Quantum Information (ELEC-C9440) Lecture 3-4

Matti Raasakka

Spring 2023

Aalto University

Mixed states, density operator, reduced density operator

Mixed states

Up to this point we have assumed that we know the state vector $|\psi\rangle$ of the system which contains the complete description (as per Postulate 1) of our physical system. Often however, we encounter situations where our knowledge of the state is probabilistic. Consider a couple common situations in our context of quantum information/computation

- A quantum computer is prepared to some initial state $|0\rangle$. During the computation, we perform deliberate actions to manipulate this state in a way of our choosing. However, the manipulation mechanism has finite precision: we end up with a probability distribution of final states.
- The isolation of the qubits and their environment is not perfect: the qubits become entangled with the environment and since we are not observing the environment, our description of the qubits is imprecise.

When we know the state vector precisely, we say that the state is *pure*. If we don't know the state exactly, the state is *mixed* and we use *density operators* in their description.

Suppose a quantum state is in a state $|\psi_i\rangle$ with probability p_i. We call $\{p_i, |\psi_i\rangle\}$ an *ensemble of pure states*. Clearly we must have $\sum_i p_i = 1$. Consider some measurement $\{M_m\}$. We would like a description of this statistical mixture which gives a the probability for the outcome m

$$p(m) = \sum_i p_i p(m|i) = \sum_i p_i \langle\psi_i| M_m^\dagger M_m |\psi_i\rangle \ , \tag{1}$$

so a weighted mixture of probabilities in each state $|\psi_i\rangle$. Similarly, we would like the expectation value of an observable A to be

$$\langle A\rangle = \sum_i p_i \langle\psi_i|A|\psi_i\rangle \ . \tag{2}$$

Such a description is given by *density operators*.

Density operators

A density operator ρ is an operator on the state space which we use to describe a quantum state. It satisfies the following properties:

1. Hermiticity/self-adjointness: $\rho^\dagger = \rho$
2. Unit trace: $\operatorname{tr} \rho = 1$
3. Non-negativity: $\langle\psi|\rho|\psi\rangle \geq 0$ for all $|\psi\rangle$
4. $\rho^2 = \rho$ if state is pure, $\rho^2 \neq \rho$ if state is mixed

A pure state can be written $\rho = |\psi\rangle\langle\psi|$ for some $|\psi\rangle$. Since ρ is Hermitian, its eigenvalues are real. Proof: let $|\psi\rangle$ be an eigenstate of Hermitian A with eigenvalue λ. Note that $\langle\psi|A|\psi\rangle = \lambda \langle\psi|\psi\rangle = \lambda$. Therefore $\lambda^* = \langle\psi|A|\psi\rangle^* = \langle\psi|A^\dagger|\psi\rangle = \langle\psi|A|\psi\rangle = \lambda$, which implies that λ is real. Using the spectral decomposition, we can always write

$$\rho = \sum_i p_i \, |\psi_i\rangle\langle\psi_i| \ . \tag{3}$$

This can be interpreted as a ensemble of quantum states $\{p_i, |\psi_i\rangle\}$. The trace condition guarantees that probabilities sum up to one: $\operatorname{tr} \rho = \sum_i p_i = 1$.

The positivity condition is satisfied if all $p_i \geq 0$. Proof: let $|\psi\rangle$ be an arbitrary pure quantum state. Then
$\langle\psi|\rho|\psi\rangle = \sum_i p_i \langle\psi|\psi_i\rangle \langle\psi_i|\psi\rangle = \sum_i p_i |\langle\psi_i|\psi\rangle|^2 \geq 0$ if each $p_i \geq 0$.
Remember that the trace of a matrix is the sum of the values on its diagonal.
If $\{|\psi_i\rangle\}$ is an orthonormal basis, then $\mathrm{tr}\,\rho = \sum_i \langle\psi_i|\rho|\psi_i\rangle = \sum_i p_i$.
Example: Is the matrix

$$\rho = \begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix} \tag{4}$$

a density matrix? It is clearly Hermitian $\rho^\dagger = \rho$ and its trace is $\mathrm{tr}\,\rho = 1/2 + 1/2 = 1$. It is non-negative if its eigenvalues are non-negative. The characteristic equation gives

$$\left(\frac{1}{2} - \lambda\right)^2 - \frac{1}{4} = 0 \to \lambda = \{0, 1\} \;, \tag{5}$$

so ρ is also positive and therefore a valid density matrix. Furthermore, it is a pure state because it only has one non-zero eigenvalue.

Postulates for density operators

The postulates of quantum mechanics can be reformulated to use density operators. Postulates

1. A physical system is completely described by a density operator acting on the state space of the system.
2. A closed quantum system evolves unitarily

   $$\rho_{t'} = U \rho_t U^\dagger \ , \tag{6}$$

   for two times $t' > t$.
3. A measurement is described by a collection of measurement operators $\{M_m\}$. The probability of outcome m is

   $$p(m) = \mathrm{tr}\left(M_m^\dagger M_m \rho \right) \tag{7}$$

   and the state after the measurement collapses $\rho \mapsto M_m \rho M_m^\dagger / p(m)$. The expectation value of an observable A is $\langle A \rangle = \mathrm{tr}\,(A\rho)$.
4. If a joint system consists of n parts, each prepared to state ρ_i, then their joint system is in the state $\rho_1 \otimes \rho_2 \otimes \ldots \otimes \rho_n$.

