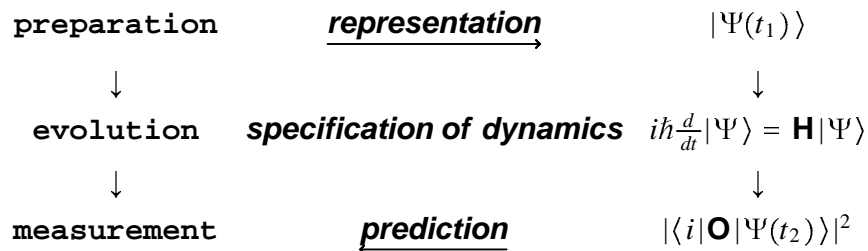


Some examples and concept-reinforcement

1. The big picture: preparation, evolution, and measurement
2. The density operator – what *is* this thing?
3. Working with tensor products

The big picture / what's a density operator?

- ! The essential function of physical theory is to make quantitative predictions about experiments.



- State vectors $|\Psi\rangle$ and density operators ρ are mathematical objects that summarize our knowledge about the state (preparation+subsequent evolution) of a physical system.

Given that we fix a complete orthonormal basis for the Hilbert space, kets can be represented as column vectors

$$|\Psi\rangle = c_0|0\rangle + c_1|1\rangle \leftrightarrow \begin{pmatrix} c_0 \\ c_1 \end{pmatrix}, \quad 1$$

bras can be represented as row vectors

$$\langle\Psi| = c_0^*\langle 0| + c_1^*\langle 1| \leftrightarrow \begin{pmatrix} c_0^* & c_1^* \end{pmatrix}, \quad 2$$

and density operators can be represented as square matrices

$$\rho = |\Psi\rangle\langle\Psi| = (c_0|0\rangle + c_1|1\rangle)(c_0^*\langle 0| + c_1^*\langle 1|)$$

$$\leftrightarrow \begin{pmatrix} c_0 \\ c_1 \end{pmatrix} \begin{pmatrix} c_0^* & c_1^* \end{pmatrix} = \begin{pmatrix} |c_0|^2 & c_0c_1^* \\ c_0^*c_1 & |c_1|^2 \end{pmatrix}.$$

Pretty compact!

- The use of representing knowledge in this way is that it efficiently allows us to make statistical predictions about the outcome of **any** measurement that could be made on the system.

There is a distinct standard measurement *procedure* corresponding to every complete set of orthogonal projectors. We say that measurement A is 'distinct' from measurement B if there exists some preparation of the system such that A and B have different measurement statistics.

There are an uncountable number of distinct measurements (even in two Hilbert-space dimensions), so $|\Psi\rangle$ and ρ are pretty compact ways to summarize the statistics of *all* of them!

- State vectors can be used when the preparation and evolution of a system are as sharply defined as possible – in such cases, we have only **intrinsic** uncertainties regarding the outcome of future measurements.
- Density operators must be used when there are **extrinsic** uncertainties in the problem, or when the system of interest is entangled with an inaccessible system.
- Intrinsic uncertainties are associated with coherent superposition of eigenstates, while extrinsic uncertainties arise when we are not sure which of several preparations or dynamics has actually occurred.

Recall that for essentially all states in a Hilbert space,

$$\Delta\mathbf{O}_q = \langle |\mathbf{O}_q - \langle\mathbf{O}_q\rangle| \rangle = \sqrt{\langle\mathbf{O}_q^2\rangle - \langle\mathbf{O}_q\rangle^2} > 0$$

4

for every observable (physical quantity) \mathbf{O}_q except those that are proportional to the identity operator.

In other words, for any given state in the Hilbert space there are plenty of questions one can ask to which there is no definite answer! Contrast this to the classical situation – if you know the position and momentum of a classical point particle, there's no measurement I could perform for which you can't predict the exact answer.

The only way to have $\Delta\mathbf{O}_q = 0$ is to prepare the system in an eigenstate of \mathbf{O}_q . Any time we have a *superposition* of eigenstates corresponding to distinct eigenvalues of \mathbf{O}_q , we have **intrinsic** quantum-mechanical uncertainty.

A simple example of extrinsic uncertainty would be if we were unsure which of two pure

states corresponded to the actual preparation of a physical system. That is, if we were unsure which of two preparation procedures was actually performed.

Let's say we're working in a two-dimensional Hilbert space, just to be concrete. Suppose we have a mixed ensemble of quantum states, with only two members:

$$\begin{aligned} p_0 &: |\Psi_0\rangle, \\ p_1 &: |\Psi_1\rangle. \end{aligned} \tag{5}$$

If we want to make predictions about the outcome of a measurement $\{\mathbf{P}_x, \mathbf{P}_y\}$, we have two choices:

$$\Pr(x) = p_0 \langle \Psi_0 | \mathbf{P}_x | \Psi_0 \rangle + p_1 \langle \Psi_1 | \mathbf{P}_x | \Psi_1 \rangle, \tag{6}$$

or

$$\Pr(x) = \text{Tr}[\rho \mathbf{P}_x], \tag{7}$$

where $\rho = p_0 |\Psi_0\rangle\langle\Psi_0| + p_1 |\Psi_1\rangle\langle\Psi_1|$.

