qubits, $\sigma_x^B = I \otimes \sigma_x \otimes I$ and so on.

Let us also talk about how to combine other kinds of objects. Remember that all we need is for the multiplications in the composition rule (3.4) to make sense. For instance, an operation that makes no sense is

$$(\langle 0| \otimes |0\rangle)(\sigma_x \otimes \sigma_x) = \text{crazy nonsense},$$

because even though $\langle 0|\sigma_x$ makes sense, the operation $|0\rangle\sigma_x$ does not.

An operation which does make sense is

$$\langle k, \ell|i, j\rangle = (\langle k| \otimes \langle \ell|)(|i\rangle \otimes |j\rangle) = (\langle k|i\rangle) \otimes (\langle \ell|j\rangle).$$

The objects that remain here are two numbers and the tensor product of two numbers is also a number. Thus, we arrive at a rule for the inner product:

$$\langle k, \ell|i, j\rangle = \langle k|i\rangle\langle \ell|j\rangle. \tag{3.5}$$

Outer products are similarly defined:

$$|k, \ell\rangle\langle i, j| = |k\rangle\langle i| \otimes |\ell\rangle\langle j|. \tag{3.6}$$

One can also come up with somewhat weird operations which nonetheless make sense. For instance,

$$(\langle k| \otimes |\ell\rangle)(|i\rangle \otimes \langle j|) = (\langle k|i\rangle) \otimes |\ell\rangle\langle j| = (\langle k|i\rangle)|\ell\rangle\langle j|.$$

In the last equality I used the fact that $\langle k|i\rangle$ is just a number.

Be cool about notation

Here is a really really really good tip: *be cool duuuuuude*. There are many ways of expressing quantum states and operators for composite systems. *Don't be rigid about notation. Just be clear so that people know what you mean.* For instance, if we talk about states, the following notations are equivalent:

$$\boxed{|i, j\rangle_{AB} = |i\rangle_A \otimes |j\rangle_B = |i\rangle_A|j\rangle_B.} \tag{3.7}$$

In the third notation adding the suffixes A and B is essential. Otherwise one would not know if $|i\rangle$ belongs to A or B. For completeness I also added the suffixes to the first two notations. Sometimes that is redundant. But if there is ever room for confusion, add it: it doesn't cost much.

A notation like $|i\rangle_A|j\rangle_B$ also allows you to move things around and write, for instance, $|j\rangle_B|i\rangle_A$. There is no room for confusion because you know one symbol belongs to A and the other to B. The same is true for operator multiplication. For instance,

$$\sigma_x^B|i, j\rangle_{AB} = |i\rangle_A\sigma_x^B|j\rangle_B.$$

Notice that there is zero room for misinterpretation: the notation is not rigid, but no one will interpret it wrong.

I *strongly* recommend you be cool about the notation. Each notation is useful for a different thing, so feel free to change them at will. Just make sure there is no room for misinterpretation.

Matrix representation of the Kronecker product

When using the Kronecker product in a computer, it is standard to order the basis elements $|i, j\rangle$ in *lexicographic order*: for each entry of the first, you loop over all elements of the last. For instance, if A and B have each dimension 3, we get

$$|0,0\rangle, \quad |0,1\rangle, \quad |0,2\rangle, \quad |1,0\rangle, \quad |1,1\rangle, \quad |1,2\rangle, \quad |2,0\rangle, \quad |2,1\rangle, \quad |2,2\rangle.$$

Conversely, if we have 3 qubits, we would order the basis elements as

$$|0,0,0\rangle, \qquad |0,0,1\rangle, \qquad |0,1,0\rangle, \qquad |0,1,1\rangle \qquad |1,0,0\rangle \quad \ldots$$

This ordering is not mandatory. But is extremely convenient for the following reason.

We then associate to each element a unit vector. For instance, for 2 qubits we would have,

$$|0,0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad |0,1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \qquad |1,0\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \qquad |1,1\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{3.8}$$

The matrix elements of an operator of the form $A \times B$ then becomes, using the property (3.4)

$$\langle k, \ell | A \otimes B | i, j \rangle = \langle k|A|i\rangle \langle \ell | B | j \rangle = A_{ki} B_{\ell j}.$$

If we now present these guys in a matrix, since we loop over all elements of the second index, for each element of the first, the matrix form of this will look like

$$A \otimes B = \begin{pmatrix} A_{0,0}B & \ldots & a_{0,d_A-1}B \\ \vdots & \ddots & \vdots \\ a_{d_A-1,0}B & \ldots & a_{d_A-1,d_A-1}B \end{pmatrix}. \tag{3.9}$$

This is just an easy of visualizing the matrix: for each A_{ki} we introduce a full block B. To be clear what is meant by this, consider for instance

$$\sigma_x \otimes \sigma_x = \begin{pmatrix} 0\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & 1\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ 1\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & 0\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \tag{3.10}$$

This provides an automated way to construct tensor product matrices. The final result is not very intuitive. But computationally, it is quite trivial. Specially since the Kronecker product is implemented in any library. In MATLAB they call it `kron()` whereas in Mathematica they call it `KroneckerProduct[]`. These functions are really useful. You should really try to play with them a bit.