Postulates for density operators

Let's check that the measurement rule gives the expected probabilities. The density operator corresponding to the ensemble $\{p_i, |\psi_i\rangle\}$ is $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$. Suppose our measurement operators are $\{M_m\}$. According to the 3rd postulate, the probability of outcome m is

$$p(m) = \operatorname{tr}\left(M_m^\dagger M_m \sum_i p_i |\psi_i\rangle\langle\psi_i|\right) = \sum_i p_i \operatorname{tr}\left(M_m^\dagger M_m |\psi_i\rangle\langle\psi_i|\right) \tag{8}$$

$$= \sum_i p_i \sum_j \langle\psi_j| M_m^\dagger M_m |\psi_i\rangle \langle\psi_i|\psi_j\rangle = \sum_i p_i \sum_j \langle\psi_j| M_m^\dagger M_m |\psi_i\rangle \delta_{ij} \tag{9}$$

$$= \sum_i p_i \langle\psi_i|M_m^\dagger M_m|\psi_i\rangle = \sum_i p_i p(m|i) \ . \tag{10}$$

This is exactly the ensemble average we wanted. How about observable expectation values:

$$\langle A\rangle = \operatorname{tr}(A\rho) = \sum_i p_i \operatorname{tr}(A|\psi_i\rangle\langle\psi_i|) = \sum_i p_i \langle\psi_i|A|\psi_i\rangle \ , \tag{11}$$

again the correct weighted average. Therefore the measurement probabilities work intuitively.

Last week we saw that the Bloch sphere is useful for visualizing one-qubit pure states. Let's generalize this for arbitrary one-qubit mixed states. The density matrix can be written

$$\rho = \begin{pmatrix} a & b \\ b^* & d \end{pmatrix} \ . \tag{12}$$

Because $\rho^\dagger = \rho$, we must have $a, d \in \mathbb{R}$. Because $\operatorname{tr}\rho = a + d = 1$, we define $a = (1 + r_3)/2$ and $d = (1 - r_3)/2$. Also, let $b = (r_1 - ir_2)/2$ for real r_i. Then

$$\rho = \frac{1}{2}\begin{pmatrix} 1 + r_3 & r_1 - ir_2 \\ r_1 + ir_2 & 1 - r_3 \end{pmatrix} = \frac{1}{2}(\mathbb{I} + r_1\sigma_1 + r_2\sigma_2 + r_3\sigma_3) = \frac{1}{2}(\mathbb{I} + \vec{r}\cdot\vec{\sigma}) \ , \tag{13}$$

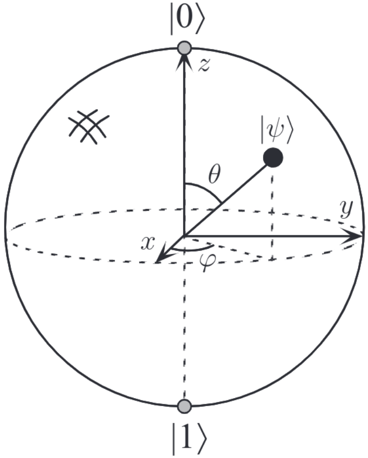
where σ_i are the Pauli matrices. Non-negativity of ρ gives a condition on $|\vec{r}|$. After a straightforward calculation, the eigenvalues of ρ are $(1 \pm |\vec{r}|)/2$. This means that in order for ρ to be positive, we must have $|\vec{r}| \leq 1$. Therefore, any qubit state can be written

$$\rho = \frac{1}{2}(\mathbb{I} + \vec{r}\cdot\vec{\sigma}) \tag{14}$$

and $\vec{r}$ is known as the *Bloch vector*.

Another straightforward computation shows that $\rho^2 = \rho$ if and only if $|\vec{r}| = 1$. Then we can conclude:

- $|\vec{r}| = 1$ corresponds to pure ρ (surface of the Bloch sphere)
- $|\vec{r}| = 0$ corresponds to the completely mixed state $\rho = \mathbb{I}/2$ (origin)



Reminder: Trace of an operator

The trace is a vital operation in quantum information.

Definition (Trace)
The trace $\operatorname{tr} A$ *or a operator* $A : \mathcal{H} \mapsto \mathcal{H}$ *is defined*

$$\operatorname{tr} A = \sum_i \langle i|A|i\rangle = \sum_i A_{ii} \ , \tag{15}$$

where $\{|i\rangle\}$ *is an orthonormal basis for* $\mathcal{H}$.

The trace is independent of the orthonormal basis you use to calculate it. To see this, consider two different orthogonal bases $\{|i\rangle\}$ and $\{|\phi_j\rangle\}$:

$$\operatorname{tr} A = \sum_i \langle i|A|i\rangle = \sum_i \langle i|\, A \left(\sum_j |\phi_j\rangle\langle\phi_j| \right) |i\rangle = \sum_{i,j} \langle \phi_j|i\rangle \, \langle i|\, A \,|\phi_j\rangle \tag{16}$$

$$= \sum_j \langle \phi_j| \left(\sum_i |i\rangle\langle i| \right) A \,|\phi_j\rangle = \sum_j \langle \phi_j|A|\phi_j\rangle \ , \tag{17}$$

where we used the completeness relation $\mathbb{I} = \sum_i |i\rangle\langle i| = \sum_j |\phi_j\rangle\langle\phi_j|$.