For this two-member ensemble, it's not clear that either choice is better than the other. But what if we have, e.g., a fifty-member ensemble? We can either compute fifty expectation values every time someone gives us a projection operator, or we can simply take the trace of that projector with the 2×2 density operator

$$\rho = \sum_{i=1}^{50} p_i |\Psi_i\rangle\langle\Psi_i|. \tag{8}$$

Surely ρ is a more compact representation of our knowledge about the system state! We need only compute it once, after all, to simplify an arbitrarily large number of statistical predictions. The same advantage applies when we want to evolve the system state...

Recall that many different ensembles have the same density operator. What this tells us is that there is no measurement that can distinguish between two mixed ensembles of quantum states that have exactly the same density operator.

- Coherent superpositions can lead to interference effects; probabilistic mixtures cannot.

Consider the following two states:

$$\begin{aligned} |\Psi_A\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \\ |\Psi_B\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \end{aligned} \tag{9}$$

Clearly each of these states has $\langle \mathbf{P}_1 \rangle = \frac{1}{2}$. Even in a situation where we are not sure which of these two quantum states has been prepared, we are sure that $\langle \mathbf{P}_1 \rangle$ should still be $\frac{1}{2}$. In other words, the two possible scenarios should not 'interfere' with one another when we make predictions about measurement results.

Suppose we know that $|\Psi_A\rangle$ and $|\Psi_B\rangle$ are prepared with equal probability. Let's convince ourselves that this does not correspond to a coherent superposition of $|\Psi_A\rangle$ and $|\Psi_B\rangle$:

$$\begin{aligned}
|\Psi_{coh}\rangle &= \frac{1}{\sqrt{2}}(|\Psi_A\rangle + |\Psi_B\rangle) \\
&= \frac{1}{2}(|0\rangle + |1\rangle + |0\rangle - |1\rangle) \\
&= |0\rangle.
\end{aligned}
\tag{10}$$

Coherent superpositions **do** exhibit interference, and this is why we can't use them to represent *extrinsic* uncertainty about the preparation of a quantum system – hence the need for density operators!

- Entanglement with an **inaccessible** system has the same type of effect on our ability to predict measurement outcomes as does the presence of extrinsic uncertainties.

Another reason we need to have density operators is the phenomenon of entanglement. Recall that there exist nonfactorizable pure states in the joint Hilbert space of two subsystems. For example,

$$\begin{aligned}
|\Psi_{AB}\rangle &= \frac{1}{\sqrt{2}}(|0_A\rangle \otimes |0_B\rangle + |1_A\rangle \otimes |1_B\rangle) \\
&\neq |\Psi_A\rangle \otimes |\Psi_B\rangle.
\end{aligned}
\tag{11}$$

There is no way to assign a pure state to system *A* or *B* individually.

Suppose our friend Charlie comes into our lab and takes away system *B*, after $|\Psi_{AB}\rangle$ has been prepared. Clearly we can still make measurements on system *A*. It is also true, although perhaps not entirely obvious, that the statistics of any measurements we might choose to make on system *A* will be independent of whatever Charlie happens to do with system *B* – we assume that *A* and *B* can no longer physically interact after they have been separated.

Hence we would like to have a compact representation of everything we know about system *A* alone, starting from the statement that the initial joint state of the *AB* system was $|\Psi_{AB}\rangle$. What do we mean by compact? It may not be entirely clear if both *A* and *B* are two-dimensional systems – pure states in H_{AB} have four complex coefficients, the same as a reduced density operator on H_A alone. But consider the more general entangled state,

$$\begin{aligned}
|\Psi_{AB}\rangle &= \frac{1}{\sqrt{2}}(|0_A\rangle \otimes |\Psi_B^0\rangle + |1_A\rangle \otimes |\Psi_B^1\rangle), \\
\langle \Psi_B^0 | \Psi_B^1 \rangle &= 0,
\end{aligned}
\tag{12}$$

when H_A is still two-dimensional but H_B is (for example) fifty-dimensional. Then state vectors in H_{AB} have one hundred complex coefficients, but the reduced density operator still only has four.

