Composite Systems

Composite systems

Here we consider composite quantum systems consisting of more than one distinct physical system.

Postulate: the state space of a composite physical system is the tensor product ($\otimes$) of the component physical spaces. For systems $i = 1,2,\ldots N$, each prepared in state $|\psi_i\rangle$, the joint state of the system is $|\psi_1\rangle\otimes|\psi_2\rangle\otimes\ldots\otimes|\psi_N\rangle$.

For 2×2 matrices, the tensor product is defined as

$$A\otimes B = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \otimes \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{pmatrix}$$

We denote the tensor product simply as

$$|\psi_1\psi_2\ldots\psi_N\rangle \equiv |\psi_1\rangle|\psi_2\rangle\ldots|\psi_N\rangle$$
$$\equiv |\psi_1\rangle\otimes|\psi_2\rangle\otimes\ldots\otimes|\psi_N\rangle$$

For two qubits, in matrix notation

$$|AB\rangle \equiv |A\rangle\otimes|B\rangle = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \otimes \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1b_1 \\ a_1b_2 \\ a_2b_1 \\ a_2b_2 \end{pmatrix}$$

For example, a two-qubit system with the first qubit in state $|0\rangle$ and the second in state $|1\rangle$ is denoted $|01\rangle$.

$$|00\rangle \equiv |0\rangle\otimes|0\rangle = \begin{pmatrix}1\\0\end{pmatrix}\otimes\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}$$

$$|01\rangle \equiv |0\rangle\otimes|1\rangle = \begin{pmatrix}1\\0\end{pmatrix}\otimes\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0\\1\\0\\0\end{pmatrix}$$

$$|10\rangle \equiv |1\rangle\otimes|0\rangle = \begin{pmatrix}0\\1\end{pmatrix}\otimes\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}0\\0\\1\\0\end{pmatrix}$$

$$|11\rangle \equiv |1\rangle\otimes|1\rangle = \begin{pmatrix}0\\1\end{pmatrix}\otimes\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}$$

Note that these form an orthonormal basis.

States and operators

Consider a system composed of two distinguishable parts A and B. Starting from an orthonormal basis $|k\rangle$ ($k = 1,2,\ldots,N_A$) for system A and $|m\rangle$ ($m = 1,2,\ldots,N_B$) for system B, a joint state

$$|\psi\rangle = \sum_{k,m} c_{k,m}\,|km\rangle$$

of the composite system can be written by using combined basis states $|km\rangle$, where $\sum_{k,m}\left|c_{k,m}\right|^2 = 1$. The corresponding dual basis vectors are denoted by $\langle km|$. The Hilbert space dimension of the composite system is $N = N_A N_B$. General operators F can be written as

$$F = \sum_{k,m,\ell,n} F_{k,m,\ell,n}\, |km\rangle\langle \ell n|$$

where $F_{k,m,\ell,n} = \langle km|F|\ell n\rangle$. Operators acting on subsystem F only will be denoted by F_A, and have representation

$$F_A = \sum_{k,\ell,m} F^{(A)}_{k,\ell}\, |km\rangle\langle \ell m|$$

Operators acting on subsystem B will be denoted by F_B, and have representation

$$F_B = \sum_{m,n,k} F^{(B)}_{m,n}\, |km\rangle\langle kn|$$

If we order basis states as $|11\rangle, |12\rangle, \ldots, |1N_B\rangle, |21\rangle, |22\rangle, \ldots$, then convenient block matrix forms for A_A and A_B are

$$F_A = \begin{pmatrix} F^{(A)}_{1,1} I & F^{(A)}_{1,2} I & \cdots \\ F^{(A)}_{2,1} I & F^{(A)}_{2,2} I & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

$$F_B = \begin{pmatrix} F^{(B)} & 0 & \cdots \\ 0 & F^{(B)} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

where I is the $N_B \times N_B$-dimensional identity matrix.

Two coupled qubits

Let's write these equations for two qubits A and B, $N_A = N_B = 2$ and the Hilbert space dimension of the composite system is $N = N_A N_B = 4$. Note that the indexing of qubit states is usually taken to be 0,1 (not 1,2) so $1 \to 0$ and $2 \to 1$ in the indexing. An arbitrary composite state may be written as

$$|\psi\rangle = c_{00}|00\rangle + c_{01}|01\rangle + c_{10}|10\rangle + c_{11}|11\rangle$$

with normalization $|c_{00}|^2 + |c_{01}|^2 + |c_{10}|^2 + |c_{11}|^2$.

An arbitrary operator F can be written as

$$F = F_{0,0,0,0}|00\rangle\langle 00| + F_{0,1,0,0}|01\rangle\langle 00| + F_{1,0,0,0}|10\rangle\langle 00| + F_{1,1,0,0}|11\rangle\langle 00| + \ldots\ldots$$

The composite basis may be written as $|00\rangle, |01\rangle, |10\rangle, |11\rangle$, so the operator F in matrix form is

$$F = \begin{pmatrix} F_{0,0,0,0} & F_{0,0,0,1} & F_{0,0,1,0} & F_{0,0,1,1} \\ F_{0,1,0,0} & F_{0,1,0,1} & F_{0,1,1,0} & F_{0,1,1,1} \\ F_{1,0,0,0} & F_{1,0,0,1} & F_{1,0,1,0} & F_{1,0,1,1} \\ F_{1,1,0,0} & F_{1,1,0,1} & F_{1,1,1,0} & F_{1,1,1,1} \end{pmatrix}$$

and

$$F_A = \begin{pmatrix} F_{0,0}^{(A)} & 0 & F_{0,1}^{(A)} & 0 \\ 0 & F_{0,0}^{(A)} & 0 & F_{0,1}^{(A)} \\ F_{1,0}^{(A)} & 0 & F_{1,1}^{(A)} & 0 \\ 0 & F_{1,0}^{(A)} & 0 & F_{1,1}^{(A)} \end{pmatrix}$$

$$F_B = \begin{pmatrix} F_{0,0}^{(B)} & F_{0,1}^{(B)} & 0 & 0 \\ F_{1,0}^{(B)} & F_{1,1}^{(B)} & 0 & 0 \\ 0 & 0 & F_{0,0}^{(B)} & F_{0,1}^{(B)} \\ 0 & 0 & F_{1,0}^{(B)} & F_{1,1}^{(B)} \end{pmatrix}$$

Entanglement

Separable: the state $|\psi\rangle$ of a composite system is separable when it can be written as a product of states of the individual systems $|\psi\rangle = |AB\rangle$.

Entanglement: the state $|\psi\rangle$ of a composite system is entangled when it can't be written as a product of states of the individual systems $|\psi\rangle \neq |AB\rangle$.

This means that the individual systems are correlated.

For example, consider qubits A and B to be entangled as

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \equiv \frac{1}{\sqrt{2}}(|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B)$$

where we explicitly write individual basis states as $|0\rangle_A$, $|1\rangle_A$ and $|0\rangle_B$, $|1\rangle_B$.

We follow a common scenario: Alice and Bob prepare the composite qubits A and B in the state $|\psi\rangle$. They then separate (perhaps light years apart).

Later, Alice makes a measurement of her qubit by projecting onto the $|0\rangle_A$, $|1\rangle_A$ basis (e.g. using the Z_A operator). With probability $1/2$ the measurement prepares the result $|0\rangle_A|0\rangle_B$, and with probability $1/2$ it prepares $|1\rangle_A|1\rangle_B$. If Alice measured $|0\rangle_A$, then Bob must measure $|0\rangle_B$ in a subsequent measurement of qubit B. Or, if Alice measured $|1\rangle_A$, then Bob must measure $|1\rangle_B$ in a subsequent measurement of qubit B.

When Alice measures her qubit, Bob's qubit immediately jumps to one of two states ($|0\rangle_B$ or $|1\rangle_B$). These states depend on the result of Alice's measurement. Alice's and Bob's measurements are correlated.

When Alice makes her measurement, Bob's wavefunctions instantaneously "collapses" to $|0\rangle_B$ or $|1\rangle_B$. Somebody (e.g. Albert Einstein) might worry that this is an example of "superluminal communication": faster-than-light communication which violates special relativity. In a 1947 letter to Max Born, Einstein wrote

of “spooky action at a distance”. However, the modern view is that although the measurements are correlated, there is no “action” between the measurements or way to send any information instantaneously.

When Bob measures his qubit, he doesn’t know that Alice has already measured and that his wavefunction has collapsed to either $|0\rangle_B$ or $|1\rangle_B$. From his point of view, his measurement will yield either $|0\rangle_B$ or $|1\rangle_B$ with a 50-50 chance.

If Bob had actually measured his qubit first, by projecting onto the $|0\rangle_B$, $|1\rangle_B$ basis (e.g. using the Z_B operator), then with probability $1/2$ the measurement prepares the result $|0\rangle_A|0\rangle_B$, and with probability $1/2$ it prepares $|1\rangle_A|1\rangle_B$. This is exactly the same result as if he measured second.

There is no way for Bob to tell from his measurement who went first, and there is no way of sending any information instantaneously between Alice and Bob. For example, Alice has no way to control the outcome of her measurement, which would be a way to send information: she will always measure either $|0\rangle_A$ or $|1\rangle_A$ with a 50-50 chance.

As a further thought: consider a classical example of correlation. I have a pair of shoes (a left and a right

shoe). I put them into separate sealed boxes which look identical from the outside. I give one box to Alice and one box to Bob. They then separate (by several light years). When Alice opens her box and finds, say, a left shoe then she knows immediately that Bob must have a right shoe. But, at that moment, Bob doesn't know about her measurement and, from his point of view, he still has a 50-50 chance of getting a right or a left shoe.

Bell states

A two-qubit state can be written generically as

$$|\psi\rangle = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle$$

with normalisation $|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 = 1$. For such a state, the degree of entanglement is characterised by the concurrence C,

$$C = 2|\alpha\delta - \beta\gamma|$$

where $0 \leq C \leq 1$.

For separable states, $C = 0$. This means that $\alpha\delta = \beta\gamma$; note that α, δ are the coefficients of $|00\rangle$, $|11\rangle$ whereas β, γ are the coefficients of $|01\rangle$, $|10\rangle$.