Reminder: Trace of an operator

Other basic properties of the trace:

- Linearity: $\operatorname{tr}(zA + wB) = z \operatorname{tr} A + w \operatorname{tr} B$, for $z, w \in \mathbb{C}$
- Cyclic property: $\operatorname{tr}(AB) = \operatorname{tr}(BA)$

This simple relation will be useful when working with density operators. Consider any two unit vectors $|\psi\rangle$ and $|\phi\rangle$. Then the trace of $|\psi\rangle\langle\phi|$ is

$$
\begin{aligned}
\operatorname{tr} |\psi\rangle\langle\phi| &= \sum_i \langle i|\psi\rangle \langle\phi|i\rangle = \sum_i \langle\phi|i\rangle \langle i|\psi\rangle && (18)\\
&= \langle\phi| \left(\sum_i |i\rangle\langle i| \right) |\psi\rangle = \langle\phi| \mathbb{I} |\psi\rangle && (19)\\
&= \langle\phi|\psi\rangle \ . && (20)
\end{aligned}
$$

The cyclic property also implies the useful fact that the trace of an operator is invariant under unitary similarity transformations $A \mapsto UAU^\dagger$ because

$$\operatorname{tr}\left(UAU^\dagger\right) = \operatorname{tr}\left(U^\dagger UA\right) = \operatorname{tr}(\mathbb{I}A) = \operatorname{tr} A \ . \qquad (21)$$

Ensemble interpretation of density operators

It is often useful to think of density matrices as ensembles of quantum states: $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ would correspond to the ensemble where the state $|\psi_i\rangle$ occurs with probability p_i. It is however important to notice that this interpretation is far from unique. In fact there is infinitely many different ensembles that give rise to the same density matrix ρ (NC Theorem 2.6).

Example: You might say that

$$\rho = \frac{3}{4}|0\rangle + \frac{1}{4}|1\rangle\langle 1| \tag{22}$$

describes a system where the state is $|0\rangle$ with probability 3/4 and $|1\rangle$ with probability 1/4. Consider the ensemble

$$|a\rangle = \sqrt{\frac{3}{4}}|0\rangle + \sqrt{\frac{1}{4}}|1\rangle \ , \quad |b\rangle = \sqrt{\frac{3}{4}}|0\rangle - \sqrt{\frac{1}{4}}|1\rangle \ , \tag{23}$$

where both states have probability 1/2. The corresponding ensemble is

$$\rho = \frac{1}{2}|a\rangle\langle a| + \frac{1}{2}|b\rangle\langle b| = \frac{3}{4}|0\rangle\langle 0| + \frac{1}{4}|1\rangle\langle 1| \ , \tag{24}$$

exactly the same as the first ensemble!

We often need to study subsystems of larger quantum systems. An important tool for this is the *reduced density matrix* and the *partial trace*. Suppose we have two physical systems A and B. We denote their joint state by ρ_{AB}. Now say we are interested only on A or only have access to A. The density operator which describes the subsystem A is

$$\rho_A = \mathrm{tr}_B\, \rho_{AB} \ . \tag{25}$$

The operator tr_B is the partial trace and we say that we "trace out/over B". The definition of the partial trace is

$$\mathrm{tr}_B : \mathcal{H}_A \otimes \mathcal{H}_B \mapsto \mathcal{H}_A \tag{26}$$

$$\mathrm{tr}_B(|a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2|) \equiv |a_1\rangle\langle a_2|\, \mathrm{tr}(|b_1\rangle\langle b_2|) \ , \tag{27}$$

for any $|a_1\rangle, |a_2\rangle \in \mathcal{H}_A$ and $|b_1\rangle, |b_2\rangle \in \mathcal{H}_B$. The trace on the RHS is the usual trace in $\mathcal{H}_B$:

$$\mathrm{tr}\left(|b_1\rangle\langle b_2|\right) = \langle b_1|b_2\rangle \ . \tag{28}$$

Reduced density operator

Example: Let $\rho_{AB} = \rho_A \otimes \rho_B$ be the state of a joint system AB. If we trace over B we get the state $\mathrm{tr}_B(\rho_{AB}) = \rho_A$.

Example: Let $|\psi\rangle = (|00\rangle + |11\rangle)/\sqrt{2} \in \mathcal{H}_{AB}$. The density matrix is

$$\rho = |\psi\rangle\langle\psi| = \left(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\right)\left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right) \tag{29}$$

$$= \frac{|00\rangle\langle 00| + |11\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 11|}{2} \; . \tag{30}$$

The reduced density matrix is then

$$\rho_A = \frac{|0\rangle\langle 0|\,\mathrm{tr}\,(|0\rangle\langle 0|) + |1\rangle\langle 0|\,\mathrm{tr}\,(|1\rangle\langle 0|) + |0\rangle\langle 1|\,\mathrm{tr}\,(|0\rangle\langle 1|) + |1\rangle\langle 1|\,\mathrm{tr}\,(|1\rangle\langle 1|)}{2} \tag{31}$$

$$= \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2} = \frac{1}{2}\mathbb{I} \; . \tag{32}$$

It is interesting to note that the joint system is in a definite state $|\psi\rangle$ but still if we look only at A the state is a coin flip between the orthogonal $|0\rangle$ and $|1\rangle$. Later we will see that because ρ_A is proportional to $\mathbb{I}$, the joint qubit state is maximally entangled.