As a consistency check, we can verify that the same logic also holds for vectors. For instance,

$$|0,0\rangle = |0\rangle \otimes |0\rangle = \begin{pmatrix} 1\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ 0\begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{3.11}$$

Proceeding similarly leads to all elements in Eq. (3.8).

3.2 Entanglement and mixed states

So far we have talked about how to represent kets and operators of composite systems using the tensor product. Now let's see what kind of physics this produces. Suppose we have two qubits, A and B. If qubit A is on Earth and qubit B is on Mars, it is reasonable to assume that they are each in *local states*, such as

$$|\eta\rangle_A = \alpha|0\rangle_A + \beta|1\rangle_A, \qquad |\phi\rangle_B = \gamma|0\rangle_B + \delta|1\rangle_B.$$

Then, the global state of AB will be

$$\begin{aligned}|\eta\rangle_A \otimes |\phi\rangle_B &= \Big[\alpha|0\rangle_A + \beta|1\rangle_A\Big] \otimes \Big[\gamma|0\rangle_B + \delta|1\rangle_B\Big] \\ &= \alpha\gamma|0,0\rangle_{AB} + \alpha\delta|0,1\rangle_{AB} + \beta\gamma|1,0\rangle_{AB} + \beta\delta|1,1\rangle_{AB}.\end{aligned}$$

If we look at the second line, this state seems like simply a linear combination of the four basis elements $|i,j\rangle_{AB}$. However, this is not an arbitrary linear combination. It contains a very special choice of parameters which are such that you can perfectly factor the state into something related to A times something related to B. Cases like this are what we call a **product state**. If A and B are in a product state, they are completely independent of each other.

However, quantum theory also allows us to have more general linear combinations which are not necessarily factorable into a product. Such a general linear combination has the form

$$|\psi\rangle_{AB} = \sum_{i,j} \psi_{ij}|i,j\rangle_{AB}, \tag{3.12}$$

where $\psi_{i,j}$ are any set of complex numbers satisfying $\sum_{ij}|\psi_{ij}|^2 = 1$. When a state like this cannot be written as a product,[1] we say A **and** B **are entangled**. An important set of entangled states are the so called **Bell states**:

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}\Big[|0,0\rangle + |1,1\rangle\Big], \tag{3.13}$$

$$|\Phi^-\rangle = \frac{1}{\sqrt{2}}\Big[|0,0\rangle - |1,1\rangle\Big], \tag{3.14}$$

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}\Big[|0,1\rangle + |1,0\rangle\Big], \tag{3.15}$$

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}\Big[|0,1\rangle - |1,0\rangle\Big]. \tag{3.16}$$

These states *cannot* be factored into a product of local states (please take a second to convince yourself of that!). In fact, we will learn soon that they are **maximally**

[1] That is, when we cannot decompose $\psi_{ij} = f_j g_j$.

entangled states. If you are familiar with the theory of angular momentum, you will also notice that these states (specially $|\Psi^{\pm}\rangle$) are exactly to the singlet and triplet states of two spin 1/2 particles. Moreover, it is useful to note that they form an orthonormal basis for the Hilbert space of the two qubits.

The Controlled NOT (CNOT)

We usually entangle systems by applying *gates*. That is, unitary transformations stemming from the interaction between the systems. A popular entangling gate for two qubits is the CNOT. It is defined by the unitary

$$U_{\text{CNOT}} = |0\rangle\langle 0|_A \otimes I_B + |1\rangle\langle 1|_A \otimes \sigma_x^B. \tag{3.17}$$

Qubit A is the *control bit*. If it is in $|0\rangle$, we do nothing on B. But if it is in $|1\rangle$, we apply the bit flip operation σ_x on B:

$$\begin{aligned} U_{\text{CNOT}}|0\rangle_A|\psi\rangle_B &= |0\rangle_A|\psi\rangle_B, \\ U_{\text{CNOT}}|0\rangle_A|\psi\rangle_B &= |0\rangle_A\Big(\sigma_x^B|\psi\rangle_B\Big). \end{aligned}$$

Suppose we now start with two qubits reset to $|0\rangle_A|0\rangle_B$. We can prepare the two qubits in a Bell state by applying two gates. First, we apply a Hadamard gate to A:

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{3.18}$$

This produces

$$H_A|0\rangle_A|0\rangle_B = |+\rangle_A|0\rangle_B = \left(\frac{|0\rangle_A + |1\rangle_A}{\sqrt{2}}\right)|0\rangle_B.$$