For entangled states $C > 0$. States with $C = 1$ are called maximally entangled. Examples of maximally entangled states are the four Bell pair states,

$$|\beta_{00}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

$$|\beta_{01}\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$$

$$|\beta_{10}\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle)$$

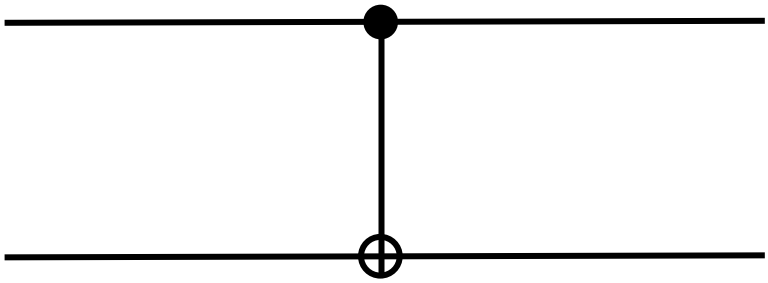
$$|\beta_{11}\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$$

They form an orthonormal basis. To transform from the computational basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ to the

Bell state basis $\{|\beta_{00}\rangle, |\beta_{01}\rangle, |\beta_{10}\rangle, |\beta_{11}\rangle\}$ we have to use two-qubit gates. One of them is the CNOT (conditional NOT) gate C_{AB} with A as the control qubit,

$$C_{AB} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Schematically, this is



If qubit A is in state $|0\rangle$, nothing happens. But if qubit A is in state $|1\rangle$, then qubit B flips. In both cases, qubit A is not changed:

$$C_{AB}|00\rangle = |00\rangle$$
$$C_{AB}|01\rangle = |01\rangle$$
$$C_{AB}|10\rangle = |11\rangle$$
$$C_{AB}|11\rangle = |10\rangle$$

We combine this with a Hadamard gate acting on qubit A only:

$$H_A = H_A \otimes I_B = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix}$$

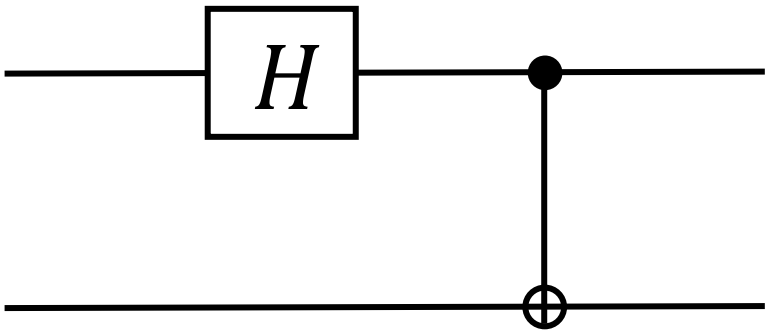
Note, in particular, that

$$H_A = H_A \otimes I_B \neq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & \sqrt{2} \end{pmatrix}$$

Finally, transformation from the computational basis to the Bell state basis is given by the combination $C_{AB}H_A$

$$C_{AB}H_A = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \end{pmatrix}$$

Schematically, this is



This is sometimes called the “Bell circuit”. Note that H_A acts before C_{AB}. With this combination,

$$C_{AB}H_A|00\rangle = |\beta_{00}\rangle$$
$$C_{AB}H_A|01\rangle = |\beta_{01}\rangle$$

$$C_{AB}H_A|10\rangle = |\beta_{10}\rangle$$
$$C_{AB}H_A|11\rangle = |\beta_{11}\rangle$$

Also, $H_A C_{AB} = (C_{AB}H_A)^{-1}$ so that

$$H_A C_{AB}|\beta_{00}\rangle = |00\rangle$$
$$H_A C_{AB}|\beta_{01}\rangle = |01\rangle$$
$$H_A C_{AB}|\beta_{10}\rangle = |10\rangle$$
$$H_A C_{AB}|\beta_{11}\rangle = |11\rangle$$

The combination $H_A C_{AB}$ is sometimes called the "reverse Bell circuit".

Superdense coding

An example of the potential use of entanglement is superdense coding in which the equivalent of two classical bits of information are sent by transferring only one qubit. This is done by (i) preparing a two-qubit (AB) composite system in a Bell state, then separating A and B; (ii) rotating qubit A depending on the values of the classical bits; (iii) sending qubit A to qubit B; (iv) measuring the combined state to determine the values of the classical bits. The two classical bits of information are $M_1 = 0$ or 1 and $M_2 = 0$ or 1.

(i) Preparing a two qubit composite system:

We consider the common situation of two coupled qubits A and B possessed by Alice and Bob. Previously, Alice and Bob prepared the composite system in an entangled Bell state

$$|\beta_{00}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

Alice and Bob then separate, Alice takes qubit A, Bob takes qubit B.

(ii) Rotating qubit A depending on the values of the classical bits

If $M_1 = 1$, Alice acts on qubit A with a NOT gate:

$$(X_A \otimes I_B)|\beta_{00}\rangle = \frac{1}{\sqrt{2}}(|10\rangle + |01\rangle) = |\beta_{01}\rangle$$

where I_B is the identity operator for qubit B. Then, if $M_2 = 1$, Alice acts on qubit A with a phase flip gate:

$$(Z_A \otimes I_B)|\beta_{00}\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |01\rangle) = |\beta_{10}\rangle$$

Or, if the NOT gate was already used ($M_1 = 1$):

$$(Z_A \otimes I_B)|\beta_{01}\rangle = \frac{1}{\sqrt{2}}(-|10\rangle + |01\rangle) = |\beta_{11}\rangle$$

Thus, depending on the values of the two classical bits M_1 and M_2, Alice prepares the coupled qubits in the state $|\beta_{M_1 M_2}\rangle$ where $M_1 = 0$ or 1 and $M_2 = 0$ or 1.

(iii) Sending qubit A to qubit B

Now, Alice transfers the single qubit A to Bob, so that Bob is in possession of both qubits in state $|\beta_{M_1 M_2}\rangle$.

(iv) Measuring the combined state to determine the values of the classical bits

Now, Bob has to measure $|\beta_{M_1 M_2}\rangle$ in order to distinguish the four Bell states and, thus, determine the values of M_1 and M_2. This can be done by acting with the two-qubit gate $H_A C_{AB}$ to take the Bell state basis $|\beta_{M_1 M_2}\rangle$ into the computational basis $|M_1 M_2\rangle$,

$$H_A C_{AB}|\beta_{M_1 M_2}\rangle = |M_1 M_2\rangle$$

Then, a projective measurement can be made in the computational basis.

What does this mean exactly? We consider a measurement of an observable represented by a diagonal matrix in the computational basis, e.g.

$$A = \begin{pmatrix} \lambda_{00} & 0 & 0 & 0 \\ 0 & \lambda_{01} & 0 & 0 \\ 0 & 0 & \lambda_{10} & 0 \\ 0 & 0 & 0 & \lambda_{11} \end{pmatrix}$$

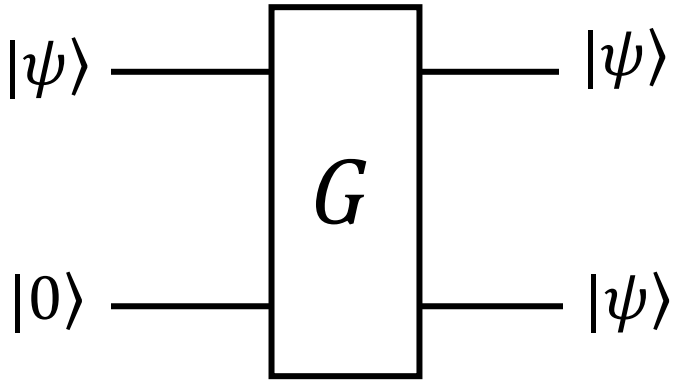
where λ_{ij} are non-degenerate eigenvalues. Then,

$$A = \sum_{i,j=0}^{1} \lambda_{ij} \, \hat{P}_{ij}$$

with projectors $\hat{P}_{ij} = |ij\rangle\langle ij|$. Thus, projective measurement of state $|M_1 M_2\rangle$ with operator A yields the eigenvalue $\lambda_{M_1 M_2}$ with probability 1 (compare the use of the operator Z for a single qubit).

No cloning theorem

We would like to make copies of an arbitrary state $|\psi\rangle$ of a qubit. Imagine we have a two-qubit gate G to do this:



The input is an arbitrary state $|\psi\rangle$ and a fixed second state $|0\rangle$ (generally called an "ancilla" qubit). The output is two copies of $|\psi\rangle$. If it exists, the gate should act as

$$G|\psi\rangle|0\rangle = |\psi\rangle|\psi\rangle$$

In particular:

$$G|0\rangle|0\rangle = |0\rangle|0\rangle \quad (1)$$

$$G|1\rangle|0\rangle = |1\rangle|1\rangle \quad (2)$$

$$G\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|0\rangle = \frac{1}{2}(|0\rangle + |1\rangle)(|0\rangle + |1\rangle)$$

$$= \frac{1}{2}(|0\rangle|0\rangle + |0\rangle|1\rangle + |1\rangle|0\rangle + |1\rangle|1\rangle) \quad (3)$$

But statements (1), (2), (3) contradict each other if we insist that G is a linear operator (as it should be):

$$G\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|0\rangle = G|0\rangle|0\rangle + G|1\rangle|0\rangle$$

Using (1) and (2):

$$G\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|0\rangle = |0\rangle|0\rangle + |1\rangle|1\rangle$$

But this is not the same as (3): a gate such as G can't exist. It is impossible to construct a gate that clones general qubits.

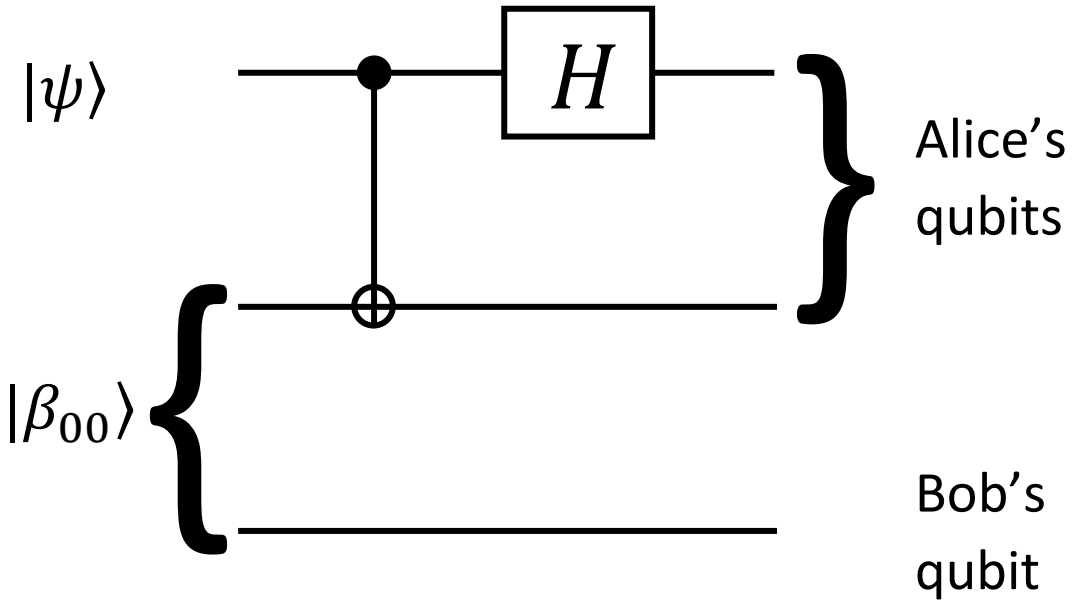
Quantum teleportation

Quantum teleportation is a technique for moving quantum states around, even in the absence of a quantum communications channel linking the sender of the quantum state to the recipient. Alice and Bob met long ago but now live far apart. While together they generated a Bell state $|\beta_{00}\rangle$, each taking one qubit of the Bell state when they separated. Now, Alice is tasked with delivering a qubit state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ to Bob. She does not know the values of α and β (although $|\psi\rangle$ is normalised as $|\alpha|^2 + |\beta|^2 = 1$), and she can only send classical information to Bob.

Quantum teleportation is a way of using the entangled Bell state in order to send $|\psi\rangle$ to Bob, with only a small overhead of classical communication.

In outline, the steps are as follows: (i) Alice interacts the qubit $|\psi\rangle$ with her half of the Bell state, and then (ii) Alice measures the two qubits in her possession, obtaining one of four possible classical results, 00, 01, 10, and 11. She sends this information to Bob. (iii) depending on Alice's classical message, Bob performs one of four operations on his half of the Bell state. By doing this he can recover the original state $|\psi\rangle$.