Example: Sometimes it is more convenient to work in component notation, e.g. when calculating on a computer. So let's do the same example with explicit components. We have $|\psi\rangle = \begin{pmatrix} 1 & 0 & 0 & 1 \end{pmatrix}^T / \sqrt{2}$. Then

$$\rho = |\psi\rangle\langle\psi| = \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} . \tag{33}$$

The partial trace is then

$$\rho_A = (\mathbb{I} \otimes \langle 0|)\rho(\mathbb{I} \otimes |0\rangle) + (\mathbb{I} \otimes \langle 1|)\rho(\mathbb{I} \otimes |1\rangle) \tag{34}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \end{pmatrix} \rho \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{35}$$

$$+ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \end{pmatrix} \rho \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{36}$$

Reduced density matrix

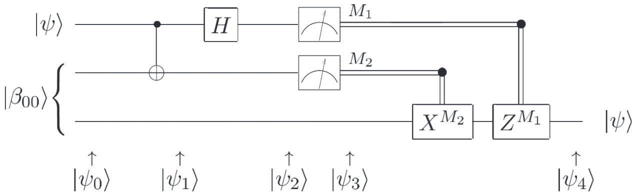
$$\rho_A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \rho \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \rho \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{37}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{2} \mathbb{I} \; . \tag{38}$$

The result is of course the same. Sometimes you'll find it more convenient to work with bras/kets and sometimes with explicit vector components.

Example: quantum teleportation

Quantum teleportation is a protocol for sending an unknown qubit state between two parties only using classical communication and a Bell state.



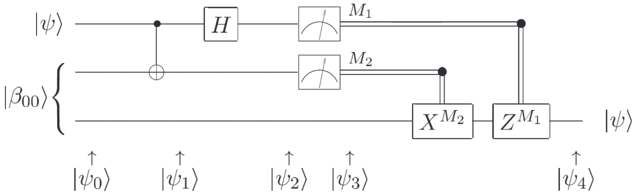
Suppose Alice (first 2 qubits) wants to send a state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ to Bob (last qubit). The protocol is shown in the above Figure. Initially, we have

$$|\psi_0\rangle = |\psi\rangle|\beta_{00}\rangle = \frac{1}{\sqrt{2}}(\alpha|0\rangle(|00\rangle + |11\rangle) + \beta|1\rangle(|00\rangle + |11\rangle)) . \tag{39}$$

After the CNOT on Alice's qubits, the state is

$$|\psi_1\rangle = |\psi\rangle|\beta_{00}\rangle = \frac{1}{\sqrt{2}}(\alpha|0\rangle(|00\rangle + |11\rangle) + \beta|1\rangle(|10\rangle + |01\rangle)) . \tag{40}$$

Example: quantum teleportation



After the Hadamard the state is

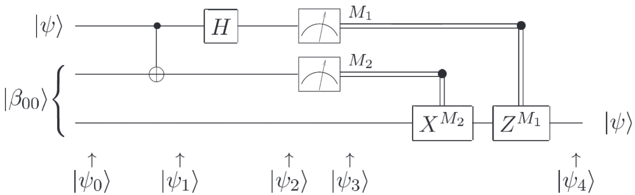
$$|\psi_2\rangle = \frac{1}{2}(\alpha(|0\rangle + |1\rangle)(|00\rangle + |11\rangle) + \beta(|0\rangle - |1\rangle)(|10\rangle + |01\rangle)) \tag{41}$$

$$= \frac{1}{2}\big(|00\rangle (\alpha |0\rangle + \beta |1\rangle) + |01\rangle (\alpha |1\rangle + \beta |0\rangle) \tag{42}$$

$$+ |10\rangle (\alpha |0\rangle - \beta |1\rangle) + |11\rangle (\alpha |1\rangle - \beta |0\rangle)\big) \ . \tag{43}$$

Next Alice measures her qubits in the computational basis. We can see that if the result is 00, then Bob's state is $|\psi\rangle = \alpha |0\rangle + \beta |1\rangle$. Other measurement results cause Bob to have different states, each of which can be transformed to $|\psi\rangle$ if Bob knows Alice's measurement result!

Example: quantum teleportation



Depending on Alice's measurement results, Bob's state collapses to

$$00 \mapsto |\psi_3\rangle = \alpha\,|0\rangle + \beta\,|1\rangle \tag{44}$$

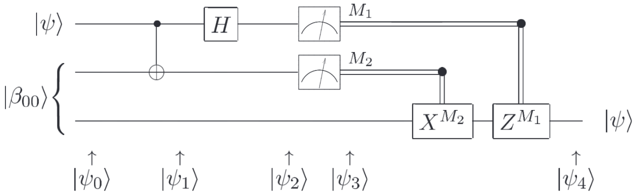
$$01 \mapsto |\psi_3\rangle = \alpha\,|1\rangle + \beta\,|0\rangle \tag{45}$$

$$10 \mapsto |\psi_3\rangle = \alpha\,|0\rangle - \beta\,|1\rangle \tag{46}$$

$$11 \mapsto |\psi_3\rangle = \alpha\,|1\rangle - \beta\,|0\rangle \quad . \tag{47}$$

Finally, Alice sends her measurement result $M_1 M_2$ to Bob, who applies $Z^{M_1} X^{M_2}$ on his qubit. Then his qubit is in state $|\psi_4\rangle = |\psi\rangle$. So, using only a shared Bell state and classical communication, Alice managed to send Bob an unknown qubit state.

Example: quantum teleportation



So what do reduced density matrices have to do with any of this? Using this formalism we can answer the glaring question: it looks like even before Bob learns Alice's measurement results, his qubit carries information about $|\psi\rangle$. Did we just transmit information faster than the speed of light? Luckily, the answer is no.