This is a gate acting only on A. It is a *local operation* and thus cannot entangle A and B. To entangle them we now apply the CNOT (3.17). It gives

$$U_{\text{CNOT}}H_A|0\rangle_A|0\rangle_B = \frac{|0,0\rangle_{AB} + |1,1\rangle_{AB}}{\sqrt{2}},$$

which is nothing but the Bell state (3.13). The other Bell states may be generated in a similar way, by starting with the four possible states $|i, j\rangle$:

$$\begin{aligned} |\Phi^+\rangle &= U_{\text{CNOT}}H_A|0\rangle_A|0\rangle_B, \\ |\Phi^-\rangle &= U_{\text{CNOT}}H_A|1\rangle_A|0\rangle_B, \\ |\Psi^+\rangle &= U_{\text{CNOT}}H_A|0\rangle_A|1\rangle_B, \\ |\Psi^-\rangle &= U_{\text{CNOT}}H_A|1\rangle_A|1\rangle_B. \end{aligned} \tag{3.19}$$

Density matrices from entanglement

Now I want you to recall our original discussion in Sec. 2.1. We saw that the concept of density matrix naturally appeared when we considered a crappy machine that produced quantum states with some classical uncertainty. What we found was that it was possible to combine quantum and classical effects by introducing an object of the form

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \tag{3.20}$$

where the $|\psi_i\rangle$ are arbitrary states and the p_i are arbitrary probabilities. This construction may have left you with the impression that the density matrix is only necessary when we want to mix quantum and classical stuff. That density matrices are not really a quantum thing. Now I want to show you that this is not the case. It is *definitely* not the case. I will show you that there is an intimate relation between mixed states and entanglement. And this relation is one the key steps relating quantum mechanics and information theory.

Essentially, the connection is made by the notion of **reduced state** or **reduced density matrix**. When a composite system is in a product state $|\psi\rangle_A \otimes |\phi\rangle_B$, it makes sense to say the state of A is simply $|\psi\rangle_A$. But if A and B are entangled, then what is exactly the "state" of A? To warm up, consider first a bipartite state of AB of the form

$$|\psi\rangle_{AB} = \sum_i c_i \, |i\rangle \otimes |i\rangle \tag{3.21}$$

for certain coefficients c_i satisfying $\sum_i |c_i|^2 = 1$. If $c_i = 1$ for some i and all other $c_j = 0$ then $|\psi\rangle = |i\rangle \otimes |i\rangle$ and we get a product state. In any other case, the state will be entangled.

Now let O_A be an operator which acts only on system A. That is, an operator which has the form $O_A = O_A \otimes I_B$. The expectation value of O_A in the state (3.21) will be

$$\langle O_A \rangle = \langle\psi|(O_A \otimes I_B)|\psi\rangle \tag{3.22}$$

Carrying out the calculation we get:

$$\begin{aligned}
\langle O_A \rangle &= \sum_{i,j} c_i^* c_j \, ((\langle i| \otimes \langle i|)(O_A \otimes I_B)(|j\rangle \otimes |j\rangle) \\
&= \sum_{i,j} c_i^* c_j \, \langle i|O_A|j\rangle\langle i|j\rangle \\
&= \sum_i |c_i|^2 \langle i|O_A|i\rangle.
\end{aligned}$$

The sandwich that remains is now performed only over the reduced state of A. However, each sandwich $\langle i|O_A|i\rangle$ is now weighted by a factor $|c_i|^2$.

We now ask the following question: can we attribute a state $|\psi_A\rangle$ for system A such that the above result can be expressed as $\langle\psi_A|O_A|\psi_A\rangle$. This is actually the same question we asked in Sec. 2.1. And we saw that the answer is *no*. In general, there is no pure

state we can associate with A. Instead, if we wish to associate a quantum state to A, it will have to be a mixed state, described by a density matrix of the form

$$\rho_A = \sum_i |c_i|^2 |i\rangle\langle i| \tag{3.23}$$

then the expectation value of A becomes

$$\langle A \rangle = \mathrm{tr}(A\rho_A) \tag{3.24}$$

This result has *extremely* important consequences. Eq. (3.23) has exactly the same form as Eq. (3.20), with the classical probabilities p_i replaced by quantum coefficients $|c_i|^2$. But there is absolutely *nothing* classical here. Nothing. We started with a pure state. We are talking about a purely quantum effect. Notwithstanding, we see that in general the state of A will be mixed. If $c_i = 1$ for some i and all other $c_j = 0$ then Eq. (3.23) reduces to $\rho_A = |i\rangle\langle i|$, which is a pure state. In all other cases, the state of A will be mixed. Thus,

> When AB are entangled, the reduced state of A and B will be mixed.

To give an example, suppose AB is in the Bell state (3.13). This state has the form of Eq. (3.21) with $c_i = 1/\sqrt{2}$. Thus, it is easy to apply Eq. (3.23), which gives

$$\rho_A = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{3.25}$$

We therefore see that the reduced state of A is actually the maximally mixed state. This is a feature of all Bell states and it is the reason we call them maximally entangled states. This is super interesting, if you think about it: A Bell state is a pure state, so we know exactly what the state of AB is. However, we know absolutely *nothing* about A alone.