Let's get back to our example with $N_A = N_B = 2$. One of the things that Charlie might choose to do with system *B* is to perform a measurement, say of the set $\{\mathbf{P}_0^B, \mathbf{P}_1^B\}$. Working from the initial state in the joint Hilbert space, we know that the outcome probabilities will be

$$\begin{aligned}
\Pr(0_B) &= \langle \Psi_{AB} | \mathbf{1}^A \otimes \mathbf{P}_0^B | \Psi_{AB} \rangle \\
&= 0.5, \\
\Pr(1_B) &= 0.5.
\end{aligned}
\tag{13}$$

Likewise, the corresponding post-measurement states of the joint system will be

$$\begin{aligned} |\Psi_{AB}\rangle &\mapsto \frac{\mathbf{1}^A \otimes \mathbf{P}_0^B |\Psi_{AB}\rangle}{\sqrt{\langle \Psi_{AB} | \mathbf{1}^A \otimes \mathbf{P}_0^B | \Psi_{AB} \rangle}} \\ &= |0_A\rangle \otimes |0_B\rangle \end{aligned} \quad 14$$

or

$$\begin{aligned} |\Psi_{AB}\rangle &\mapsto \frac{\mathbf{1}^A \otimes \mathbf{P}_1^B |\Psi_{AB}\rangle}{\sqrt{\langle \Psi_{AB} | \mathbf{1}^A \otimes \mathbf{P}_1^B | \Psi_{AB} \rangle}} \\ &= |1_A\rangle \otimes |1_B\rangle \end{aligned} \quad 15$$

But we see that the post-measurement state of system A alone is definitely either $|0_A\rangle$ or $|1_A\rangle$, each with 50% probability. **If Charlie doesn't tell us the measurement result**, we are left with a mixed ensemble of quantum states for A. The density operator that represents this ensemble is

$$\rho_A = \frac{1}{2}(|0_A\rangle\langle 0_A| + |1_A\rangle\langle 1_A|), \quad 16$$

which is in fact identical to the reduced density operator

$$\begin{aligned} \tilde{\rho}_A &= \text{Tr}_B [|\Psi_{AB}\rangle\langle \Psi_{AB}|] && 17 \\ &= \frac{1}{2} \text{Tr}_B [(|0_A 0_B\rangle + |1_A 1_B\rangle)(\langle 0_A 0_B| + \langle 1_A 1_B|)] && \# \\ &= \frac{1}{2} \text{Tr}_B [|0_A 0_B\rangle\langle 0_A 0_B| + |0_A 0_B\rangle\langle 1_A 1_B| + |1_A 1_B\rangle\langle 0_A 0_B| + |1_A 1_B\rangle\langle 1_A 1_B|] && \# \\ &= \frac{1}{2} \left\{ \begin{array}{l} \langle 0_B | [|0_A 0_B\rangle\langle 0_A 0_B| + |0_A 0_B\rangle\langle 1_A 1_B| + |1_A 1_B\rangle\langle 0_A 0_B| + |1_A 1_B\rangle\langle 1_A 1_B|] |0_B\rangle \\ + \langle 1_B | [|0_A 0_B\rangle\langle 0_A 0_B| + |0_A 0_B\rangle\langle 1_A 1_B| + |1_A 1_B\rangle\langle 0_A 0_B| + |1_A 1_B\rangle\langle 1_A 1_B|] |1_B\rangle \end{array} \right\} && \# \\ &= \frac{1}{2} \{|0_A\rangle\langle 0_A| + |1_A\rangle\langle 1_A|\}. && \# \end{aligned}$$

What if Charlie decides to perform the measurement corresponding to $\{\mathbf{P}_x^B, \mathbf{P}_y^B\}$, where

$$\begin{aligned} \mathbf{P}_x^B &= \frac{1}{2}(|0_B\rangle\langle 0_B| + |0_B\rangle\langle 1_B| + |1_B\rangle\langle 0_B| + |1_B\rangle\langle 1_B|), && 18 \\ \mathbf{P}_y^B &= \frac{1}{2}(|0_B\rangle\langle 0_B| - |0_B\rangle\langle 1_B| - |1_B\rangle\langle 0_B| + |1_B\rangle\langle 1_B|). \end{aligned}$$

Then $\langle \mathbf{1}^A \otimes \mathbf{P}_x^B \rangle = \langle \mathbf{1}^A \otimes \mathbf{P}_y^B \rangle = 0.5$ still, but the resulting post-measurement states are