Consider the joint state before Alice's measurement:

$$\rho = |\psi_2\rangle\langle\psi_2| = \frac{1}{4}\big(|00\rangle\langle 00| (\alpha |0\rangle + \beta |1\rangle)(\alpha^* \langle 0| + \beta^* \langle 1|) \tag{48}$$

$$+ |01\rangle\langle 01| (\alpha |1\rangle + \beta |0\rangle)(\alpha^* \langle 1| + \beta^* \langle 0|) \tag{49}$$

$$+ |10\rangle\langle 10| (\alpha |0\rangle - \beta |1\rangle)(\alpha^* \langle 0| - \beta^* \langle 1|) \tag{50}$$

$$+ |11\rangle\langle 11| (\alpha |1\rangle - \beta |0\rangle)(\alpha^* \langle 1| - \beta^* \langle 0|)\big) . \tag{51}$$

Example: quantum teleportation

To study this situation from Bob's perspective, we compute his reduced density matrix. That is, we trace out Alice's qubits

$$\rho_B = \mathrm{tr}_A\, \rho = \frac{1}{2}\big((|\alpha|^2 + |\beta|^2)\, |0\rangle\langle 0| + (|\alpha|^2 + |\beta|^2)\, |1\rangle\langle 1|\,\big) \tag{52}$$

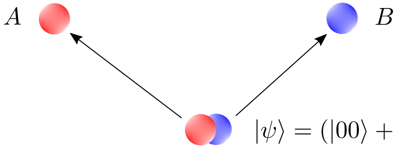
$$= \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) = \mathbb{I}/2\ . \tag{53}$$

So, Bob has the maximally mixed state which doesn't depend in any way on the teleported state $|\psi\rangle$ (α, β do not appear in ρ_B). The quantum information is effectively teleported at the speed of light: only after Bob learns Alice's measurement results can he extract $|\psi\rangle$. Therefore causality is preserved.

Entanglement, entanglement entropy

Entanglement

Entanglement is the fundamental property that distinguishes quantum mechanics from classical theories of physics and its presence is thus responsible for many of the strange quantum properties of nature. Entanglement causes different parts of a system to behave as one in a fundamental way, even when the parts are not interacting with one another. Classical intuition is that one should be able to study the behavior of the non-interacting parts separately and thus understand the behavior of the full system. This intuition fails if the parts of the system are entangled, regardless of the parts not interacting with one another or even being separated by a vast distance.



$|\psi\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$

Entanglement is used as a resource in quantum information. We have already seen this: superdense coding was possible because we used this quantum resource. An another example would be quantum teleportation in which we use pre-shared entanglement and classical communication to teleport a qubit.

Product, separable and entangled states

Which states have entanglement and which do not? The following classification of states is useful.

Definition (Product state)
A product state is a tensor product of two or more density operators.

The state $|\psi\rangle = |00\rangle$, $\rho = |00\rangle\langle 00| = |0\rangle\langle 0| \otimes |0\rangle\langle 0|$ is a product state because $|0\rangle\langle 0|$ is a density operator.

Definition (Separable state)
A state ρ_{AB} is separable if it can be written as an ensemble of product states: $\rho_{AB} = \sum_i p_i \rho_A^{(i)} \otimes \rho_B^{(i)}$.

The state $\rho_{AB} = \frac{1}{2}|0\rangle\langle 0| \otimes |1\rangle\langle 1| + \frac{1}{2}|+\rangle\langle +| \otimes |-\rangle\langle -|$ would be an example of a separable state because each of the factors is a density operator. Note that a product state is also separable.

Definition (Entangled state)
A state ρ_{AB} is entangled if it is not separable.

A prominent example of an entangled state is the Bell state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \to \rho_{AB} = \frac{1}{2}(|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|) \ . \tag{54}$$

There are many different ways to detect and quantify quantum entanglement. One such way is given by the *Schmidt number* which uses the widely useful *Schmidt decomposition*.

Theorem (Schmidt decomposition)
Let $|\psi\rangle \in \mathcal{H}_{AB}$ be a pure state on a bipartite quantum system AB. Then there exist orthonormal bases $\{|i_A\rangle\}$ and $\{|i_B\rangle\}$ for A and B, respectively, such that

$$|\psi\rangle = \sum_i \lambda_i |i_A\rangle |i_B\rangle \ , \tag{55}$$

where $\lambda_i \geq 0$ and $\sum_i \lambda_i^2 = 1$. The coefficients λ_i are the Schmidt coefficients.

One nice thing about this decomposition is that the reduced density matrices

$$\rho_A = \mathrm{tr}_B |\psi\rangle\langle\psi| = \sum_i \lambda_i^2 |i_A\rangle\langle i_A| \ , \quad \rho_B = \sum_i \lambda_i^2 |i_B\rangle\langle i_B| \tag{56}$$

have the same eigenvalues! Therefore any functions on reduced density matrices which depend only on their eigenvalues give the same result for ρ_A and ρ_B. For example, if a bipartite state is pure $|\psi\rangle$, then $\mathrm{tr}\big(\rho_A^2\big) = \mathrm{tr}\big(\rho_B^2\big)$, $\mathrm{tr}(\rho_A \log \rho_A) = \mathrm{tr}(\rho_B \log \rho_B)$ etc.