(i) Alice interacts the qubit $|\psi\rangle$ with her half of the Bell state



Bob has the second qubit of the Bell state

$$|\beta_{00}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

Alice has the first qubit of the Bell state, plus the unknown state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$. As shown in the circuit diagram, she interacts $|\psi\rangle$ with her part of the Bell state using the reverse Bell circuit of a CNOT gate with $|\psi\rangle$ as control followed by a Hadamard gate on the first qubit.

Let us write the initial state of the three qubits as

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

$$= \frac{1}{\sqrt{2}}(\alpha|000\rangle + \alpha|011\rangle + \beta|100\rangle + \beta|111\rangle)$$

The action of the CNOT gate is

$$C_{12}|00\rangle = |00\rangle$$
$$C_{12}|01\rangle = |01\rangle$$
$$C_{12}|10\rangle = |11\rangle$$
$$C_{12}|11\rangle = |10\rangle$$

so that

$$C_{12}|\psi\rangle\otimes|\beta_{00}\rangle$$
$$= \frac{1}{\sqrt{2}}(\alpha|000\rangle + \alpha|011\rangle + \beta|110\rangle + \beta|101\rangle)$$

The action of the Hadamard gate is

$$H_1|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$
$$H_1|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

so that

$$H_1 C_{12}|\psi\rangle\otimes|\beta_{00}\rangle$$
$$= \frac{1}{2}(\alpha|000\rangle + \alpha|100\rangle + \alpha|011\rangle + \alpha|111\rangle$$
$$+ \beta|010\rangle - \beta|110\rangle + \beta|001\rangle - \beta|101\rangle)$$
$$= \frac{1}{2}|00\rangle(\alpha|0\rangle + \beta|1\rangle) + \frac{1}{2}|01\rangle(\alpha|1\rangle + \beta|0\rangle)$$
$$+ \frac{1}{2}|10\rangle(\alpha|0\rangle - \beta|1\rangle) + \frac{1}{2}|11\rangle(\alpha|1\rangle - \beta|0\rangle)$$

(ii) Alice measures the two qubits in her possession

Alice now measures her two qubits in the computational basis, e.g. with operator

$$A = \sum_{i,j=0}^{1} \lambda_{ij}\, \hat{P}_{ij}$$

with non-degenerate eigenvalues λ_{ij} and projectors $\hat{P}_{ij} = |ij\rangle\langle ij|$. In the full three-qubit space, these should be written as

$$\hat{P}_{ij} = (|ij\rangle \otimes I_3)(\langle ij| \otimes I_3)$$

where I_3 is the identity operator acting on the third qubit.

For example, the probability associated with λ_{00} is

$$\begin{aligned} P(\lambda_{00}) &= |\langle 00| \otimes I_3 H_1 C_{12} |\psi\rangle \otimes |\beta_{00}\rangle|^2 \\ &= \left| \frac{1}{2} (\alpha|0\rangle + \beta|1\rangle) \right|^2 \\ &= \frac{1}{4} (|\alpha|^2 + |\beta|^2) \\ &= \frac{1}{4} \end{aligned}$$

using $|\alpha|^2 + |\beta|^2 = 1$. And the resulting state is

$$\begin{aligned} |\psi'_{00}\rangle &= \frac{1}{\sqrt{P(\lambda_{00})}} \hat{P}_n |\psi\rangle \\ &= 2(|00\rangle \otimes I_3)\langle 00| \otimes I_3 H_1 C_{12} |\psi\rangle \otimes |\beta_{00}\rangle \end{aligned}$$

$$= 2(|00\rangle \otimes I_3)\frac{1}{2}(\alpha|0\rangle + \beta|1\rangle)$$
$$= |00\rangle \otimes (\alpha|0\rangle + \beta|1\rangle)$$

There are four possible outcomes to Alice's measurement, $\lambda_{00}, \lambda_{01}, \lambda_{10}, \lambda_{11}$, and each of them occurs with probability ¼. Afterwards, Bob's qubit is left in different states:

$$\lambda_{00} \Leftrightarrow \alpha|0\rangle + \beta|1\rangle$$
$$\lambda_{01} \Leftrightarrow \alpha|1\rangle + \beta|0\rangle$$
$$\lambda_{10} \Leftrightarrow \alpha|0\rangle - \beta|1\rangle$$
$$\lambda_{11} \Leftrightarrow \alpha|1\rangle - \beta|0\rangle$$

Alice needs to let Bob know which situation he is in. So she sends him two classical bits of information 00, 01, 10 or 11.

(iii) depending on Alice's classical message, Bob performs one of four operations on his half of the Bell state

If Bob receives 00, he knows that his qubit is in the correct form $\alpha|0\rangle + \beta|1\rangle$ and so he does nothing.

If Bob receives 01, he knows that his qubit is $\alpha|1\rangle + \beta|0\rangle$. He applies the gate X to it.

If Bob receives 10, he knows that his qubit is $\alpha|0\rangle - \beta|1\rangle$. He applies the gate Z to it.

If Bob receives 11, he knows that his qubit is $\alpha|1\rangle - \beta|0\rangle$. He applies the gate ZX to it.

Let's check this explicitly:

$$\begin{aligned} ZX(\alpha|1\rangle - \beta|0\rangle) &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} -\beta \\ \alpha \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} -\beta \\ \alpha \end{pmatrix} \\ &= \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \\ &= \alpha|0\rangle + \beta|1\rangle \end{aligned}$$

In every case Bob's qubit ends in state $\alpha|0\rangle + \beta|1\rangle$, the original state of the qubit that Alice wanted to teleport.

It is important to note that there is only one qubit in state $\alpha|0\rangle + \beta|1\rangle$ at any point during the process. Initially, Alice has it. At the end Bob has it, but as the no cloning theorem tells us, we can't copy, so only one of them can have it at a time.

When Alice measures her qubits, Bob's qubit instantaneously jumps to one of the four states. But he has to wait for Alice to send him the two classical bits before he can determine which of the four qubits correspond to Alice's original qubit. It is the fact that the two bits have to be sent by some conventional transportation method that prevents instantaneous transmission of information.

EPR and the Bell inequality

In the Copenhagen interpretation, principally attributed to Niels Bohr and Werner Heisenberg, observables generally have no definite value until a measurement is made. An example of this is Heisenberg's uncertainty principle in which a measurement of a particle's momentum, say, limits the ability to know its position.

Albert Einstein, in particular, took issue with the probabilistic nature of quantum mechanics and the role of measurement. In a 1926 letter to Max Born, Einstein wrote *"God does not play dice"* and in a 1947 letter to Max Born, Einstein wrote of *"spooky action at a distance"* (referring to entanglement).

EPR

"Can quantum-mechanical description of physical reality be considered complete?"

Albert Einstein, Boris Podolsky, and Nathan Rosen, Physical Review **47**, 777 (1935).

Published in 1935, the Einstein-Podolsky-Rosen (EPR) paper questioned whether quantum mechanics is a complete theory.

From their 1935 paper:

"Is the description given by the theory complete? ...every element of the physical reality must have a counterpart in the physical theory…If, without in any way disturbing a system, we can predict with certainty (i.e. with probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity."

For example, in quantum mechanics *"when the momentum of a particle is known, its coordinate has no physical reality."*

"More generally, it is shown in quantum mechanics that, if the operators corresponding to two physical quantities, say A and B, do not commute, that is, if $AB \neq BA$, then the precise knowledge of one of them precludes such a knowledge of the other. Furthermore, any attempt to determine the latter experimentally will alter the state of the system in such a way as to destroy the knowledge of the first."

Hence *"either (1) the quantum-mechanical description of reality given by the wave function is not complete or (2) when the operators corresponding to two physical quantities do not commute the two quantities cannot have simultaneous reality."*

EPR uses the example of entanglement of two particles with position and momentum – this discussion can easily be translated into one about two qubits and spin measurements in different directions (see, e.g. D. Bohm, *Quantum Theory*, Prentice-Hall, Englewood Cliffs, NJ (1951); D. Bohm and Y. Aharonov, Phys. Rev. **108**, 1070 (1957)).

EPR believed in local realism: the properties of qubit B, say, should be independent of what happens to qubit A which is far away.

- local: no *"spooky action at a distance"*
- realism: *"God does not play dice"*

EPR considered that quantum mechanics is incomplete. Local realism implies a more complete specification of a state (as compared to quantum mechanics): there must be "hidden variables", precisely definable variables determining the actual behaviour of a particle (as in classical physics), and not merely its probable behaviour.

Bell inequality

"On the Einstein Podolsky Rosen Paradox"

John Bell, Physics **1**, 195 (1964).

Bell developed a test of hidden variable theory, based on the model of a two-qubit system. The key idea is to consider measurements of spin in different directions. The test is in the form of an inequality that should be obeyed by hidden variable theories. Following Bell's paper, several versions of the inequality were developed. Here we'll describe the commonly-used CHSH inequality.

"Proposed Experiment to Test Local Hidden-Variable Theories"

John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt, Phys. Rev. Lett. **23**, 880 (1969).

Previously, we saw that an operator which measures spin along the direction of unit vector $\hat{n}$ is

$$\sigma_{\hat{n}} = \hat{n} \cdot \vec{\sigma}$$

The outcomes of a measurement (the eigenvalues) are always $+1$ or -1, independent of $\hat{n}$.

Now consider a system of two qubits, A and B. On each qubit, we can choose to carry out one of two different experiments which are measurements of the spin in different directions, corresponding to observables

denoted $\hat{A}_1$, $\hat{A}_2$ for qubit A and $\hat{B}_1$, $\hat{B}_2$ for qubit B. Each measurement will have the value $+1$ or -1. Now we consider an observable defined as

$$\hat{F} = (\hat{A}_1 + \hat{A}_2)\hat{B}_1 - (\hat{A}_1 - \hat{A}_2)\hat{B}_2$$

We will now consider the possible values that the expectation value $\langle F \rangle$ could take.

Hidden variables: for each measurement, the system has well defined values for the outcomes a_1, a_2 and b_1, b_2, namely $+1$ or -1:

$$F_{hv} = (a_1 + a_2)b_1 - (a_1 - a_2)b_2$$

For each measurement, we can have $a_1 = \pm 1$ and $a_2 = \pm 1$:

a_1	a_2	$a_1 + a_2$	$a_1 - a_2$
$+1$	$+1$	2	0
$+1$	-1	0	2
-1	$+1$	0	-2
-1	-1	-2	0

These can be reduced to either of:

- $a_1 + a_2 = 0 \Rightarrow a_1 - a_2 = \pm 2 \Rightarrow F_{hv} = \pm 2b_2$
- $a_1 - a_2 = 0 \Rightarrow a_1 + a_2 = \pm 2 \Rightarrow F_{hv} = \pm 2b_1$

But we can also only have $b_1 = \pm 1$ and $b_2 = \pm 1$, so each individual measurement corresponds to either $F_{hv} = +2$ or $F_{hv} = -2$. On averaging over many experiments, the expectation value should satisfy

$$-2 \leq \langle F_{hv} \rangle \leq 2$$

This is the CHSH inequality.