$$\begin{aligned}
x : |\Psi_{AB}\rangle &\mapsto \mathbf{1}^A \otimes \mathbf{P}_x^B |\Psi_{AB}\rangle \\
&\propto \mathbf{1}^A \otimes \mathbf{P}_x^B \frac{1}{\sqrt{2}} (|0_A\rangle \otimes |0_B\rangle + |1_A\rangle \otimes |1_B\rangle) \quad \# \\
&\propto \mathbf{1}^A \otimes \frac{1}{2} (|0_B\rangle\langle 0_B| + |0_B\rangle\langle 1_B| + |1_B\rangle\langle 0_B| + |1_B\rangle\langle 1_B|) \frac{1}{\sqrt{2}} (|0_A\rangle \otimes |0_B\rangle + |1_A\rangle \otimes |1_B\rangle) \quad \# \\
&\propto \{|0_A\rangle \otimes |0_B\rangle + |1_A\rangle \otimes |0_B\rangle + |0_A\rangle \otimes |1_B\rangle + |1_A\rangle \otimes |1_B\rangle\} \quad \# \\
&= \frac{1}{2} (|0_A\rangle \otimes |0_B\rangle + |0_A\rangle \otimes |1_B\rangle + |1_A\rangle \otimes |0_B\rangle + |1_A\rangle \otimes |1_B\rangle) \quad \# \\
&= \frac{1}{2} (|0_A\rangle + |1_A\rangle) \otimes (|0_B\rangle + |1_B\rangle), \\
y : |\Psi_{AB}\rangle &\mapsto \frac{1}{2} (|0_A\rangle \otimes |0_B\rangle - |0_A\rangle \otimes |1_B\rangle - |1_A\rangle \otimes |0_B\rangle + |1_A\rangle \otimes |1_B\rangle) \\
&= \frac{1}{2} (|0_A\rangle - |1_A\rangle) \otimes (|0_B\rangle - |1_B\rangle). \quad 19
\end{aligned}$$

Hence the post-measurement state of system A considered alone will be either $\frac{1}{\sqrt{2}}(|0_A\rangle + |1_A\rangle)$ or $\frac{1}{\sqrt{2}}(|0_A\rangle - |1_A\rangle)$ with $p = 0.5$, and the density operator for this ensemble is still

$$\begin{aligned}
\tilde{\rho}_A &= \frac{1}{4} (|0_A\rangle + |1_A\rangle)(\langle 0_A| + \langle 1_A|) + \frac{1}{4} (|0_A\rangle - |1_A\rangle)(\langle 0_A| - \langle 1_A|) \\
&= \frac{1}{2} (|0_A\rangle\langle 0_A| + |1_A\rangle\langle 1_A|), \quad 20
\end{aligned}$$

which is once again our good old reduced density operator for A .

We thus see that taking a partial trace over subsystem B is like performing an ‘imaginary’ complete measurement on B and then ignoring the result.

Finally, recall that a pair of systems that starts out in a factorizable initial state can evolve into an entangled state via Hamiltonian evolution. Here’s a simple example.

Let’s work again with our same two two-dimensional quantum systems. Let the initial state be

$$\begin{aligned}
|\Psi_{AB}(0)\rangle &= \frac{1}{2} (|0_A\rangle + |1_A\rangle) \otimes (|0_B\rangle + |1_B\rangle) \\
&= \frac{1}{2} (|0_A\rangle \otimes |0_B\rangle + |0_A\rangle \otimes |1_B\rangle + |1_A\rangle \otimes |0_B\rangle + |1_A\rangle \otimes |1_B\rangle), \quad 21
\end{aligned}$$

and suppose the Hamiltonian is

$$\mathbf{H} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad 22$$

where the ordering of basis states in H_{AB} is

$$|0_A\rangle \otimes |0_B\rangle, \quad |0_A\rangle \otimes |1_B\rangle, \quad |1_A\rangle \otimes |0_B\rangle, \quad |1_A\rangle \otimes |1_B\rangle. \quad 23$$

The unitary evolution operator is

$$\mathbf{T}(t, 0) = \exp\left(\frac{-i}{\hbar} \mathbf{H}t\right). \quad 24$$

After a time $t = \frac{1}{2}\pi\hbar$,

$$\begin{aligned} \mathbf{T}(t,0) &= \begin{pmatrix} e^{-i\pi/2} & 0 & 0 & 0 \\ 0 & e^{-i\pi/2} & 0 & 0 \\ 0 & 0 & e^{-i\pi/2} & 0 \\ 0 & 0 & 0 & e^{+i\pi/2} \end{pmatrix} \\ &= \begin{pmatrix} -i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \end{pmatrix}, \end{aligned} \quad 25$$

and

$$\begin{aligned} |\Psi_{AB}(t)\rangle &= \mathbf{T}(t,0)|\Psi_{AB}(0)\rangle \\ &= \frac{i}{2}(|0_A\rangle \otimes |0_B\rangle + |0_A\rangle \otimes |1_B\rangle + |1_A\rangle \otimes |0_B\rangle - |1_A\rangle \otimes |1_B\rangle), \end{aligned} \quad 26$$

which is an entangled state.