Let's prove the theorem. Any pure bipartite state can be written

$$|\psi\rangle = \sum_{j,k} a_{jk} |j\rangle |k\rangle = \sum_{i,j,k} u_{ji} d_{ii} v_{i,k} |j\rangle |k\rangle \ , \tag{57}$$

where $\{|j\rangle\}$ and $\{|k\rangle\}$ are orthonormal bases on A and B, respectively. The second step uses the singular value decomposition of the matrix a_{jk}: $a_{jk} = u_{ji} d_{ii} v_{ik}$ for unitary matrices u and v. Then we define $d_{ii} \equiv \lambda_i$, $|i_A\rangle \equiv \sum_j u_{ji} |j\rangle$, and $|i_B\rangle \equiv \sum_k v_{ik} |k\rangle$:

$$|\psi\rangle = \sum_i d_{ii} \left(\sum_j u_{ji} |j\rangle \right) \left(\sum_k v_{ik} |k\rangle \right) = \sum_i \lambda_i |i_A\rangle |i_B\rangle \ , \tag{58}$$

which proves the theorem.

The *Schmidt number* or *Schmidt rank* is the number of non-zero Schmidt coefficients λ_i in the decomposition of a state $|\psi\rangle$. A pure state is a product state if its Schmidt number is exactly one, otherwise it is entangled.
Example: If the Schmidt number is one, then we can write

$$|\psi\rangle = |0_A\rangle\,|0_B\rangle \in \mathcal{H}_{AB}\ , \tag{59}$$

for some states $|0_A\rangle \in \mathcal{H}_A$ and $|0_B\rangle \in \mathcal{H}_B$. This state is clearly a product state and therefore not entangled.
Example: If we have the Schmidt decomposition

$$|\psi\rangle = \frac{1}{\sqrt{2}}\,|0\rangle\,|0\rangle + \frac{1}{\sqrt{2}}\,|1\rangle\,|1\rangle\ \ , \tag{60}$$

the Schmidt number is 2 and the state is entangled.

Entanglement entropy

The Schmidt number gives a very crude characterization of entanglement. A widely used quantity for measuring the amount of entanglement between subsystems of a quantum state is the *von Neumann entropy*. It is often called simply *entanglement entropy*.

Definition (Entanglement entropy)
The entanglement entropy $S(\rho)$ of a quantum state ρ is

$$S(\rho) = -\operatorname{tr}(\rho \log \rho) = -\sum_i p_i \log p_i \ , \tag{61}$$

where $\{p_i\}$ are the eigenvalues of ρ.

We will take the logarithm to be in base 2. A few basic properties of entanglement entropy:

- Non-negative: $S(\rho) \geq 0$, equality iff ρ is pure.
- Invariant under unitary transformations: $S(U\rho U^\dagger) = S(\rho)$.
- $S(\rho) \leq \log d$, where $d = dim(\mathcal{H})$. Equality iff $\rho = \mathbb{I}/d$.
- If ρ_{AB} is pure, then $S(\rho_A) = S(\rho_B)$.
- Subadditivity: $S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B)$ with equality iff $\rho_{AB} = \rho_A \otimes \rho_B$.
- Araki-Lieb inequality: $S(\rho_{AB}) \geq |S(\rho_A) - S(\rho_B)|$.

Let's show that $S(\rho) = -\sum_i p_i \log p_i$. With the spectral decomposition we can write any state $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$. Then

$$\log\rho = \sum_i \log p_i\, |\psi_i\rangle\langle\psi_i| \tag{62}$$

$$\to \rho\log\rho = \sum_{i,j} p_j\, |\psi_j\rangle\langle\psi_j| \log p_i\, |\psi_i\rangle\langle\psi_i| = \sum_{i,j} p_j \log p_i\, |\psi_j\rangle\, \langle\psi_j|\psi_i\rangle\, \langle\psi_i| \tag{63}$$

$$= \sum_{i,j} p_j \log p_i\, |\psi_j\rangle\, \delta_{ji}\, \langle\psi_i| = \sum_i p_i \log p_i\, |\psi_i\rangle\langle\psi_i| \tag{64}$$

$$-\operatorname{tr}(\rho\log\rho) = -\sum_i p_i \log p_i \operatorname{tr}|\psi_i\rangle\langle\psi_i| = -\sum_i p_i \log p_i = S(\rho)\ . \tag{65}$$

The eigenvalue expression shows explicitly some of the listed basic properties: non-negativity because $0 \leq p_i \leq 1$, [1] invariance under unitary transformations (these do not change eigenvalues), $S(\rho) = 0$ for pure states (one $p_i = 1$, others zero).

[1] We always define $0\log 0 \equiv \lim_{x\to 0} x\log x = 0$

Example: Bell state entanglement

Example: Consider again the Bell state $|\psi\rangle = (|00\rangle + |11\rangle)/\sqrt{2} \in \mathcal{H}_{AB}$. We already know this is an entangled state because the Schmidt number is 2. Let's study the entanglement between subsystems A and B by computing $S(\rho_A)$ and $S(\rho_B)$. The reduced density matrix ρ_A is

$$
\begin{aligned}
|\psi\rangle\langle\psi| &= \frac{1}{2}(|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|) && (66)\\
\to \mathrm{tr}_B\,|\psi\rangle\langle\psi| &= \frac{1}{2}(|0\rangle\langle 0|\,\langle 0|0\rangle + |0\rangle\langle 1|\,\langle 0|1\rangle + |1\rangle\langle 0|\,\langle 1|0\rangle + |1\rangle\langle 1|\,\langle 1|1\rangle) && (67)\\
&= \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) = \rho_A\ . && (68)
\end{aligned}
$$

We see that the eigenvalues of ρ_A are $p_1 = 1/2$ and $p_2 = 1/2$. Therefore the entanglement entropy is

$$
S(\rho_A) = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2} = -\frac{1}{2}(-1) - \frac{1}{2}(-1) = 1\ . \qquad (69)
$$

Because the joint state $|\psi\rangle\langle\psi|$ is pure, we know that $S(\rho_A) = S(\rho_B)$. This also follows from the Araki-Lieb inequality and $S(|\psi\rangle\langle\psi|) = 0$. We notice that the subsystems A and B are maximally entangled.