Quantum mechanics: owing to correlations arising from entanglement, it is possible to violate the CHSH inequality. Let's consider a particular example by choosing

$$\hat{A}_1 = Z_A, \qquad \hat{A}_2 = X_A$$
$$\hat{B}_1 = -\frac{1}{\sqrt{2}}(Z_B + X_B), \qquad \hat{B}_2 = \frac{1}{\sqrt{2}}(Z_B - X_B)$$

Hence

$$\begin{aligned}\hat{F} &= -\frac{1}{\sqrt{2}}(Z_A + X_A)(Z_B + X_B) \\ &\quad -\frac{1}{\sqrt{2}}(Z_A - X_A)(Z_B - X_B) \\ &= -\sqrt{2}(X_A X_B + Z_A Z_B)\end{aligned}$$

We can use

$$F_A = \begin{pmatrix} F_{0,0}^{(A)} & 0 & F_{0,1}^{(A)} & 0 \\ 0 & F_{0,0}^{(A)} & 0 & F_{0,1}^{(A)} \\ F_{1,0}^{(A)} & 0 & F_{1,1}^{(A)} & 0 \\ 0 & F_{1,0}^{(A)} & 0 & F_{1,1}^{(A)} \end{pmatrix}$$

$$F_B = \begin{pmatrix} F_{0,0}^{(B)} & F_{0,1}^{(B)} & 0 & 0 \\ F_{1,0}^{(B)} & F_{1,1}^{(B)} & 0 & 0 \\ 0 & 0 & F_{0,0}^{(B)} & F_{0,1}^{(B)} \\ 0 & 0 & F_{1,0}^{(B)} & F_{1,1}^{(B)} \end{pmatrix}$$

to write

$$X_A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$X_B = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$X_A X_B = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

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

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

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

$$\hat{F} = -\sqrt{2}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$

Let's assume the system is prepared in the Bell state

$$|\beta_{11}\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}$$

so the expectation value is

$$\langle F \rangle$$

$$= -\frac{\sqrt{2}}{2}(0 \quad 1 \quad -1 \quad 0)\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}$$

$$= -\frac{\sqrt{2}}{2}(0 \quad 1 \quad -1 \quad 0)\begin{pmatrix} 0 \\ -2 \\ 2 \\ 0 \end{pmatrix}$$

$$= 2\sqrt{2}$$

So $\langle F \rangle > 2$, and this violates the CHSH inequality.

Experiments: do experiments satisfy the CHSH inequality (in agreement with hidden variable theory) or violate it (in agreement with quantum mechanics)?

Experiments are usually performed with the polarisation of photons instead of the spin of electrons. A series of experiments have been performed to improve their compliance with necessary conditions (e.g. spatial separation of A and B) and to reduce statistical uncertainty. They have generally shown violation of the CHSH inequality and agreement with the predictions of quantum mechanics. See, for example:

- S. J. Freedman and J. F. Clauser, Phys. Rev. Lett. **28**, 938 (1972).
- Aspect, J. Dalibard and G. Roger, Phys. Rev. Lett. **49**, 1804 (1982).
- G. Weihs, T. Jennewein, C. Simon, H. Weinfurter and A. Zeilinger, Phys. Rev. Lett. **81**, 5039 (1998).
- "Scientific Background on the Nobel Prize in Physics 2022" https://www.nobelprize.org/prizes/physics/2022/advanced-information/

The Nobel Prize in Physics 2022 was awarded jointly to Alain Aspect, John F. Clauser and Anton Zeilinger "for

experiments with entangled photons, establishing the violation of Bell inequalities and pioneering quantum information science"

Density operator

We consider an ensemble which is a collection of physically identical quantum systems, which however can be described by different states. Generally, we know the probability that a certain state arises within the ensemble but, because this is only known to the level of probability, there is incomplete knowledge of the state of the whole system.

An example might be: you have a friend who prepares a certain quantum state $|\psi_0\rangle$. They then operate on the state to create $|\psi_1\rangle$, but only with probability P. Then, you have a mix of state $|\psi_0\rangle$ with probability $1 - P$ and state $|\psi_1\rangle$ with probability P.

This is a classical probability, so we have to consider two types of randomness: (i) statistical interpretation of the wave function (quantum mechanics) (ii) incomplete knowledge of the state of the system (statistical mechanics). This leads to the topic of quantum statistical mechanics.

Pure: all the states in the ensemble are identical. The ensemble can be described with a single state.

Mixed: all the states in the ensemble are not identical. The ensemble can't be described with a single state.

Density operator

Imagine that the total system is an ensemble of N sub-systems, labelled $\alpha = 1,2, \ldots N$, each of which is in a state described by a normalised state vector $|\alpha\rangle$ where $\langle\alpha|\alpha\rangle = 1$. We use the statistical weight P_α of the state$|\alpha\rangle$: i.e. P_α is the probability of finding the system in state $|\alpha\rangle$. We assume that

$$0 \leq P_\alpha \leq 1\,; \qquad \sum_{\alpha=1}^{N} P_\alpha = 1$$

In the ensemble, expectation values are defined by

$$\langle A\rangle = \sum_{\alpha=1}^{N} P_\alpha \langle A\rangle_\alpha$$

where $\langle A\rangle_\alpha = \langle\alpha|\hat{A}|\alpha\rangle$ is the expectation value of an operator $\hat{A}$ in the state $|\alpha\rangle$.

If we define the density operator $\hat{\rho}$ as

$$\hat{\rho} = \sum_{\alpha=1}^{N} P_\alpha |\alpha\rangle\langle\alpha|$$

then the expectation value can be conveniently written as

$$\langle A\rangle = \mathrm{tr}\left(\hat{\rho}\hat{A}\right)$$

where $\mathrm{tr}\left(\hat{B}\right)$ denotes the trace of an operator $\hat{B}$, which in any given orthonormal basis $|n\rangle$ can be calculated as

$\mathrm{tr}(\hat{B}) = \sum_n \langle n|\hat{B}|n\rangle = \sum_n B_{nn}$. The density operator leads us to the ensemble average of the operator $\hat{A}$. Note that the names "density operator" and "density matrix" are often used synonymously.

Some properties of the density operator are:

- the trace of the density operator is one: $\mathrm{tr}(\hat{\rho}) = 1$
- the density operator is positive-definite: $\langle\varphi|\hat{\rho}|\varphi\rangle \geq 0$.
- the density matrix is Hermitian: $(\hat{\rho}_{n'n})^* = \hat{\rho}_{nn'}$
- diagonal matrix elements give the probability to find the system in the state $|n\rangle$: $\hat{\rho}_{nn} = \sum_{\alpha=1}^{N} P_\alpha |\langle n|\alpha\rangle|^2$ such that $0 \leq \hat{\rho}_{nn} \leq 1$.
- when the system is in a pure state $|\beta\rangle$, $P_\alpha = \delta_{\alpha,\beta}$ for all α, $\hat{\rho}$ is a projection operator $\hat{\rho} = |\beta\rangle\langle\beta|$ with the idempotent property, $\hat{\rho}^2 = \hat{\rho}$.
- when the system is in a mixed state, $\hat{\rho}^2 \neq \hat{\rho}$.

Mathematical details

Now let's add some mathematical details behind these results. We can choose a complete set of orthonormal basis vectors [e.g. the eigenstates of some operator] $|n\rangle$ where

$$\langle n'|n\rangle = \delta_{n'n}$$

$$\sum_n |n\rangle\langle n| = \hat{I}$$

We can expand the pure state $|\alpha\rangle$ in the basis $|n\rangle$:

$$|\alpha\rangle = \sum_n c_n^{(\alpha)} |n\rangle$$

where the expansion coefficient is $c_n^{(\alpha)} = \langle n|\alpha\rangle$. The normalisation of $|\alpha\rangle$ leads to

$$\sum_n \left|c_n^{(\alpha)}\right|^2 = 1$$

$$\text{LHS} = \sum_n \left|c_n^{(\alpha)}\right|^2 = \sum_n \langle\alpha|n\rangle\langle n|\alpha\rangle = \langle\alpha|\alpha\rangle = 1$$

The expectation value of an operator $\hat{A}$ in the state $|\alpha\rangle$ is

$$\begin{aligned}\langle A\rangle_\alpha = \langle\alpha|\hat{A}|\alpha\rangle &= \sum_{n,n\prime} \left(c_{n\prime}^{(\alpha)}\right)^* c_n^{(\alpha)} \langle n'|\hat{A}|n\rangle \\ &= \sum_{n,n\prime} \langle n|\alpha\rangle\,\langle\alpha|n'\rangle\langle n'|\hat{A}|n\rangle\end{aligned}$$

Now consider the ensemble average:

$$\langle A\rangle = \sum_{\alpha=1}^{N} P_\alpha \langle A\rangle_\alpha$$

where P_α is the statistical weight of the pure state $|\alpha\rangle$. Using the expression for $\langle A\rangle_\alpha$ above gives

$$\langle A\rangle = \sum_{\alpha n,n\prime} P_\alpha \left(c_{n\prime}^{(\alpha)}\right)^* c_n^{(\alpha)} \langle n'|\hat{A}|n\rangle$$
$$= \sum_{\alpha,n,n\prime} \langle n|\alpha\rangle\, P_\alpha \langle\alpha|n'\rangle\langle n'|\hat{A}|n\rangle$$

Now we simplify this expression using the density operator:

$$\hat{\rho} = \sum_{\alpha=1}^{N} P_\alpha |\alpha\rangle\,\langle\alpha|$$

Then we have

$$\langle A\rangle = \sum_{\alpha,n,n\prime} \langle n|\alpha\rangle\, P_\alpha \langle\alpha|n'\rangle\langle n'|\hat{A}|n\rangle$$
$$= \sum_{n,n\prime} \langle n|\,\hat{\rho}|n'\rangle\langle n'|\hat{A}|n\rangle$$
$$= \sum_{n} \langle n|\,\hat{\rho}\hat{A}|n\rangle$$
$$= \sum_{n} \left(\hat{\rho}\hat{A}\right)_{nn}$$
$$= \mathrm{tr}\left(\hat{\rho}\hat{A}\right)$$

A possible choice for $\hat{A}$ is the unit operator, giving:

$$\langle I\rangle = \sum_{\alpha=1}^{N} P_\alpha \langle I\rangle_\alpha = \sum_{\alpha=1}^{N} P_\alpha \langle\alpha|I|\alpha\rangle = \sum_{\alpha=1}^{N} P_\alpha = 1$$

But $\langle A\rangle = \mathrm{tr}\left(\hat{\rho}\hat{A}\right)$, so

$$\langle I \rangle = \mathrm{tr}(\hat{\rho})$$
$$\Rightarrow \quad \mathrm{tr}(\hat{\rho}) = 1$$

The density matrix is Hermitian:

$$\hat{\rho}_{nn'} = \langle n|\rho|n'\rangle = \sum_{\alpha=1}^{N} \langle n|\alpha\rangle \, P_\alpha \langle \alpha|n'\rangle$$

$$\hat{\rho}_{n'n} = \langle n'|\rho|n\rangle = \sum_{\alpha=1}^{N} \langle n'|\alpha\rangle \, P_\alpha \langle \alpha|n\rangle$$

$$(\hat{\rho}_{n'n})^* = \sum_{\alpha=1}^{N} (\langle n'|\alpha\rangle)^* \, P_\alpha^* (\langle \alpha|n\rangle)^*$$

$$= \sum_{\alpha=1}^{N} \langle \alpha|n'\rangle \, P_\alpha \langle n|\alpha\rangle$$

$$= \sum_{\alpha=1}^{N} \langle n|\alpha\rangle \, P_\alpha \langle \alpha|n'\rangle = \hat{\rho}_{nn'}$$

Diagonal matrix elements give the probability to find the system in the state $|n\rangle$:

$$\hat{\rho}_{nn} = \langle n|\hat{\rho}|n\rangle = \sum_{\alpha=1}^{N} \langle n|\alpha\rangle \, P_\alpha \langle \alpha|n\rangle = \sum_{\alpha=1}^{N} P_\alpha \left| c_n^{(\alpha)} \right|^2$$

Each term in the summation is the probability to be in ensemble $|\alpha\rangle$ multiplied by probability that $|\alpha\rangle$ is in state $|n\rangle$.