3.3 Reduced density matrices and the partial trace

The state ρ_A in Eq. (3.23) is called a **reduced density matrix**. And the procedure that led us from $|\psi\rangle_{AB}$ to ρ_A is called the partial trace. This is the quantum analog of computing the marginal $P(x)$ of a joint probability distribution $P(x, y)$. In this section I will teach you how to make this procedure in a more algorithmic way.

The partial trace

Consider a bipartite system AB. Let $|a\rangle$ and $|b\rangle$ be basis sets for A and B. Then a possible basis for AB is the tensor basis $|a, b\rangle$. What I want to do is investigate the trace operation within the full AB space. Any operator in AB can always be decomposed as

$$O = \sum_\alpha A_\alpha \otimes B_\alpha, \tag{3.26}$$

for some index α and some set of operators A_α and B_α. So, to start, let us consider simply an operator of the form $O = A \otimes B$. Then, by linearity it will be easy to extend to more general operators.

We begin by computing the trace of $O = A \otimes B$ in the $|a, b\rangle$ basis:

$$\begin{aligned}
\mathrm{tr}(O) &= \sum_{a,b} \langle a, b|O|a, b\rangle \\
&= \sum_{a,b} (\langle a| \otimes \langle b|)(A \otimes B)(|a\rangle \otimes |b\rangle) \\
&= \sum_{a,b} \langle a|A|a\rangle \otimes \langle b|B|b\rangle \\
&= \sum_{a} \langle a|A|a\rangle \sum_{b} \langle b|B|b\rangle.
\end{aligned}$$

I got rid of the $\otimes$ in the last line because the kron of two numbers is a number. The two terms in this formula are simply the trace of the operators A and B in their respective Hilbert spaces. Whence, we conclude that

$$\boxed{\mathrm{tr}(A \otimes B) = \mathrm{tr}(A)\,\mathrm{tr}(B).} \tag{3.27}$$

We started with an operator having support on two Hilbert spaces and ended up tracing everything, so that we are left with only a single number.

We can now imagine an operation where we only trace over one of the Hilbert spaces and obtain an operator still having support on the other part. This is what we call the **partial trace**. It is defined as

$$\boxed{\mathrm{tr}_A(A \otimes B) = \mathrm{tr}(A)B, \qquad \mathrm{tr}_B(A \otimes B) = A\,\mathrm{tr}(B)} \tag{3.28}$$

When you "trace over A", you eliminate the variables pertaining to A and what you are left with is an operator acting only on B. This is something we often forget, so please pay attention: the result of a partial trace is still an operator. More generally, for an arbitrary operator O as defined in Eq. (3.26), we have

$$\mathrm{tr}_A\, O = \sum_\alpha \mathrm{tr}(A_\alpha)B_\alpha, \qquad \mathrm{tr}_B\, O = \sum_\alpha A_\alpha\,\mathrm{tr}(B_\alpha). \tag{3.29}$$

An important example is the partial trace of an operator of the form $|a, b\rangle\langle a', b'|$. To take the partial trace, remember that this can be written as

$$|a, b\rangle\langle a', b'| = |a\rangle\langle a'| \otimes |b\rangle\langle b'|.$$

The partial trace over B, for instance, will simply go right through the first part and act only on the second part; i.e.,

$$\mathrm{tr}_B\, |a, b\rangle\langle a', b'| = |a\rangle\langle a'|\,\mathrm{tr}\left\{|b\rangle\langle b'|\right\}$$

Thus, we conclude that

$$\boxed{\mathrm{tr}_A\,|a,b\rangle\langle a',b'| = \delta_{a,a'}\;|b\rangle\langle b'|, \qquad \mathrm{tr}_B\,|a,b\rangle\langle a',b'| = |a\rangle\langle a'|\;\delta_{b,b'}.} \tag{3.30}$$

Reduced density matrices

We are now ready to introduce the idea of a reduced density matrix in a more formal way. Given a bipartite system ρ_{AB} we define the reduced density matrix of A and B as

$$\boxed{\rho_A = \mathrm{tr}_B\,\rho_{AB}, \qquad \rho_B = \mathrm{tr}_A\,\rho_{AB}} \tag{3.31}$$

Let us now practice with some examples.

Example: Bell states

To practice, consider the Bell state example that led us from the bipartite state (3.13) to the reduced state (3.25). The global density matrix is

$$\rho_{AB} = |\Phi_1\rangle\langle\Phi_1| = \frac{1}{2}\Big\{|0,0\rangle\langle 0,0| + |0,0\rangle\langle 1,1| + |1,1\rangle\langle 0,0| + |1,1\rangle\langle 1,1|\Big\} \tag{3.32}$$

To take the partial trace we use Eq. (3.30) to find:

$$\rho_A = \frac{1}{2}\Big\{|0\rangle\langle 0| + |1\rangle\langle 1|\Big\} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{3.33}$$

with an identical result for ρ_B.