How do we know that this is an entangled state? Probably by inspection, but let's also check the trace of $\tilde{\rho}_A^2$. First off, the joint density operator is

$$\begin{aligned} \rho_{AB} &= |\Psi_{AB}(t)\rangle\langle\Psi_{AB}(t)| \\ &= \frac{1}{4}(|0_A0_B\rangle\langle0_A0_B| + |0_A0_B\rangle\langle0_A1_B| + |0_A0_B\rangle\langle1_A0_B| - |0_A0_B\rangle\langle1_A1_B| \\ &\quad + |0_A1_B\rangle\langle0_A0_B| + |0_A1_B\rangle\langle0_A1_B| + |0_A1_B\rangle\langle1_A0_B| - |0_A1_B\rangle\langle1_A1_B| \\ &\quad + |1_A0_B\rangle\langle0_A0_B| + |1_A0_B\rangle\langle0_A1_B| + |1_A0_B\rangle\langle1_A0_B| - |1_A0_B\rangle\langle1_A1_B| \\ &\quad - |1_A1_B\rangle\langle0_A0_B| - |1_A1_B\rangle\langle0_A1_B| - |1_A1_B\rangle\langle1_A0_B| + |1_A1_B\rangle\langle1_A1_B|). \end{aligned} \quad 27$$

Next we take the partial trace over B :

$$\begin{aligned} \tilde{\rho}_A &= \text{Tr}_B [\rho_{AB}] \\ &= \langle 0_B | \rho_{AB} | 0_B \rangle + \langle 1_B | \rho_{AB} | 1_B \rangle \\ &= \frac{1}{4}(|0_A\rangle\langle 0_A| + |0_A\rangle\langle 1_A| + |1_A\rangle\langle 0_A| + |1_A\rangle\langle 1_A| \\ &\quad + |0_A\rangle\langle 0_A| - |0_A\rangle\langle 1_A| - |1_A\rangle\langle 1_A| + |1_A\rangle\langle 1_A|) \\ &= \frac{1}{2}(|0_A\rangle\langle 0_A| + |1_A\rangle\langle 1_A|) \\ &= \frac{1}{2}\mathbf{1}^A. \end{aligned} \quad 28$$

Hence $\tilde{\rho}_A^2 = \frac{1}{4}\mathbf{1}^A$, and $\text{Tr}[\tilde{\rho}_A^2] = \frac{1}{2}$, which is clearly less than one. Since density operators that correspond to pure states are projectors, we conclude that no pure state can be assigned to subsystem A when the joint state of the AB system is $|\Psi_{AB}(t)\rangle$.

- The formal rules for manipulating state vectors $|\Psi\rangle$ are different from those for manipulating density operators ρ , but these two mathematical objects play identical **roles** in quantum theory.

- Mutually exclusive measurement-outcomes correspond to orthogonal projection operators $\{\mathbf{P}_0, \mathbf{P}_1, \dots\}$, and the probability of a particular outcome i is given by $|\mathbf{P}_i \Psi|^2$.

Say I have N copies of a two-dimensional system prepared in the state

$$|\Psi\rangle = c_0|0\rangle + c_1|1\rangle, \quad 29$$

and that I want to perform the measurement corresponding to $\{\mathbf{P}_0 = |0\rangle\langle 0|, \mathbf{P}_1 = |1\rangle\langle 1|\}$ on every copy of the system.

Each time I measure one copy, the result I get is **random** – it could either be 0 or 1.

What quantum measurement theory allows me to predict is the **relative frequency** with which I will obtain 0's and 1's in this random string of results:

$$\begin{aligned} n_0 &= N\langle \mathbf{P}_0 \rangle = N|c_0|^2, \\ n_1 &= N\langle \mathbf{P}_1 \rangle = N|c_1|^2, \\ n_0 + n_1 &= N(|c_0|^2 + |c_1|^2) = N. \end{aligned} \quad 30$$

Hence in any given single measurement along the way, we say that the probability to get outcome i is $\text{Pr}(i) = n_i/N = \langle \mathbf{P}_i \rangle$.

Working with tensor products

Let's work with our favorite example of two two-dimensional Hilbert spaces H_A and H_B , with given complete orthonormal bases $\{|0_A\rangle, |1_A\rangle\}$ and $\{|0_B\rangle, |1_B\rangle\}$. Let's also choose the simple tensor-product basis for H_{AB} , $\{|0_A 0_B\rangle, |0_A 1_B\rangle, |1_A 0_B\rangle, |1_A 1_B\rangle\}$.

Suppose we are given vectors $|\Psi_A\rangle \in H_A$ and $|\Psi_B\rangle \in H_B$:

$$\begin{aligned} |\Psi_A\rangle &= a_0|0_A\rangle + a_1|1_A\rangle \leftrightarrow \begin{pmatrix} a_0 \\ a_1 \end{pmatrix}, \\ |\Psi_B\rangle &= b_0|0_B\rangle + b_1|1_B\rangle \leftrightarrow \begin{pmatrix} b_0 \\ b_1 \end{pmatrix}. \end{aligned} \quad 31$$