Since $P_\alpha \geq 0$ and

$$\mathrm{tr}(\hat{\rho}) = \sum_n \hat{\rho}_{nn} = 1$$

we have $0 \leq \hat{\rho}_{nn} \leq 1$.

If the system is in a pure state $|\beta\rangle$ then we have $P_\alpha = \delta_{\alpha,\beta}$ for all α. This gives

$$\hat{\rho} = \sum_{\alpha=1}^{N} |\alpha\rangle\, \delta_{\alpha,\beta} \langle\alpha| = |\beta\rangle\langle\beta|$$

Then, the density operator is a projection operator $\hat{\rho} = |\beta\rangle\langle\beta|$ with the idempotent property, $\hat{\rho}^2 = \hat{\rho}$:

$$\hat{\rho}^2 = |\beta\rangle\langle\beta|\beta\rangle\langle\beta| = |\beta\rangle\langle\beta| = \hat{\rho}$$

If acting on an arbitrary state $|\chi\rangle$:

$$\hat{\rho}|\chi\rangle = |\beta\rangle\langle\beta|\chi\rangle = \langle\beta|\chi\rangle|\beta\rangle$$

the operator $\hat{\rho}$ forces the result to be parallel to $|\beta\rangle$ with amplitude given by the scalar product $\langle\beta|\chi\rangle$ (because $\hat{\rho}$ is a projection operator).

Finally, the density operator is positive-definite: $\langle\varphi|\hat{\rho}|\varphi\rangle \geq 0$. For an arbitrary vector $|\varphi\rangle$,

$$\begin{aligned}\langle\varphi|\hat{\rho}|\varphi\rangle &= \sum_{\alpha=1}^{N} P_\alpha \langle\varphi|\alpha\rangle\langle\alpha|\varphi\rangle \\ &= \sum_{\alpha=1}^{N} P_\alpha |\langle\varphi|\alpha\rangle|^2 \geq 0\end{aligned}$$

Eigenrepresentation

If we know the eigenvalues p_m and eigenstates $|m\rangle$ of the density operator, then its eigenrepresentation is

$$\hat{\rho} = \sum_m p_m \, |m\rangle\langle m|$$

Properties of hermicity and positive-definite mean that all eigenvalues are real and positive $p_m > 0$. Also, $\mathrm{tr}(\hat{\rho}) = 1$ means

$$\sum_m p_m = 1$$

Note that the eigenvalues p_m are only identical to the probabilities P_α if the states $|\alpha\rangle$ used to define the ensemble are orthogonal to each other.

For a pure ensemble, $p_m = 1$ for one state, while all the other $p_j = 1 \;\; (j \neq m)$. In this case, $\hat{\rho} = |m\rangle\langle m|$ is a projection operator, and therefore fulfills $\hat{\rho}^2 = \hat{\rho}$. It follows that for a pure state $\mathrm{tr}(\hat{\rho}^2) = 1$. For a mixed state, however,

$$\mathrm{tr}(\hat{\rho}^2) = \sum_m |p_m|^2 < 1$$

Thus we define the purity $\wp$

$$\wp = \mathrm{tr}(\hat{\rho}^2)$$

which distinguishes between pure ($\wp = 1$) and mixed ($\wp < 1$) states.

The maximally mixed state is described by the density matrix

$$\hat{\rho} = \frac{1}{N}\hat{I}$$

(where N is the Hilbert space dimension), and it has purity $\wp = 1/N$.

Example

Consider the density matrix

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

This describes a system in state $|0\rangle$ with probability ¾ and in state $|1\rangle$ with probability ¼. However, this assignment of the density matrix to a particular ensemble of states is not unique. For example, we could consider two other states:

$$|a\rangle = \sqrt{\frac{3}{4}}|0\rangle + \sqrt{\frac{1}{4}}|1\rangle$$

$$|b\rangle = \sqrt{\frac{3}{4}}|0\rangle - \sqrt{\frac{1}{4}}|1\rangle$$

We then prepare an ensemble with state $|a\rangle$ with probability ½ and in the state $|b\rangle$ with probability ½. The density matrix is

$$\hat{\rho} = \frac{1}{2}|a\rangle\langle a| + \frac{1}{2}|b\rangle\langle b|$$

$$= \frac{1}{2}\left(\sqrt{\frac{3}{4}}|0\rangle + \sqrt{\frac{1}{4}}|1\rangle\right)\left(\sqrt{\frac{3}{4}}\langle 0| + \sqrt{\frac{1}{4}}\langle 1|\right)$$

$$+ \frac{1}{2}\left(\sqrt{\frac{3}{4}}|0\rangle - \sqrt{\frac{1}{4}}|1\rangle\right)\left(\sqrt{\frac{3}{4}}\langle 0| - \sqrt{\frac{1}{4}}\langle 1|\right)$$

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

These two ensembles have the same density matrix. Generally, the eigenvectors and eigenstates of a density matrix indicate only one of the many possible ensembles described by that density matrix.

We can calculate the purity $\wp$ with

$$\hat{\rho}^2 = \frac{9}{16}|0\rangle\langle 0| + \frac{1}{16}|1\rangle\langle 1|$$

so that

$$\wp = \mathrm{tr}(\hat{\rho}^2) = \frac{5}{8}$$

Example

Consider a single qubit state

$$|\psi_0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

This is an eigenstate of σ_x with eigenvalue 1. Let's calculate the expectation value of σ_x using the density matrix:

$$\begin{aligned}\hat{\rho} &= |\psi_0\rangle\langle\psi_0| \\ &= \frac{1}{2}(|0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| + |1\rangle\langle 1|)\end{aligned}$$

Note that $\hat{\rho}^2 = \hat{\rho}$ because this is a pure state.

With $\sigma_x = |0\rangle\langle 1| + |1\rangle\langle 0|$,

$$\sigma_x\hat{\rho} = \frac{1}{2}(|0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| + |1\rangle\langle 1|)$$

so that

$$\langle\sigma_x\rangle = \mathrm{tr}(\sigma_x\hat{\rho}) = 1$$

as expected (because $|\psi_0\rangle$ is an eigenstate of σ_x).

Consider instead a mixture with state $|0\rangle$ occurring with probability ½ and state $|1\rangle$ occurring with probability ½:

$$\hat{\rho} = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|)$$

Note that $\hat{\rho}^2 \neq \hat{\rho}$ because this is a mixed state.

With $\sigma_x = |0\rangle\langle 1| + |1\rangle\langle 0|$,

$$\sigma_x\hat{\rho} = \frac{1}{2}(|0\rangle\langle 1| + |1\rangle\langle 0|)$$

so that

$$\langle\sigma_x\rangle = \mathrm{tr}(\sigma_x\hat{\rho}) = 0$$

This illustrates that there is an essential difference between a coherent superposition of the states $|0\rangle$ and $|1\rangle$ and a probabilistic ensemble in which $|0\rangle$ and $|1\rangle$ can each occur with specified probabilities. In the latter case (a mixed state), we say that $\hat{\rho}$ is an incoherent superposition of the states $|0\rangle$ and $|1\rangle$, incoherent meaning that the relative phases of $|0\rangle$ and $|1\rangle$ are experimentally inaccessible.

The Bloch sphere and polarisation

Let's consider a single qubit as described by an arbitrary Hermitian 2×2 density matrix:

$$\begin{aligned}\hat{\rho} &= \frac{1}{2}\left(I + \vec{P}\cdot\vec{\sigma}\right) \\ &= \frac{1}{2}\left(I + P_x\sigma_x + P_y\sigma_y + P_z\sigma_z\right) \\ &= \frac{1}{2}\begin{pmatrix} 1+P_z & P_x - iP_y \\ P_x + iP_y & 1-P_z \end{pmatrix}\end{aligned}$$

where P_x, P_y, P_z are all real, and the factor of ½ is chosen so that $\mathrm{tr}\hat{\rho} = 1$. We also have

$$\det\hat{\rho} = \frac{1}{4}(1 - P^2)$$

where $P^2 = \vec{P}\cdot\vec{P} = P_x^2 + P_y^2 + P_z^2$. Then,

$$\begin{aligned}&\langle\vec{\sigma}\rangle = \mathrm{tr}(\hat{\rho}\vec{\sigma}) \\ &= \mathrm{tr}\left[\frac{1}{2}\left(I + P_x\sigma_x + P_y\sigma_y + P_z\sigma_z\right)\left(\sigma_x\vec{\imath} + \sigma_y\vec{\jmath} + \sigma_z\vec{k}\right)\right] \\ &= \mathrm{tr}\left[\frac{1}{2}\left(IP_x\vec{\imath} + IP_y\vec{\jmath} + IP_z\vec{k}\right)\right] \\ &= \vec{P}\end{aligned}$$

which shows that $\vec{P}$ is the polarisation $\vec{P} = \langle\vec{\sigma}\rangle$.

The purity of $\hat{\rho}$ is given by

$$\wp = \mathrm{tr}(\hat{\rho}^2) = \frac{1}{2}(1 + P^2)$$

Pure states

Normalised eigenstates $|\psi_{\pm}\rangle$ can be written as

$$|\psi_{+}\rangle = \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2)\, e^{i\varphi} \end{pmatrix};\ |\psi_{-}\rangle = \begin{pmatrix} \sin(\theta/2) \\ -\cos(\theta/2)\, e^{i\varphi} \end{pmatrix}$$

so that the density matrix of a pure state can be written as

$$\hat{\rho}_{\pm} = |\psi_{\pm}\rangle\langle\psi_{\pm}|$$

Recall that

$$|a\rangle\langle b| = \begin{pmatrix} a_0 b_0^* & a_0 b_1^* \\ a_1 b_0^* & a_1 b_1^* \end{pmatrix}$$

so

$$\hat{\rho}_{+} = \begin{pmatrix} cos^2(\theta/2) & \cos(\theta/2)\sin(\theta/2)\, e^{-i\varphi} \\ \cos(\theta/2)\sin(\theta/2)\, e^{i\varphi} & sin^2(\theta/2) \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1+\cos\theta & \sin\theta(\cos\varphi - i\sin\varphi) \\ \sin\theta(\cos\varphi + i\sin\varphi) & 1-\cos\theta \end{pmatrix}$$

$$\hat{\rho}_{-} = \begin{pmatrix} sin^2(\theta/2) & -\cos(\theta/2)\sin(\theta/2)\, e^{-i\varphi} \\ -\cos(\theta/2)\sin(\theta/2)\, e^{i\varphi} & cos^2(\theta/2) \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1-\cos\theta & -\sin\theta(\cos\varphi - i\sin\varphi) \\ -\sin\theta(\cos\varphi + i\sin\varphi) & 1+\cos\theta \end{pmatrix}$$

These show that

$$\vec{P} = \pm\hat{n}$$

for these pure states. We can thus write

$$\hat{\rho}_{\pm} = \frac{1}{2}(I \pm \hat{n} \cdot \vec{\sigma})$$

With $\vec{P} = \pm\hat{n}$, then $P^2 = \vec{P} \cdot \vec{P} = \hat{n} \cdot \hat{n} = 1$, so that

$$\det\hat{\rho} = 0, \qquad \wp = \mathrm{tr}(\hat{\rho}^2) = 1$$

Mixed states

For a mixed state, it is still possible to write

$$\hat{\rho} = \frac{1}{2}\left(I + \vec{P} \cdot \vec{\sigma}\right)$$

with polarisation $\vec{P} = \langle \vec{\sigma} \rangle$, although now $\vec{P} \neq \pm\hat{n}$.