Example: Partially entangled states

Consider now a state of the form

$$|\psi\rangle_{AB} = \sqrt{p}|0,1\rangle_{AB} + \sqrt{1-p}|0,1\rangle_{AB}, \tag{3.34}$$

for some number $p \in [0,1]$. If $p = 1/2$ we recover the Bell state (3.13). To take the partial trace we proceed as before:

$$\rho_{AB} = |\psi\rangle\langle\psi| = p|0,1\rangle\langle 0,1| + (1-p)|1,0\rangle\langle 1,0| + \sqrt{p(1-p)}\Big(|0,1\rangle\langle 1,0| + |1,0\rangle\langle 0,1|\Big).$$

Due to Eq. (3.30), the last two terms will always give 0 when we take the partial trace. We are then left with

$$\rho_A = p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1|,$$

$$\rho_B = (1-p)|0\rangle\langle 0| + p|1\rangle\langle 1|.$$

Example: X states

X states of two qubits are density matrices of the form

$$\rho = \begin{pmatrix} p_1 & 0 & 0 & \beta \\ 0 & p_2 & \alpha & 0 \\ 0 & \alpha^* & p_3 & 0 \\ \beta^* & 0 & 0 & p_4 \end{pmatrix}. \tag{3.35}$$

Normalization requires $\sum_i p_i = 1$ and positivity imposes constraints on the allowed values of α and β. The ordering of states here is the lexicographic order we discussed in Sec. 3.1. Namely, $|0,0\rangle, |0,1\rangle, |1,0\rangle, |1,1\rangle$. Thus, we can write this X state more explicitly as

$$\begin{aligned} \rho = {} & p_1|0,0\rangle\langle 0,0| + p_2|0,1\rangle\langle 0,1| + p_3|1,0\rangle\langle 1,0| + p_4|1,1\rangle\langle 1,1| \\ & + \alpha|0,1\rangle\langle 1,0| + \alpha^*|1,0\rangle\langle 0,1| \\ & + \beta|0,0\rangle\langle 1,1| + \beta^*|1,1\rangle\langle 0,0|. \end{aligned} \tag{3.36}$$

The meaning of α and β now become a bit more clear: they represent, respectively, the non-local coherences between $\{|0,1\rangle, |1,0\rangle\}$ and $\{|0,0\rangle, |1,1\rangle\}$. From Eq. (3.36) it is easy to take the partial trace:

$$\rho_A = \mathrm{tr}_B\, \rho = \begin{pmatrix} p_1 + p_2 & 0 \\ 0 & p_3 + p_4 \end{pmatrix}, \tag{3.37}$$

$$\rho_B = \mathrm{tr}_A\, \rho = \begin{pmatrix} p_1 + p_3 & 0 \\ 0 & p_2 + p_4 \end{pmatrix}. \tag{3.38}$$

We see that for X states, the reduced density matrices are diagonal. The entries which are set to zero in Eq. (3.35) are precisely the ones that would lead to non-zero diagonals in the reduced state. If we now look for observables, for instance, we will then find that

$$\langle \sigma_x^A \rangle = \langle \sigma_x^B \rangle = \langle \sigma_y^A \rangle = \langle \sigma_y^B \rangle = 0.$$

Non-local observables, on the other hand, can be non-zero. For instance, one may check that

$$\langle \sigma_x^A \sigma_x^B \rangle = \alpha + \alpha^* + \beta + \beta^*.$$

Partial trace looses information

If we have a state which is of the form $\rho_{AB} = \rho_A \otimes \rho_B$, then Eq. (3.28) directly gives us $\mathrm{tr}_B\, \rho_{AB} = \rho_A$ and $\mathrm{tr}_A\, \rho_{AB} = \rho_B$, as of course expected. So any density matrix which is a product of the form $\rho_{AB} = \rho_A \otimes \rho_B$ represents *uncorrelated* systems, irrespective of whether the state is pure or not. However, it is very important to note that in general we *cannot* recover the full density matrix ρ_{AB} from the reduced density matrices ρ_A and ρ_B. The operation of taking the partial trace is **irreversible** and in general looses

information. To put that more precisely, given a general ρ_{AB} and its reduced density matrices (3.31), we have

$$\rho_A \otimes \rho_B \neq \rho_{AB} \tag{3.39}$$

This is only true when ρ_{AB} was already originally uncorrelated. Thus, in general, we see that information is lost whenever AB are correlated.

Example: separable states

A state is called separable when it can be written in the form

$$\rho = \sum_i p_i \, \rho_A^i \otimes \rho_B^i, \tag{3.40}$$

for a set of probabilities $p_i \in [0, 1]$, $\sum_i p_i = 1$ and an arbitrary set of density matrices $\rho^i_{A(B)}$. Of course, in light of Eq. (3.26), any density matrix of AB can be decomposed as a sum of products. But usually each term in the sum is not a valid density matrix with a valid probability. The reason why a state of the form (3.40) is physically interesting is because it represents a classical statistical mixture of states of A and B.