Then $|\Psi_A \Psi_B\rangle \in H_{AB}$ has the vector representation

$$\begin{aligned} |\Psi_A\rangle \otimes |\Psi_B\rangle &= (a_0|0_A\rangle + a_1|1_A\rangle) \otimes (b_0|0_B\rangle + b_1|1_B\rangle) \\ &= a_0b_0|0_A 0_B\rangle + a_0b_1|0_A 1_B\rangle + a_1b_0|1_A 0_B\rangle + a_1b_1|1_A 1_B\rangle \\ &\leftrightarrow \begin{pmatrix} a_0b_0 \\ a_0b_1 \\ a_1b_0 \\ a_1b_1 \end{pmatrix}. \end{aligned} \quad 32$$

Likewise,

$$\langle \Psi_A \Psi_B | \leftrightarrow \begin{pmatrix} a_0^*b_0^* & a_0^*b_1^* & a_1^*b_0^* & a_1^*b_1^* \end{pmatrix}. \quad 33$$

Moving on to operators, let's compute a matrix representation for $\sigma_x^A \otimes \sigma_x^B$, where

$\sigma_x = |0\rangle\langle 1| + |1\rangle\langle 0|$. So

$$\begin{aligned}\sigma_x^A \otimes \sigma_x^B &= (|0_A\rangle\langle 1_A| + |1_A\rangle\langle 0_A|) \otimes (|0_B\rangle\langle 1_B| + |1_B\rangle\langle 0_B|) \\ &= |0_A 0_B\rangle\langle 1_A 1_B| + |0_A 1_B\rangle\langle 1_A 0_B| + |1_A 0_B\rangle\langle 0_A 1_B| + |1_A 1_B\rangle\langle 0_A 0_B| \\ &\leftrightarrow \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.\end{aligned}\tag{34}$$

Given the ordering we have chosen for the basis of H_{AB} , we could have also used

$$\mathbf{A} \otimes \mathbf{B} \leftrightarrow \begin{pmatrix} a_{00}B & a_{01}B \\ a_{10}B & a_{11}B \end{pmatrix},\tag{35}$$

where in this case

$$\begin{aligned}A &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ \mathbf{A} \otimes \mathbf{B} &\leftrightarrow \begin{pmatrix} 0 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & 1 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ 1 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & 0 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.\end{aligned}\tag{36}$$

More about the mysteries of entanglement

Entanglement vs. classical correlation

Still working with two-dimensional H_A and H_B , etc.

Consider the joint density operator

$$\rho_{AB} = \frac{1}{2}(|0_A 0_B\rangle\langle 0_A 0_B| + |1_A 1_B\rangle\langle 1_A 1_B|).\tag{37}$$

This corresponds to a mixed preparation where, with equal probability, the joint pure state is either $|0_A 0_B\rangle$ or $|1_A 1_B\rangle$. Hence, we don't know whether A or B is $|0\rangle$ or $|1\rangle$, but we do know that **they have to be the same**. Formally,

$$\begin{aligned}\langle \mathbf{P}_0^A \otimes \mathbf{P}_0^B \rangle &= 0.5, & \langle \mathbf{P}_0^A \otimes \mathbf{P}_1^B \rangle &= 0, \\ \langle \mathbf{P}_1^A \otimes \mathbf{P}_0^B \rangle &= 0, & \langle \mathbf{P}_1^A \otimes \mathbf{P}_1^B \rangle &= 0.5.\end{aligned}\tag{38}$$

This kind of situation, where some property of A is random but nonetheless tied to some property of B , is called 'classical' or 'probabilistic' correlation.

Now consider the entangled pure state

$$|\Psi_{AB}\rangle = \frac{1}{\sqrt{2}}(|0_A 0_B\rangle + |1_A 1_B\rangle). \quad 39$$

It is still the case that

$$\begin{aligned} \langle \mathbf{P}_0^A \otimes \mathbf{P}_0^B \rangle &= 0.5, & \langle \mathbf{P}_0^A \otimes \mathbf{P}_1^B \rangle &= 0, \\ \langle \mathbf{P}_1^A \otimes \mathbf{P}_0^B \rangle &= 0, & \langle \mathbf{P}_1^A \otimes \mathbf{P}_1^B \rangle &= 0.5. \end{aligned} \quad 40$$

Is there any difference between the type of correlation embodied in $|\Psi_{AB}\rangle$ and that embodied in ρ_{AB} ?