The purity of $\hat{\rho}$ is given by

$$\wp = \mathrm{tr}(\hat{\rho}^2) = \frac{1}{2}(1 + P^2)$$

where $\wp < 1$ when $P^2 < 1$. Thus, for a mixed state, the density matrix (of a single qubit) corresponds to a point *within* the Bloch sphere with polarisation $\vec{P}$ where $0 \leq \left|\vec{P}\right| < 1$. Vector $\vec{P}$ is sometimes called the "Bloch vector". If we're within the Bloch sphere, we might describe it as the "Bloch ball".

Time evolution of the density matrix

If we apply a unitary transformation $\hat{U}$ to *all* states $|\alpha\rangle$,

$$|\beta\rangle = \hat{U}|\alpha\rangle \quad \Leftrightarrow \quad \langle\beta| = \langle\alpha|\hat{U}^+$$

then the density matrix transforms as

$$\hat{\rho}' = \sum_{\alpha=1}^{N} P_\alpha \hat{U}|\alpha\rangle\langle\alpha|\hat{U}^+ = \hat{U}\hat{\rho}\hat{U}^+$$

In particular, if we consider time evolution with the time evolution operator

$$|\alpha(t)\rangle = \hat{U}(t,t_0)|\alpha(t_0)\rangle$$

then, at time t,

$$\hat{\rho}(t) = \hat{U}(t,t_0)\hat{\rho}(t_0)\hat{U}^+(t,t_0)$$

With

$$\hat{U}(t,t_0) = e^{-iH(t-t_0)/\hbar}$$

calculate

$$\frac{d\hat{\rho}}{dt} = \frac{d\hat{U}(t,t_0)}{dt}\hat{\rho}(t_0)\hat{U}^+(t,t_0) + \hat{U}(t,t_0)\hat{\rho}(t_0)\frac{d\hat{U}^+(t,t_0)}{dt}$$

$$= -\frac{i}{\hbar}(H\hat{\rho} - \hat{\rho}H)$$

$$\frac{d\hat{\rho}}{dt} = \frac{i}{\hbar}[\hat{\rho}, H]$$

where $[\hat{\rho}, H] = \hat{\rho}H - H\hat{\rho}$. This is the quantum Liouville equation.

Measurement

Suppose we make a measurement described by projection operator $\hat{P}_n = |n\rangle\langle n|$ where $|n\rangle$ is a state with associated eigenvalue a_n. When acting on state $|\alpha\rangle$, the probability $P(n|\alpha)$ to obtain outcome a_n (given α) is

$$P(n|\alpha) = \langle\alpha|\hat{P}_n|\alpha\rangle = \langle\alpha|n\rangle\langle n|\alpha\rangle$$

With $\mathrm{tr}\left(\hat{B}|\alpha\rangle\langle\alpha|\right) = \sum_\beta \langle\beta|\hat{B}|\alpha\rangle\langle\alpha|\beta\rangle = \langle\alpha|\hat{B}|\alpha\rangle$, we can write

$$P(n|\alpha) = \langle\alpha|n\rangle\langle n|\alpha\rangle = \mathrm{tr}(|n\rangle\langle n|\alpha\rangle|\langle\alpha|)$$

When acting on the ensemble, the probability $P(n)$ to obtain outcome a_n is

$$\begin{aligned} P(n) &= \sum_{\alpha=1}^{N} P(n|\alpha)\, P_\alpha \\ &= \sum_{\alpha=1}^{N} \mathrm{tr}(|n\rangle\langle n|\alpha\rangle|\langle\alpha|)\, P_\alpha \\ &= \sum_{\alpha=1}^{N} \sum_i \langle i|n\rangle\langle n|\alpha\rangle\langle\alpha|i\rangle\, P_\alpha \\ &= \sum_i \langle i|n\rangle\langle n|\hat{\rho}|i\rangle \\ &= \mathrm{tr}(|n\rangle\langle n|\hat{\rho}) \end{aligned}$$

After acting on state $|\alpha\rangle$, the post-measurement state is

$$|\alpha_n\rangle = \frac{1}{\sqrt{P(n|\alpha)}}\hat{P}_n|\alpha\rangle = \frac{|n\rangle\langle n|\alpha\rangle}{\sqrt{\langle\alpha|n\rangle\langle n|\alpha\rangle}}$$

For the ensemble, we have an ensemble of states $|\alpha_n\rangle$ each with probability $P(\alpha|n)$. Hence, the density operator $\hat{\rho}_n$ is

$$\begin{aligned}\hat{\rho}_n &= \sum_{\alpha=1}^{N} P(\alpha|n)\,|\alpha_n\rangle\langle\alpha_n| \\ &= \sum_{\alpha=1}^{N} P(\alpha|n)\,\frac{|n\rangle\langle n|\alpha\rangle\langle\alpha|n\rangle\langle n|}{\langle\alpha|n\rangle\langle n|\alpha\rangle}\end{aligned}$$

Now use $P(\alpha|n)P(n) = P(n|\alpha)P_\alpha$ to write

$$P(\alpha|n) = P_\alpha\frac{P(n|\alpha)}{P(n)} = P_\alpha\frac{\langle\alpha|n\rangle\langle n|\alpha\rangle}{\mathrm{tr}(|n\rangle\langle n|\hat{\rho})}$$

so that

$$\begin{aligned}\hat{\rho}_n &= \sum_{\alpha=1}^{N} P_\alpha\frac{|n\rangle\langle n|\alpha\rangle\langle\alpha|n\rangle\langle n|}{tr(|n\rangle\langle n|\hat{\rho})} \\ &= \frac{|n\rangle\langle n|\hat{\rho}|n\rangle\langle n|}{\mathrm{tr}(|n\rangle\langle n|\hat{\rho})}\end{aligned}$$

Reduced density matrix

Suppose we have two physical systems A and B whose state is described by density operator $\hat{\rho}_{AB}$. The reduced density operator for system A is defined by

$$\hat{\rho}_A = \mathrm{tr}_B(\hat{\rho}_{AB})$$

where tr_B is the partial trace over system B. Likewise,

$$\hat{\rho}_B = \mathrm{tr}_A(\hat{\rho}_{AB})$$

The partial trace is defined by

$$\mathrm{tr}_B(F) = \sum_j \left(I_A \otimes {}_B\langle j|\right) F \left(I_A \otimes |j\rangle_B\right)$$

where I_A is the identity operator acting on A and $\{|j\rangle_B\}$ is any orthonormal basis for B. Likewise,

$$\mathrm{tr}_A(F) = \sum_j \left({}_A\langle j| \otimes I_B\right) F \left(|j\rangle_A \otimes I_B\right)$$

Consider a general term, $|a_1 b_1\rangle\langle a_2 b_2|$, appearing in an operator, then

$$\begin{aligned}
&\mathrm{tr}_B(|a_1 b_1\rangle\langle a_2 b_2|) \\
&\quad = \sum_j \left(I_A \otimes {}_B\langle j|\right)|a_1 b_1\rangle\langle a_2 b_2| \left(I_A \otimes |j\rangle_B\right) \\
&\quad = |a_1\rangle\langle a_2| \sum_j \langle j|b_1\rangle\langle b_2|j\rangle \\
&\quad = |a_1\rangle\langle a_2| \mathrm{tr}(|b_1\rangle\langle b_2|)
\end{aligned}$$

$$= |a_1\rangle\langle a_2| \sum_j \langle b_2|j\rangle\langle j|b_1\rangle$$
$$= |a_1\rangle\langle a_2|\langle b_2|b_1\rangle$$
$$= |a_1\rangle\langle a_2|\delta_{b_2,b_1}$$

where the last line holds if $|b_1\rangle$ and $|b_2\rangle$ are orthonormal.

Likewise,

$$\mathrm{tr}_B(|a_1b_1\rangle\langle a_2b_2|) = |b_1\rangle\langle b_2|\langle a_2|a_1\rangle$$
$$= |b_1\rangle\langle b_2|\delta_{a_2,a_1}$$

where the last line holds if $|a_1\rangle$ and $|a_2\rangle$ are orthonormal.

Note that, with the general form of an operator F,

$$F = \sum_{k,m,\ell,n} F_{k,m,\ell,n}\, |km\rangle\langle \ell n|$$

where $F_{k,m,\ell,n} = \langle km|F|\ell m\rangle$, then it is possible to write

$$\mathrm{tr}_B(F) = \sum_{k,\ell,m} F_{k,m,\ell,m}\, |k\rangle\langle \ell|$$

$$\mathrm{tr}_A(F) = \sum_{m,n,k} F_{k,m,k,n}\, |m\rangle\langle n|$$

Why use the reduced density matrix?

The reduced density matrix operator for A provides the correct measurement statistics for measurements made on system A.

Suppose M_A is a Hermitian operator corresponding to an observable on system A. For the combined system AB, the measurement has the form $M_A \otimes I_B$, and the expectation value of M_A is given by

$$\langle M_A \rangle = \mathrm{tr}\big(\hat{\rho}_{AB}(M_A \otimes I_B)\big)$$

where the trace is over A and B. However, the reduced density matrix gives an expectation value as

$$\langle M_A \rangle = \mathrm{tr}(\hat{\rho}_A M_A)$$

where the trace is over A only.

Explicitly, consider AB to be prepared in a pure state

$$|\psi\rangle_{AB} = \sum_{k,m} c_{k,m}\,|k\rangle_A |m\rangle_B$$

where $\sum_{k,m}\left|c_{k,m}\right|^2 = 1$. Then,

$$\begin{aligned}\langle M_A \rangle &= {}_{AB}\langle\psi|M_A \otimes I_B|\psi\rangle_{AB} \\ &= \sum_{k,\ell,m,n} c_{k,m}c^*_{\ell,n}\ {}_A\langle\ell|\,{}_B\langle n|M_A \otimes I_B|k\rangle_A|m\rangle_B \\ &= \sum_{k,\ell,m,n} c_{k,m}c^*_{\ell,n}\ {}_A\langle\ell|M_A|k\rangle_A \delta_{n,m}\end{aligned}$$

$$= \sum_{k,\ell,m} c_{k,m} c^*_{\ell,n} \; {}_A\langle \ell | M_A | k \rangle_A$$

The reduced density matrix $\hat{\rho}_A$ is given by

$$\begin{aligned}
\hat{\rho}_A &= \mathrm{tr}_B(\rho_{AB}) = \mathrm{tr}_B\big(|\psi\rangle_{AB\,AB}\langle\psi|\big) \\
&= \mathrm{tr}_B\left(\sum_{k,\ell,m,n} c_{k,m} c^*_{\ell,n} \, |k\rangle_A |m\rangle_B \; {}_A\langle \ell| \; {}_B\langle n| \right) \\
&= \sum_{k,\ell,m,n} c_{k,m} c^*_{\ell,n} \, \mathrm{tr}_B\big(|k\rangle_A |m\rangle_B \; {}_A\langle \ell| \; {}_B\langle n|\big) \\
&= \sum_{k,\ell,m,n} c_{k,m} c^*_{\ell,n} \, |k\rangle_A \; {}_A\langle \ell| \delta_{n,m} \\
&= \sum_{k,\ell,m} c_{k,m} c^*_{\ell,m} \, |k\rangle_A \; {}_A\langle \ell|
\end{aligned}$$

Thus

$$\begin{aligned}
\mathrm{tr}(\hat{\rho}_A M_A) &= \mathrm{tr}\left(\sum_{k,\ell,m} c_{k,m} c^*_{\ell,m} \, |k\rangle_A \; {}_A\langle \ell | M_A \right) \\
&= \sum_{k,\ell,m,j} c_{k,m} c^*_{\ell,m} \; {}_A\langle j | k \rangle_A \; {}_A\langle \ell | M_A | j \rangle_A \\
&= \sum_{k,\ell,m} c_{k,m} c^*_{\ell,m} \; {}_A\langle \ell | M_A | k \rangle_A
\end{aligned}$$

which is the same as the expression for $\langle M_A \rangle$ above.