This is just like the crappy machine of Sec. 2.1. With some probability p_i the machine prepares a state ρ_A^i for A and ρ_B^i for B. The two systems will in general be correlated: if we learn something about A, we can usually infer something about B. But this is only because they share a **common ignorance** about the machine that produced them. The states of A and B may very well be quantum. But their correlations are purely classical. We say *separable states of the form (3.40) are not entangled.*

Classical states and Quantum Discord

The partial trace is the quantum analog of *marginalizing* a probability distribution. To see that, consider a bipartite state of the form

$$\rho_{AB} = \sum_{i,j} p_{i,j} \, |i\,j\rangle\langle i\,j| \tag{3.41}$$

which will be a valid quantum state provided $p_{i,j} \in [0, 1]$ and $\sum_{i,j} p_{i,j} = 1$. This state is nothing but a *classical* probability distribution encoded in a density matrix. To compute the partial trace over B we use Eq. (3.30), which gives

$$\rho_A = \mathrm{tr}_B \, \rho_{AB} = \sum_{i,j} p_{i,j} \, |i\rangle\langle i|_A = \sum_i p_i^A \, |i\rangle\langle i|_A$$

In the last equality I carried out the sum over j and defined

$$p_i^A = \sum_j p_{i,j}. \tag{3.42}$$

This is precisely the marginalization procedure in classical probability theory. We simply sum over all probabilities of B to obtain a reduced probability distribution only for A. A state which is not of the form (3.41) is said to have some *quantum discord*. We will talk more about what this means soon, when we talk about measurements.

Interactions between physical systems

Now that we know how to describe mathematically composite systems, we can start talking about how to model the interactions between physical systems.

Consider two systems A and B, with Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$. The Hamiltonian of A, which we call H_A, lives only on $\mathcal{H}_A$. Same for B and H_B. Interactions, on the other hand, live on both spaces. They are operators like $A \otimes B$ which act non-trivially on both $\mathcal{H}_A$ and $\mathcal{H}_B$.

A typical Hamiltonian for two systems will then have the form

$$H = H_A + H_B + V \tag{1}$$

where $V \in \mathcal{H}_A \otimes \mathcal{H}_B$ is the interaction.

There are two usual situations we have to usually deal with interactions: dynamics and equilibrium. We saw in lecture 3 that density matrices evolve according to the von Neumann equation

$$\rho(t) = e^{-iHt}\, \rho_0\, e^{iHt} \tag{2}$$

which thus involves the exponential of H. On the other hand, thermal equilibrium reads

$$\rho_{eq} = \frac{e^{-\beta H}}{Z} \tag{3}$$

It also involves the exponential of H. But now with a real coefficient. We say equilibrium is like an evolution in imaginary time:

$$(\text{evolution op}) \quad it \longrightarrow \beta \quad (\text{thermal eq state}) \tag{4}$$

this is also called a wick rotation. Thus, if we understand one, we also understand the other.

what we need to understand is therefore how to deal with $e^{-\beta H}$ when H lives on more than one subspace. Let's try to do this by means of examples.

Spin-spin interactions

Ising interaction:

Let us first consider two spin 1/2 (qubit) particles, each described by Pauli operators σ_α^i, $i = 1,2$, $\alpha = x,y,z$. The simplest spin-spin interaction is the Ising interaction:

$$V = -J\,\sigma_z^1\,\sigma_z^2 \tag{5}$$

This interaction is easy to analyze because it is already diagonal. Let us introduce a new notation that I like. We know that

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Define a variable $\sigma = \pm 1$ and let us denote the eigenvectors as $|\uparrow\rangle = |+1\rangle$ and $|\downarrow\rangle = |-1\rangle$. Then we may summarize

$$\sigma_z|\uparrow\rangle = +1\,|\uparrow\rangle \qquad \sigma_z|\downarrow\rangle = -1\,|\downarrow\rangle$$

into a single pretty equation with many σ's:

$$\sigma_z|\sigma\rangle = \sigma|\sigma\rangle \qquad \sigma = \pm 1. \tag{6}$$

Going back now to Eq. (5), we already know what will be the eigenthings of V:

$$V\,|\sigma_1 \sigma_2\rangle = -J\,\sigma_1\sigma_2\,|\sigma_1\sigma_2\rangle \tag{7}$$

Let us analyze these eigenvalues, $-J\sigma_1\sigma_2$.

$$-J\sigma_1\sigma_2 = \begin{cases} -J & \sigma_1 = \sigma_2 \\ +J & \sigma_1 \neq \sigma_2 \end{cases} \tag{8}$$

The energy is smaller when the spins are parallel. It doesn't matter if both are up or both are down, as long as they are both parallel.