Distance measures for quantum states

How close are two quantum states?

A natural question in the study of quantum information is: *what does it mean that quantum information is preserved in some physical process?*
This question arises in for example in the following situations:

- We store a quantum state in some "quantum memory" to be retrieved and used later. How to determine whether the information has degraded and to what extent?
- Alice and Bob communicate through a noisy quantum channel. How close is the message received by Bob and the message sent by Alice?
- We run a program on a noisy quantum computer. Suppose we apply a gate on some quantum state. What is the "distance" between the actual output state and the theoretically correct output state?

In order to study such questions, we need to develop a *distance measure* which tells us how similar/close two quantum states are to each other. We will introduce the two most important distance measures:

- Trace distance
- Fidelity

Trace distance for classical distributions

We have seen that density operators can be interpreted as ensembles of quantum states: each state occurring with some probability. In order to gain intuition about distance measures, let's first consider classical probability distributions

$$\{p_x\} \quad \text{and} \quad \{q_x\} \ , \tag{70}$$

over the same set x. There is no unique way to describe their similarity but the *trace distance* $D(p_x, q_x)$ is widely used

$$D(p_x, q_x) \equiv \frac{1}{2} \sum_x |p_x - q_x| \ . \tag{71}$$

This is a sensible metric on probability distributions because it is

1. Symmetric: $D(p_x, q_x) = D(q_x, p_x)$
2. Satisfies the triangle inequality: $D(p_x, q_x) \leq D(p_x, r_x) + D(r_x, q_x)$

We also see that it is never negative and zero if and only if $p_x = q_x$, so calling it a "distance" is reasonable.

Trace distance for classical distributions

Example: What is the trace distance between the probability distributions $(1/2, 1/3, 1/6)$ and $(3/4, 1/8, 1/8)$?

$$D(p_x, q_x) = \frac{1}{2}\left(\left|\frac{1}{2} - \frac{3}{4}\right| + \left|\frac{1}{3} - \frac{1}{8}\right| + \left|\frac{1}{6} - \frac{1}{8}\right|\right) \tag{72}$$

$$= \frac{1}{2}\left(\frac{1}{4} + \frac{5}{24} + \frac{1}{24}\right) = \frac{1}{4} \; . \tag{73}$$

Example: What is the trace distance between the probability distributions $(p, 1-p)$ and $(q, 1-q)$?

$$D(p_x, q_x) = \frac{1}{2}(|p-q| + |(1-p) - (1-q)|) \tag{74}$$

$$= \frac{1}{2}(|p-q| + |-p+q|) = |p-q| \; . \tag{75}$$

Fidelity for classical distributions

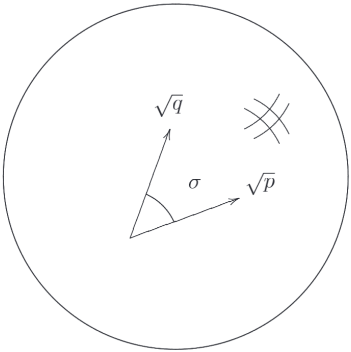
The other useful distance measure in quantum information is the *fidelity*. Let's again first introduce it for classical probability distributions. The fidelity $F(p_x, q_x)$ is

$$F(p_x, q_x) \equiv \sum_x \sqrt{p_x q_x} \ . \tag{76}$$

Note that this has quite different properties compared to trace distance: if $\{p_x\}$ and $\{q_x\}$ are the same, then $F(p_x, p_x) = \sum_x \sqrt{p_x^2} = 1$. Interpretation: $F(p_x, q_x)$ is the inner product between two vectors with components $\sqrt{p_x}$ and $\sqrt{q_x}$.

Example: The fidelity of probability distributions $(1/2, 1/3, 1/6)$ and $(3/4, 1/8, 1/8)$ is

$$F(p_x, q_x) = \sqrt{3/8} + \sqrt{1/24} \tag{77}$$
$$+ \sqrt{1/48} \approx 0.96 \ . \tag{78}$$



$F(p,q) = \sqrt{p}.\sqrt{q} = \cos(\sigma)$

Trace distance for quantum states

We'll now extend the classical trace distance for quantum states. The *trace distance* between two quantum states ρ and σ is defined to be

$$D(\rho,\sigma) \equiv \frac{1}{2}\,\mathrm{tr}\,|\rho-\sigma|\ , \tag{79}$$

where the absolute value of a matrix is defined $|A| \equiv \sqrt{A^\dagger A}$. Lets first see how this reduces to the classical case. Consider a situation where ρ and σ commute $[\rho,\sigma]=0$, which implies that they have the same eigenstates. Then

$$\rho = \sum_i r_i\,|i\rangle\langle i|\ , \quad \sigma = \sum_i s_i\,|i\rangle\langle i| \tag{80}$$

$$\to \rho-\sigma = \sum_i (r_i - s_i)\,|i\rangle\langle i| \tag{81}$$

$$\to (\rho-\sigma)^\dagger(\rho-\sigma) = \sum_i (r_i - s_i)^2\,|i\rangle\langle i| \tag{82}$$

$$\to \sqrt{(\rho-\sigma)^\dagger(\rho-\sigma)} = \sum_i |(r_i - s_i)|\,|i\rangle\langle i| \tag{83}$$

$$\to \frac{1}{2}\,\mathrm{tr}\,\sqrt{(\rho-\sigma)^\dagger(\rho-\sigma)} = \frac{1}{2}\sum_i |(r_i - s_i)| = D(r_i, s_i)\ . \tag{84}$$