Example

Consider a system prepared in the product state $\hat{\rho}_{AB} = \hat{\rho}_1 \otimes \hat{\rho}_2$, where $\hat{\rho}_1$ is a density operator for system A, and $\hat{\rho}_2$ is a density operator for system B. Then

$$\begin{aligned}\hat{\rho}_A &= \mathrm{tr}_B(\hat{\rho}_1 \otimes \hat{\rho}_2) \\ &= \sum_j \left(I_A \otimes {}_B\langle j|\right)(\hat{\rho}_1 \otimes \hat{\rho}_2)(I_A \otimes |j\rangle_B) \\ &= I_A \hat{\rho}_1 I_A \sum_j {}_B\langle j|\hat{\rho}_2|j\rangle_B \\ &= \hat{\rho}_1\end{aligned}$$

as expected (because $\mathrm{tr}_B(\hat{\rho}_2) = 1$).

Example

Consider the Bell state

$$|\beta_{00}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

It is described by a single state, so is a pure state (although it is entangled). Its density operator is

$$\begin{aligned}\hat{\rho}_{AB} &= |\beta_{00}\rangle\langle\beta_{00}| \\ &= \left[\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)\right]\left[\frac{1}{\sqrt{2}}(\langle 00| + \langle 11|)\right] \\ &= \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 11|)\end{aligned}$$

Note that $\langle 00|00\rangle = \langle 11|11\rangle = 1$ and $\langle 00|11\rangle = 0$ so that it is easy to show $\hat{\rho}_{AB}^2 = \hat{\rho}_{AB}$ ($|\beta_{00}\rangle$ is a pure state).

The reduced density operator for system A is

$$\hat{\rho}_A = \mathrm{tr}_B(\hat{\rho}_{AB})$$

To evaluate this, we can use

$$\mathrm{tr}_B(|a_1 b_1\rangle\langle a_2 b_2|) = |a_1\rangle\langle a_2|\delta_{b_2,b_1}$$

so that, for example,

$$\mathrm{tr}_B(|00\rangle\langle 00|) = |0\rangle\langle 0|$$
$$\mathrm{tr}_B(|00\rangle\langle 11|) = 0$$

Hence, for this Bell state,

$$\begin{aligned}\hat{\rho}_A &= \mathrm{tr}_B(\hat{\rho}_{AB}) \\ &= \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) \\ &= \frac{I}{2}\end{aligned}$$

where I is the identity operator.

Now consider

$$\hat{\rho}_A^2 = \left(\frac{I}{2}\right)^2 = \frac{I}{4}$$

Hence $\hat{\rho}_A^2 \neq \hat{\rho}_A$, so this state is a mixed state. The state of the joint system of two qubits is a pure state and is known exactly; however, the first qubit is in a mixed

state, that is, a state about which we do not have maximal knowledge. This strange property, that the joint state of a system can be completely known, yet a subsystem be in mixed states, is a hallmark of quantum entanglement. When system A interacts with system B, A and B become entangled (i.e. correlated). The entanglement destroys the coherence of a superposition of states of A so that some of the phases in the superposition become inaccessible if we look at A alone. The state of system A is in one of a set of alternative states, each of which can be assigned a probability.

Starting from

$$|\beta_{00}\rangle = \frac{1}{\sqrt{2}}|0\rangle_A|0\rangle_B + \frac{1}{\sqrt{2}}|1\rangle_A|1\rangle_B$$

we stopped looking at qubit B (assuming we would never be able to look at it again). We might have given it to Bob who then took it light-years away from us. Anything that happened to that qubit B was then completely unable to affect the state of the qubit A we kept. For all we know, Bob could have measured it. Because our state is entangled, we know that any measurement results must be correlated, so if Bob measured his qubit in state $|0\rangle_B$, then he knows our qubit must also be in state $|0\rangle_A$. Bob has a 50/50

chance of measuring $|0\rangle_B$ or $|1\rangle_B$, so, that tells us that our qubit is in state $|0\rangle_A$ with probability 1/2 or in state $|1\rangle_A$ with probability 1/2. Hence, the density matrix is

$$\hat{\rho}_A = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|)$$

Whatever Bob does to qubit B can't affect qubit A as described by $\hat{\rho}_A$: there isn't any faster than light communication.

A two-qubit system in a pure state

A two-qubit state can be written generically as

$$|\psi\rangle = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle$$

with normalisation $|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 = 1$. For such a state, the degree of entanglement is characterised by the concurrence C,

$$C = 2|\alpha\delta - \beta\gamma|$$

where $0 \leq C \leq 1$: for separable states, $C = 0$, whereas for entangled states, $C > 0$.

We want to write the density matrix for this pure state:

$$\hat{\rho} = |\psi_0\rangle\langle\psi_0|$$

which will be a 4×4 matrix. This is done with the outer product $|a\rangle\langle b|$. For example,

$$|a\rangle \equiv \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix}; \qquad \langle \mathrm{b}| \equiv (b_0^* \quad b_1^* \quad b_2^* \quad b_3^*)$$

$$|a\rangle\langle b| = \begin{pmatrix} a_0 b_0^* & a_0 b_1^* & a_0 b_2^* & a_0 b_3^* \\ a_1 b_0^* & a_1 b_1^* & a_1 b_2^* & a_1 b_3^* \\ a_2 b_0^* & a_2 b_1^* & a_2 b_2^* & a_2 b_3^* \\ a_3 b_0^* & a_3 b_1^* & a_3 b_2^* & a_3 b_3^* \end{pmatrix}$$

so $|a\rangle$ gives the row, $\langle b|$ gives the column.

Using tensor products,

$$|00\rangle \equiv |0\rangle \otimes |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$|01\rangle \equiv |0\rangle \otimes |1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

$$|10\rangle \equiv |1\rangle \otimes |0\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

$$|11\rangle \equiv |1\rangle \otimes |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

we can write

$$\rho = \begin{pmatrix} |\alpha|^2 & \alpha\beta^* & \alpha\gamma^* & \alpha\delta^* \\ \beta\alpha^* & |\beta|^2 & \beta\gamma^* & \beta\delta^* \\ \gamma\alpha^* & \gamma\beta^* & |\gamma|^2 & \gamma\delta^* \\ \delta\alpha^* & \delta\beta^* & \delta\gamma^* & |\delta|^2 \end{pmatrix}$$

We can write this in terms of 2×2 blocks,

$$\rho = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

$$A = \begin{pmatrix} |\alpha|^2 & \alpha\beta^* \\ \beta\alpha^* & |\beta|^2 \end{pmatrix}, \qquad B = \begin{pmatrix} \alpha\gamma^* & \alpha\delta^* \\ \beta\gamma^* & \beta\delta^* \end{pmatrix}$$
$$C = \begin{pmatrix} \gamma\alpha^* & \gamma\beta^* \\ \delta\alpha^* & \delta\beta^* \end{pmatrix}, \qquad D = \begin{pmatrix} |\gamma|^2 & \gamma\delta^* \\ \delta\gamma^* & |\delta|^2 \end{pmatrix}$$

The reduced density operator for A is

$$\hat{\rho}_A = \mathrm{tr}_B(\hat{\rho}_{AB}) = \sum_{n=1}^{1} {}_B\langle n|\,\hat{\rho}_{AB}\,|n\rangle_B$$
$$= (|\alpha|^2 + |\beta|^2)|0\rangle\langle 0| + (\alpha\gamma^* + \beta\delta^*)|0\rangle\langle 1|$$
$$+ (\gamma\alpha^* + \delta\beta^*)|1\rangle\langle 0| + (|\gamma|^2 + |\delta|^2)|1\rangle\langle 1|$$

or, as a matrix,

$$\rho_A = \begin{pmatrix} trA & trB \\ trC & trD \end{pmatrix} = \begin{pmatrix} |\alpha|^2 + |\beta|^2 & \alpha\gamma^* + \beta\delta^* \\ \gamma\alpha^* + \delta\beta^* & |\gamma|^2 + |\delta|^2 \end{pmatrix}$$

The reduced density operator for B is

$$\hat{\rho}_B = \mathrm{tr}_A(\hat{\rho}_{AB}) = \sum_{n=1}^{1} {}_A\langle n|\,\hat{\rho}_{AB}\,|n\rangle_A$$
$$= (|\alpha|^2 + |\gamma|^2)|0\rangle\langle 0| + (\alpha\beta^* + \gamma\delta^*)|0\rangle\langle 1|$$
$$+ (\beta\alpha^* + \delta\gamma^*)|1\rangle\langle 0| + (|\beta|^2 + |\delta|^2)|1\rangle\langle 1|$$

or, as a matrix,

$$\rho_B = A + D = \begin{pmatrix} |\alpha|^2 + |\gamma|^2 & \alpha\beta^* + \gamma\delta^* \\ \beta\alpha^* + \delta\gamma^* & |\beta|^2 + |\delta|^2 \end{pmatrix}$$

The purity $\wp = \mathrm{tr}(\hat{\rho}^2)$ [which distinguishes between pure ($\wp = 1$) and mixed ($\wp < 1$) states] of these

reduced density matrices is related to the concurrence C [which distinguishes between separable ($C = 0$) and entangled ($C > 0$) states],

$$tr(\hat{\rho}_A^2) = tr(\hat{\rho}_B^2) = 1 - \frac{C^2}{2}$$

There is also the identity $det(\hat{\rho}_A) = C^2/4$.

For this system, a two-qubit composite system in a pure state, then

- if it is separable ($C = 0$), the reduced density matrices are pure ($tr(\hat{\rho}_A^2) = tr(\hat{\rho}_B^2) = 1$)
- if it is entangled ($C \neq 0$), the reduced density matrices are mixed ($tr(\hat{\rho}_A^2) = tr(\hat{\rho}_B^2) < 1$)

Entropy

Given a classical probability distribution $\{p_i\}$, the entropy S is defined by

$$S = -\sum_i p_i \ln p_i \qquad (1)$$

where $\ln$ is the natural logarithm. In information theory, this is called the Shannon entropy. In physics, this quantity is usually multiplied by the Boltzmann constant k_B and is called the Gibbs entropy.