This, in fact, is a very special symmetry of Eq (5), called $\mathbb{Z}_2$ symmetry: if we change

$$\sigma_z^i \longrightarrow -\sigma_z^i \qquad (\mathbb{Z}_2) \tag{9}$$

we keep V intact. As we will learn soon, symmetries play a very important role in statistical mechanics. So it is very useful to keep an eye always out for it.

For instance, suppose we add magnetic fields to the game:

$$H = -h\,(\sigma_z^1 + \sigma_z^2) - J\,\sigma_z^1\sigma_z^2 \tag{10}$$

This Hamiltonian no longer has the $\mathbb{Z}_2$ symmetry. The magnetic field explicitly breaks the symmetry.

Let us now consider the equilibrium state associated to V. For simplicity we set the magnetic fields to zero. Then

$$\rho_{eq} = \frac{e^{-\beta V}}{Z} = \frac{e^{\beta J \sigma_2^1 \sigma_2^2}}{Z} \tag{11}$$

This operator is already diagonal so to compute Z we can use $|\sigma_1, \sigma_2\rangle$ as a convenient basis:

$$Z = \mathrm{tr}\, e^{-\beta V} = \sum_{\sigma_1 \sigma_2} \langle \sigma_1, \sigma_2 | e^{\beta J \sigma_2^1 \sigma_2^2} | \sigma_1, \sigma_2 \rangle$$

$$= \sum_{\sigma_1, \sigma_2} e^{\beta J \sigma_1 \sigma_2} \tag{12}$$

This was step 1: write the trace as a sum of eigenvalues. Step 2 is now to carry out the sum:

$$Z = e^{\beta J} + e^{-\beta J} + e^{-\beta J} + e^{\beta J} = 4 \cosh \beta J \tag{13}$$

The equilibrium state in this case may thus be written as

$$\rho_{eq} = \frac{1}{Z} \mathrm{diag}\left(e^{\beta J}, e^{-\beta J}, e^{-\beta J}, e^{\beta J} \right) \tag{14}$$

where I'm using lexicographic order,

$$|+1,+1\rangle,\ |+1,-1\rangle,\ |-1,+1\rangle,\ |-1,-1\rangle. \tag{15}$$

It is important to notice how the state (14) is not a product state: the Ising interaction makes the two spins correlated.

Let us write (14) as

$$\rho_{eq} = \frac{1}{Z}\left\{ e^{\beta J}|++\rangle\langle++| + e^{-\beta J}|+-\rangle\langle+-| + e^{-\beta J}|-+\rangle\langle-+| + e^{\beta J}|--\rangle\langle--|\right\}$$

where $|++\rangle = |+1,+1\rangle$ is just a shorthand. Then

$$\mathrm{tr}_B \rho_{eq} = \frac{1}{Z}\left\{ (e^{\beta J} + e^{-\beta J})|+\rangle\langle+| + (e^{\beta J} + e^{-\beta J})|-\rangle\langle-|\right\}$$

$$= \frac{1}{2\cosh\beta J}\, 2\cosh\beta J\left(|+\rangle\langle+| + |-\rangle\langle-|\right)$$

or

$$\mathrm{tr}_A \rho_{eq} = \mathrm{tr}_B \rho_{eq} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

The reduced states are maximally mixed! It has no information anymore about β or J. Thus of course, it is not possible to reconstruct ρ_{eq} from ρ_A and ρ_B.

The message you should take from this is that

Interactions correlate!

Exchange interaction:

Let us now consider another type of interaction, of the form

$$V = g\,(\sigma_+^1 \sigma_-^2 + \sigma_-^1 \sigma_+^2) \qquad (16)$$

this is called an exchange interaction because it represents the two qubits exchanging excitations: $\sigma_+^1 \sigma_-^2$ lowers the state of 2 and raises that of 1.

this interaction has a different symmetry than (5). The symmetry is

$$\sigma_+^1 \to \sigma_+^1 e^{i\phi} \qquad U(1) \text{ symmetry} \qquad (17)$$
$$\sigma_+^2 \to \sigma_+^2 e^{i\phi}$$

where ϕ is an arbitrary angle. Unlike in the Ising case, this is a **continuous symmetry**. Please think about **why** this is a symmetry of (16): $\sigma_- = (\sigma_+)^\dagger$ so if we multiply σ_+^1 by $e^{i\phi}$ and σ_-^2 by $e^{-i\phi}$ nothing happens.

The interaction (16) is also not diagonal in the $|\sigma_1 \sigma_2\rangle$ basis. In fact, you may verify that

$$V = g \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad (17)$$

Let us study now the exponential of this guy. Using Mathematica we find

$$e^{V} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cosh(g) & \sinh(g) & 0 \\ 0 & \sinh(g) & \cosh(g) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{18}$$

It is now easy to adapt this to either interactions, e^{-ibV}, or thermal states, $e^{-\beta H}$.