Example: computing the trace distance

Example: Find the trace distance between the density operators $\rho = |0\rangle\langle 0|$ and $\sigma = |+\rangle\langle +|$? Let's do this explicitly in component form[2]

$$\rho = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \sigma = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \tag{85}$$

$$\rightarrow \rho - \sigma = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix} \rightarrow (\rho-\sigma)^\dagger(\rho-\sigma) = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{86}$$

$$|\rho - \sigma| = \sqrt{(\rho-\sigma)^\dagger(\rho-\sigma)} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{87}$$

$$\rightarrow D(\rho,\sigma) = \frac{1}{2}\operatorname{tr}|\rho-\sigma| = \frac{1}{\sqrt{2}} \tag{88}$$

[2] Remember that $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$.

Fidelity for quantum states

Let's now define the quantum version of fidelity. The *fidelity* $F(\rho, \sigma)$ is defined

$$F(\rho, \sigma) \equiv \mathrm{tr}\sqrt{\rho^{1/2}\sigma\rho^{1/2}}\ . \tag{89}$$

One can show that this is symmetric $F(\rho, \sigma) = F(\sigma, \rho)$, invariant under unitary transformations $F(U\rho U^\dagger, U\sigma U^\dagger) = F(\rho, \sigma)$, and $0 \leq F(\rho, \sigma) \leq 1$. We'll start by checking that this reduces to the classical case when $\rho = \sum_i r_i |i\rangle\langle i|$ and $\sigma = \sum_i s_i |i\rangle\langle i|$:

$$F(\rho, \sigma) = \mathrm{tr}\sqrt{\sum_i r_i s_i |i\rangle\langle i|} = \mathrm{tr}\left(\sum_i \sqrt{r_i s_i}\, |i\rangle\langle i|\right) \tag{90}$$

$$= \sum_i \sqrt{r_i s_i}\, \mathrm{tr}\, |i\rangle\langle i| = \sum_i \sqrt{r_i s_i} = F(r_i, s_i)\ . \tag{91}$$

We can see from the definition that in general $F(\rho, \rho) = 1$. Another useful relation for $F(\rho, \sigma)$ we can see from the definition is the case where one of the states is pure

$$F(|\psi\rangle\langle\psi|, \rho) = \mathrm{tr}\sqrt{|\psi\rangle\langle\psi|\,\rho\,|\psi\rangle\langle\psi|} \tag{92}$$

$$= \mathrm{tr}\sqrt{\langle\psi|\rho|\psi\rangle\, |\psi\rangle\langle\psi|} = \sqrt{\langle\psi|\rho|\psi\rangle}\ . \tag{93}$$

Example: Find the fidelity between $\rho = \frac{3}{4}|0\rangle\langle 0| + \frac{1}{4}|-\rangle\langle -|$ and $\sigma = |+\rangle\langle +|$.

$$\rho = \frac{1}{8}\begin{pmatrix} 7 & -1 \\ -1 & 1 \end{pmatrix} \tag{94}$$

We could directly use the definition $F(\rho, \sigma) = \text{tr}\sqrt{\rho^{1/2}\sigma\rho^{1/2}}$ but we avoid some work by noticing that σ is pure. Then the fidelity simplifies $F(\rho, \sigma) = \sqrt{\langle +|\rho|+\rangle}$.

$$\langle +|\rho|+\rangle = \frac{1}{16}\begin{pmatrix} 1 & 1 \end{pmatrix}\begin{pmatrix} 7 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{6}{16} \tag{95}$$

$$\rightarrow F(\rho, \sigma) = \frac{1}{2}\sqrt{\frac{3}{2}}\ . \tag{96}$$

Relations between trace distance and fidelity

Similar states have $D(\rho, \sigma) \approx 0$ and $F(\rho, \sigma) \approx 1$. For pure states, these are effectively equivalent. Consider two pure states $|\psi\rangle$ and $|\phi\rangle$. We can always find basis vectors $|a\rangle$ and $|b\rangle$ so that

$$|\psi\rangle = |a\rangle \ , \quad |\phi\rangle = \cos\frac{\theta}{2}|a\rangle + \sin\frac{\theta}{2}|b\rangle \ . \tag{97}$$

Think about the Bloch sphere representation of a pure qubit state to see this. Then

$$D(|\psi\rangle, |\phi\rangle) = \frac{1}{2}\mathrm{tr}\left|\begin{pmatrix} 1-\cos\left(\frac{\theta}{2}\right)^2 & -\sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right) \\ -\sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right) & -\sin\left(\frac{\theta}{2}\right)^2 \end{pmatrix}\right| = \left|\sin\frac{\theta}{2}\right| \tag{98}$$

$$F(|\psi\rangle, |\phi\rangle) = \sqrt{\langle\phi|\psi\rangle\langle\psi|\phi\rangle} = \left|\cos\frac{\theta}{2}\right| \tag{99}$$

Therefore we see that for pure states $D(|\psi\rangle, |\phi\rangle) = \sqrt{1 - F(|\psi\rangle, |\phi\rangle)^2}$. More generally, one can show the following statement about the qualitative similarity for general states ρ and σ

$$1 - F(\rho, \sigma) \le D(\rho, \sigma) \le \sqrt{1 - F(\rho, \sigma)^2} \ . \tag{100}$$