The entropy is a measure of the uncertainty encoded in the probability distribution. For example, if there's no uncertainty because, say $p_i = 1$ while all other $p_i = 0$, then $S = 0$. In contrast, if there are N possibilities, the entropy is maximised when we have no idea which is most likely, meaning that $p_i = 1/N$ for each. In this case

$$S = -\sum_i \frac{1}{N} \ln \frac{1}{N} = \ln N$$

Von Neumann entropy: for a quantum state described by a density matrix $\hat{\rho}$,

$$S = -\mathrm{tr}(\hat{\rho} \ln \hat{\rho})$$

Entanglement entropy: this is the von Neumann entropy of a reduced density matrix that came from taking a partial trace of a pure state of a larger system:

$$\mathcal{E} = -tr(\hat{\rho}_A \ln \hat{\rho}_A) = -tr(\hat{\rho}_B \ln \hat{\rho}_B)$$

If we have the eigenrepresentation of a density matrix,

$$\hat{\rho} = \sum_m p_m \, |m\rangle\langle m|$$

Then $ln\hat{\rho} = \sum_m ln(p_m) \, |m\rangle\langle m|$, and the entropy is

$$\begin{aligned} S &= -\sum_j \langle j| \left(\sum_n p_n \, |n\rangle\langle n| \sum_m \ln(p_m) \, |m\rangle\langle m| \right) |j\rangle \\ &= -\sum_{j,m,n} p_n \ln(p_m) \langle j|n\rangle\langle n|m\rangle\langle m|j\rangle \\ &= -\sum_n p_n \ln(p_n) \end{aligned}$$

which agrees with the classical definition, Eq.(1). For a pure state $S = 0$, for a mixed state $S \neq 0$. For a maximally mixed state $\hat{\rho} = \hat{I}/N$ (where N is the Hilbert space dimension) then $S = \ln N$.

Schmidt decomposition and purification

Schmidt decomposition: suppose $|\psi\rangle$ is a pure state of a composite system, AB. Then there exist orthonormal states $|i_A\rangle$ of system A, and orthonormal states $|i_B\rangle$ of system B such that

$$|\psi\rangle = \sum_i \lambda_i \, |i_A\rangle|i_B\rangle$$

where λ_i are non-negative real numbers satisfying $\sum_i |\lambda_i|^2 = 1$ known as Schmidt coefficients.

Proof (when A and B have Hilbert spaces of the same dimensions): Starting from an orthonormal basis $|k\rangle$ ($k = 1,2,\ldots,N_A$) for system A and $|m\rangle$ ($m = 1,2,\ldots,N_B$) for system B, a joint state

$$|\psi\rangle = \sum_{k,m} c_{k,m} \, |k\rangle|m\rangle$$

of the composite system can be written by using combined basis states $|k\rangle|m\rangle$, where $\sum_{k,m} \left|c_{k,m}\right|^2 = 1$. Coefficients $c_{k,m}$ can be regarded as matrix elements of a matrix c. "Singular value decomposition" can be used to write this matrix as $c = UDV$ where D is a diagonal matrix with non-negative elements, and U and V are unitary matrices (this is like diagonalisation of a Hermitian matrix). Hence

$$|\psi\rangle = \sum_{i,k,m} U_{k,i}\, D_{i,i} V_{i,m} |k\rangle |m\rangle$$

Defining $|i_A\rangle = \sum_k U_{k,i}\, |k\rangle$, $|i_B\rangle = \sum_m V_{i,m}\, |m\rangle$ and $\lambda_i = D_{i,i}$, then $|\psi\rangle = \sum_i \lambda_i\, |i_A\rangle |i_B\rangle$. These states are orthonormal due to unitarity of U and V, and the orthonormality of $|k\rangle$ and $|m\rangle$.

One consequence of Schmidt decomposition is that, for a pure state $|\psi\rangle$, reduced density matrices are

$$\rho_A = \sum_i \lambda_i^2\, |i_A\rangle\langle i_A|, \qquad \rho_B = \sum_i \lambda_i^2\, |i_B\rangle\langle i_B|$$

so they have identical eigenvalues (λ_i^2). Because of the identical eigenvalues, there is the same entanglement entropy for ρ_A and ρ_B,

$$\mathcal{E} = -\sum_i \lambda_i^2 \ln\left(\lambda_i^2\right)$$

Schmidt bases: the bases $|i_A\rangle$ and $|i_B\rangle$ for A and B, respectively.

Schmidt number (or “Schmidt rank”) R for $|\psi\rangle$: the number of non-zero values λ_i.

If $R = 1$, the pure state is separable $|\psi\rangle = |i_A\rangle |i_B\rangle$, but, if $R > 1$, the state $|\psi\rangle$ is entangled.

The Schmidt number is a useful measure of entanglement between A and B. It is useful because it

is independent of unitary transformations on either A or B.

Imagine the pure state is separable $|\psi\rangle = |i_A\rangle|i_B\rangle$ ($R = 1$). If Alice and Bob each operate on their own qubit, this is described a unitary matrix

$$U = U_A \otimes U_B$$

creating a new state $|\varphi\rangle$,

$$|\varphi\rangle = U|\psi\rangle = U_A|i_A\rangle \otimes U_B|i_B\rangle$$

This new state $|\varphi\rangle$ is still separable ($R = 1$): we cannot change the amount of entanglement by local operators which act only on part of the Hilbert space; entanglement can only be created by bringing the two subsystems together.

Purification: given system A with state defined by a (reduced) density matrix ρ_A, we define a reference system R so that the composite system AR is described by a pure state $|AR\rangle$ such that

$$\rho_A = \mathrm{tr}_R(\rho_{AR}) = \mathrm{tr}_R(|AR\rangle\langle AR|)$$

Let's see how to construct R when ρ_A is written in terms of orthonormal states $|i_A\rangle$ as

$$\rho_A = \sum_i P_i \, |i_A\rangle\langle i_A|$$

Now introduce system R with the same Hilbert space as A and with orthonormal states $|i_R\rangle$. Then define a pure state for AR as

$$|AR\rangle = \sum_i \sqrt{P_i}\,|i_A\rangle|i_R\rangle$$

Check this works:

$$\begin{aligned}
\rho_A &= \mathrm{tr}_R(\rho_{AR}) = \mathrm{tr}_R(|AR\rangle\langle AR|) \\
&= \mathrm{tr}_R\left(\sum_{i,j}\sqrt{P_iP_j}\,|i_A\rangle|i_R\rangle\langle j_A|\langle j_R|\right) \\
&= \sum_{i,j}\sqrt{P_iP_j}\,|i_A\rangle\langle j_A|\sum_k \langle k|i_R\rangle\,\langle j_R|k\rangle \\
&= \sum_{i,j}\sqrt{P_iP_j}\,|i_A\rangle\langle j_A|\delta_{i,j} \\
&= \sum_i P_i\,|i_A\rangle\langle i_A|
\end{aligned}$$

as it should be.

Note the close relationship of purification to Schmidt decomposition: the procedure used to purify a mixed state of system A is to define a pure state whose Schmidt basis for system A ($|i_A\rangle$) is just the basis in which the mixed state is diagonal, with the Schmidt coefficients $\sqrt{P_i}$ being the square root of the eigenvalues P_i of the density operator being purified.

Example

For a composite system of two qubits, determine the Schmidt decomposition of the state

$$|\psi\rangle = \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle)$$

It's actually possible to guess the answer (by inspection). But, let's imagine we can't do this, so we need to calculate. We have

$$|\psi\rangle = \sum_{k,m} c_{k,m}\,|k\rangle|m\rangle$$

and write the coefficients $c_{k,m}$ as matrix elements of a matrix c:

$$c = \begin{pmatrix} c_{0,0} & c_{0,1} \\ c_{1,0} & c_{1,1} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

This is Hermitian, so we can write its decomposition as

$$c = U\hat{\lambda}U^{-1}$$

where $\hat{\lambda}$ is a diagonal matrix of eigenvalues and U is a matrix of eigenstates: this is just diagonalisation of matrix c.

Eigenvalues are 0 or 1, with eigenstates

$$|\Phi_0\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \qquad |\Phi_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

so the matrix of eigenstates is

$$U = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \qquad U^{-1} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

Thus:

$$c = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

In terms of the Schmidt decomposition notation, $V = U^{-1}$, and the basis states are given by $|\tilde{\imath}_A\rangle = \sum_k U_{k,i}\,|k\rangle$, $|\tilde{\imath}_B\rangle = \sum_m U^{-1}_{i,m}\,|m\rangle$. Write

$$U = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \equiv \begin{pmatrix} U_{0,0} & U_{0,1} \\ U_{1,0} & U_{1,1} \end{pmatrix}$$

so

$$\left|\tilde{0}_A\right\rangle = U_{0,0}|0\rangle + U_{1,0}|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

$$\left|\tilde{1}_A\right\rangle = U_{0,1}|0\rangle + U_{1,1}|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

$$\left|\tilde{0}_B\right\rangle = U^{-1}_{0,0}|0\rangle + U^{-1}_{0,1}|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

$$\left|\tilde{1}_B\right\rangle = U^{-1}_{1,0}|0\rangle + U^{-1}_{1,1}|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

Hence the decomposition is

$$|\psi\rangle = \sum_i \lambda_i\,|\tilde{\imath}_A\rangle|\tilde{\imath}_B\rangle = \left|\tilde{1}_A\right\rangle\left|\tilde{1}_B\right\rangle$$

$$= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)_A \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)_B$$

This state is separable as there is only one non-zero eigenstate; the Schmidt number $R = 1$, concurrence $C = 0$, and the entanglement entropy $\mathcal{E} = 0$.

Different formalisms and/or interpretations of non-relativistic quantum mechanics

This list is not exhaustive. For more information see M. Jammer, *"The Philosophy of Quantum Mechanics"* (Wiley, 1974) – an old book, but quite good.

- "Copenhagen interpretation": the mainstream view taught here. Formalism based on states and operators in a Hilbert space, commutation of operators etc. Interpretation of a probabilistic view, collapse of wave function on measurement etc. This suffers from the "measurement problem" (what is wave function collapse?) which motivates a lot of alternative approaches.
- "Feynmann path integral": time evolution is expressed as a summation over an infinite continuum of possible paths, each one weighted by its classical action. Essentially a different (but equivalent) mathematical formalism to the mainstream methodology using operators and commutators.
- "Local hidden variables": following Einstein, Podolsky, Rosen (EPR), theories of local realism that say all properties of a particle should be

determined and independent of particles at very large distances away.

- “de Broglie–Bohm theory” or “pilot wave theory”: David Bohm developed a deterministic formalism (D. Bohm, Phys. Rev. **85**, 166 (1952); Phys. Rev. **85**, 180 (1952)) which he originally described as using “hidden variables” but is also non-local.
- “Many-worlds interpretation”: by Hugh Everett, III, in 1957 (H. Everett, III, Rev. Mod. Phys. **29**, 454 (1957)). Essentially an attempt to generalise the Copenhagen interpretation and circumvent the measurement problem: *“The whole issue of the transition from “possible” to “actual” is taken care of in the theory in a very simple way—there is no such transition, nor is such a transition necessary for the theory to be in accord with our experience. From the viewpoint of the theory all elements of a superposition (all “branches”) are “actual,” none any more “real” than the rest. It is unnecessary to suppose that all but one are somehow destroyed, since all the separate elements of a superposition individually obey the wave equation with complete indifference to the presence or absence (“actuality” or not) of any other elements. This total lack of effect of one branch on another also*

implies that no observer will ever be aware of any "splitting" process. Arguments that the world picture presented by this theory is contradicted by experience, because we are unaware of any branching process, are like the criticism of the Copernican theory that the mobility of the earth as a real physical fact is incompatible with the common sense interpretation of nature because we feel no such motion."

- There are mathematical formalisms to deal with systems of many interacting identical particles e.g. a sea of interacting electrons. The mainstream approach of operators and commutators can be generalised to "second quantisation" (aka "occupation number" approach), and Feynmann path integrals are generalised to many-body functional field integrals.