Consider a new set of projectors $\{\mathbf{P}_x^A \mathbf{P}_x^B, \mathbf{P}_x^A \mathbf{P}_y^B, \mathbf{P}_y^A \mathbf{P}_x^B, \mathbf{P}_y^A \mathbf{P}_y^B\}$, where

$$\begin{aligned} \mathbf{P}_x &= |x\rangle\langle x|, \\ |x\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \\ \mathbf{P}_y &= |y\rangle\langle y| = \mathbf{1} - \mathbf{P}_x, \\ |y\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \end{aligned} \quad 41$$

Then for the mixed initial state,

$$\begin{aligned} \langle \mathbf{P}_x^A \otimes \mathbf{P}_y^B \rangle &= \langle x_{AYB} | \rho_{AB} | x_{AYB} \rangle \\ &= \frac{1}{2} \langle x_{AYB} | (|0_A 0_B\rangle\langle 0_A 0_B| + |1_A 1_B\rangle\langle 1_A 1_B|) | x_{AYB} \rangle \\ &= \frac{1}{2} (|\langle x_{AYB} | 0_A 0_B \rangle|^2 + |\langle x_{AYB} | 1_A 1_B \rangle|^2) \\ &= \frac{1}{2} \left(\frac{1}{4} + \frac{1}{4} \right) = \frac{1}{4}. \end{aligned} \quad 42$$

Hence we see that the correlation between A and B is not as strong in the x, y basis as in the $0, 1$ basis, when the initial state is ρ_{AB} . What about $|\Psi_{AB}\rangle$?

$$\begin{aligned} \langle \mathbf{P}_x^A \otimes \mathbf{P}_y^B \rangle &= |\langle x_{AYB} | \Psi_{AB} \rangle|^2 \\ &= \frac{1}{\sqrt{2}} |\langle x_{AYB} | (|0_A 0_B\rangle + |1_A 1_B\rangle) \rangle|^2 \\ &= \frac{1}{\sqrt{2}} \left| \frac{1}{2} - \frac{1}{2} \right|^2 \\ &= 0. \end{aligned}$$

$$\langle \mathbf{P}_y^A \otimes \mathbf{P}_x^B \rangle = |\langle y_{AXB} | \Psi_{AB} \rangle|^2 = 0. \quad 43$$

For the entangled state, A and B remain perfectly correlated even under change of measurement basis! As my friend Chris Fuchs likes to say, entanglement is all-purpose correlation.

Entanglement is a very special property of composite *quantum* systems. As you might guess from even this simple example, there are scenarios in which entanglement can be 'utilized' to perform otherwise-impossible tasks in communication and information processing. We'll begin to see some of this in subsequent lectures on Bell Inequalities and quantum information theory.

However, the tensor-product rule for representation of joint states also places some limitations on what we can do with composite quantum systems, as we'll see in our

discussions of quantum measurement and of the no-cloning theorem.

Nonlocality and Bell Inequalities

(Based on the discussion in Chris Isham's book, *Lectures on Quantum Theory: Mathematical and Structural Foundations* (Imperial College Press, 1995).)

Say we have two experimenters, Alice and Bob, whose labs are located many kilometers apart. Their labs are basically identical, actually, each consisting of one particle 'detector' that has one meter, one switch, and a bell. The meter is for reading out the result of a measurement (which we assume to be either ± 1), while the switch is used to select which of two types of measurements the experimenter would like to make. On Alice's side we'll label the two possibilities A and A' , and on Bob's side B and B' . The bell rings each time a particle hits the detector, letting the experimenter know when he or she can read out the result of his/her selected measurement.

So where do these particles come from? Midway between Alice's lab and Bob's there is a 'pair source.' This source always produces particles in pairs, sending one to Alice and the other to Bob. We assume that the particles have some internal degree of freedom, which is what Alice's and Bob's detectors are designed to measure. The pair source prepares the internal states of the particles in some unknown, possibly random fashion.

The 'experiment' consists of the following procedure. The source prepares and emits one pair of particles per unit of time, so Alice and Bob know that they may expect to receive particles at a regular rate. Once per unit time, they each (independently) select a random setting for their switch, wait for their bell to ring, and then read off and write down the measurement result.

Hence after ten rounds, e.g., Alice's and Bob's lab books might look something like this:

Alice	Bob
$A -1$	$B' -1$
$A +1$	$B' -1$
$A' +1$	$B +1$
$A +1$	$B +1$
$A' -1$	$B -1$
$A' -1$	$B' +1$
$A +1$	$B -1$
$A' -1$	$B' +1$
$A +1$	$B' +1$
$A -1$	$B +1$

1

Although this experimental scenario seems extremely general, it turns out that we have already specified enough to derive some important predictions about the statistics of Alice's and Bob's measurement records!

Let's start by making some reasonable assumptions about the overall behavior of the experiment:

1. **Local determinism** – we might like to believe that the result of Alice's measurement (either A or A') is *locally* determined by the physical state of the particle she receives from the pair source. It should not depend on the state of Bob's particle, since in this scenario Bob could be really far away! And the result of Alice's measurement certainly should not depend on Bob's choice of measurement – that is, whether Alice's meter reads $+1$ or -1 should not depend on whether Bob has his switch set to B or B' ...
2. **Objective reality** – Even though Alice (and Bob) must choose to make one measurement or the other (A or A') on any given particle, each particle 'knows' what its value is for both measurements. That is, sufficient information to determine the outcome of either measurement is encoded in the internal state of each particle.