For thermal states we simply replace $g \to -\beta g$:

$$e^{-\beta H} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cosh\beta g & -\sinh\beta g & 0 \\ 0 & -\sinh\beta g & \cosh\beta g & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{19}$$

The partition function is then simply

$$Z = 2 + 2\cosh(\beta g) \tag{20}$$

Notice also how this state is a correlated state. In fact, it is an X state, like the one we studied in page 49 above.

Let's also adapt (18) to the case of interactions. Simply replace $g \to -igt$. We also need

$$\cosh(ix) = \cos x \qquad \sinh(ix) = i\sin x \tag{21}$$

then

$$U(t) := e^{-itV} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(gt) & -i\sin(gt) & 0 \\ 0 & -i\sin(gt) & \cos(gt) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{22}$$

You may check that this U is unitary, $U^\dagger U = UU^\dagger = 1$. This kind of interaction also goes by the name of a **partial SWAP** because, of course, it partially SWAPs the states of the 2 qubits.

The SWAP is complete when $gt = \pi/2$. Then

$$U(\pi/2g) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{23}$$

this SWAP leaves a phase behind. It is a "dirty SWAP". A clean SWAP would look like

$$U_{SWAP} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{24}$$

If you are looking for a nice way to write this SWAP, I will leave for you to check that

$$U_{SWAP} = \frac{1}{2}\left(1 + \vec{\sigma}_1 \cdot \vec{\sigma}_2\right) \tag{25}$$
$$= \frac{1}{2}\left(1 + \sigma_x^1\sigma_x^2 + \sigma_y^1\sigma_y^2 + \sigma_z^1\sigma_z^2\right)$$

this formula is attributed to Pauli.

Heisenberg, XX, XY and XXZ

A generic type of interaction between two spins can be written as

$$V = J_x \sigma_x^1 \sigma_x^2 + J_y \sigma_y^1 \sigma_y^2 + J_z \sigma_z^1 \sigma_z^2 \qquad \text{(XYZ)} \qquad (26)$$

In the literature this would be called XYZ interaction, which is a reference to the fact that the coefficients J_x, J_y, J_z are different.

If $J_x = J_y \neq J_z$ we get instead the XXZ interaction, which is usually written as

$$V = J \left(\sigma_x^1 \sigma_x^2 + \sigma_y^1 \sigma_y^2 + \Delta \sigma_z^1 \sigma_z^2 \right) \qquad \text{(XXZ)} \qquad (27)$$

where Δ is called the assymmetry parameter. If we also have $\Delta = 1$ then we get the XXX Hamiltonian

$$V = J \left(\sigma_x^1 \sigma_x^2 + \sigma_y^1 \sigma_y^2 + \sigma_z^1 \sigma_z^2 \right) = J \, \vec{\sigma}_1 \cdot \vec{\sigma}_2 \qquad (28)$$

No one calls it XXX though, because it sounds like porn. We call it instead the Heisenberg interaction.

I know it may look like (27) and (28) are not too different from each other. But when you have a lattice of spins, the two models have dramatically different physical properties.

Back to (26), when $J_z = 0$ we get the XY model

$$V = J_x \sigma_x^1 \sigma_x^2 + J_y \sigma_y^1 \sigma_y^2 \qquad (XY) \qquad (29)$$

and if $J_x = J_y = J$ we get the XX model

$$V = J (\sigma_x^1 \sigma_x^2 + \sigma_y^1 \sigma_y^2) \qquad (XX) \qquad (30)$$

These models have nice properties. We will come back to them later on. Notice how the XX model has a continuous symmetry that the XY does not. Namely, a rotation symmetry

$$\begin{aligned} \sigma_x^i &\to \cos\phi\, \sigma_x^i - \sin\phi\, \sigma_y^i \\ \sigma_y^i &\to \sin\phi\, \sigma_x^i + \cos\phi\, \sigma_y^i \end{aligned} \qquad (31)$$

The intuitive way to check that this is true is to think about the XX model (30) as the dot product between two vectors in the xy plane

$$V = J (\sigma_x^1, \sigma_y^1, 0) \cdot (\sigma_x^2, \sigma_y^2, 0) \qquad (XX)$$

Thus, if we rotate the two vectors by the same angle ϕ, nothing changes.

Finally, we have another very important model, which is the transverse field Ising model (TFIM):

$$H = -h(\sigma_1^1 + \sigma_1^2) - J\,\sigma_x^1\sigma_x^2 \qquad \text{(TFIM)} \qquad (32)$$

compare this to the usual Ising model in Eqs (5) or (10)

$$H = -h(\sigma_1^1 + \sigma_1^2) - J\,\sigma_1^1\sigma_1^2 \qquad (33)$$

The difference in that in (32) the field is transverse to the interaction. Quantum mechanically this is expressed by the fact that the two terms in (32) do not commute, whereas in (33) they do. In the TFIM this competition between non-commuting terms will be responsible for the appearance of a quantum phase transition, as we will see later on the course.