Under these assumptions, we can write down the following model for this experiment. In each round, the pair source produces a pair of particles with the following information encoded in their internal states:

$$A_n = \pm 1, \quad A'_n = \pm 1, \quad B_n = \pm 1, \quad B'_n = \pm 1. \quad 2$$

Here the four possible measurement labels are treated as random variables, with the subscript labelling the round. As a logical consequence of local determinism and objective realism, we can assume the existence of a *joint probability distribution* $P(A, A', B, B')$. Hence, it should be meaningful to consider correlation functions of all four random variables simultaneously, and these correlation functions should be measurable by Alice and Bob.

Consider the following function of the random variables,

$$g_n = A_n B_n + A'_n B_n + A_n B'_n - A'_n B'_n. \quad 3$$

Were we to tabulate the 16 possible values of g_n , we would magically find that $g_n = \pm 2$. However, an easier way to see this is to note that the last term in the sum is equal to the product of the first three, since $A_n^2 = (A'_n)^2 = B_n^2 = (B'_n)^2 = +1$:

$$\begin{aligned} A'_n B'_n &= (A_n B_n)(A'_n B_n)(A_n B'_n) \\ &= A_n^2 B_n^2 A'_n B'_n. \end{aligned} \quad 4$$

Then if $A'_n B'_n = +1$, the set $\{A_n B_n, A'_n B_n, A_n B'_n\}$ has either zero or two -1 's, hence $g_n = A_n B_n + A'_n B_n + A_n B'_n - A'_n B'_n$ must be either $+2$ or -2 . If on the other hand $A'_n B'_n = -1$, the set must have either zero or two $+1$'s, hence g_n must be either -2 or $+2$.

In any case, it follows that

$$\left| \frac{1}{N} \sum_{n=1}^N g_n \right| = \frac{1}{N} \left| \sum_{n=1}^N A_n B_n + \sum_{n=1}^N A'_n B_n + \sum_{n=1}^N A_n B'_n - \sum_{n=1}^N A'_n B'_n \right| \leq 2.$$

5

This is one form (due to Clauser, Horne, Shimony, and Holt) of Bell's famous inequality.

It should be noted that at this point, all we have relied on in our derivation is basic probability theory! Hence the Bell Inequality is a **model-independent** prediction about measurement statistics in a world that is locally deterministic and allows objective realism.

Hence experimental violations of the Inequality actually tell us something about Nature, not just quantum theory!

As it turns out, one can actually go to the lab and perform experiments of precisely the type described above, and find that this inequality is strongly violated! For example, see

- G. Weihs *et al*, "Violation of Bell's Inequality under Strict Einstein Locality Conditions," Phys. Rev. Lett. **81**, 5039-5043 (1998);
- W. Tittel *et al*, "Violation of Bell Inequalities by Photons More Than 10 km Apart," Phys. Rev. Lett. **81**, 3563-3566 (1998);
- A. Aspect, "Bell's inequality test: more ideal than ever," Nature **398**, 189-190 (1999).

In experiments of this type, the key is to construct a source that produces pairs of photons an *entangled* state such as

$$|\Psi_{ab}\rangle = \frac{1}{\sqrt{2}}(|0_a 1_b\rangle - |1_a 0_b\rangle).$$

6

In each round of the experiment, Alice's two measurements correspond to the observables $\mathbf{A} = \sigma_z^a$ and $\mathbf{A}' = \cos \phi \sigma_z^a + \sin \phi \sigma_x^a$, where

$$\sigma_z^a = |0_a\rangle\langle 0_a| - |1_a\rangle\langle 1_a|,$$

$$\sigma_x^a = |0_a\rangle\langle 1_a| + |1_a\rangle\langle 0_a|.$$

7

On Bob's side we choose $\mathbf{B} = \sigma_z^b$ and $\mathbf{B}' = \cos \phi \sigma_z^b - \sin \phi \sigma_x^b$. The eigenvalues of \mathbf{A} and \mathbf{B} are clearly ± 1 , and it turns out that those of \mathbf{A}' and \mathbf{B}' are also ± 1 . For example, the eigenstates of $\cos \phi \sigma_z^a + \sin \phi \sigma_x^a$ are simply

$$|\tilde{0}\rangle = \cos \frac{\phi}{2} |0\rangle + \sin \frac{\phi}{2} |1\rangle,$$

$$|\tilde{1}\rangle = \sin \frac{\phi}{2} |0\rangle - \cos \frac{\phi}{2} |1\rangle.$$

8

Hence \mathbf{A}' corresponds to projectors on a basis that is rotated from that of \mathbf{A} by an angle $\phi/2$ (and similarly a rotation of $-\phi/2$ for \mathbf{B}, \mathbf{B}').

Now we can compute the necessary correlation functions using the standard quantum probability rules:

$$\begin{aligned}
\frac{1}{N} \sum_{n=1}^N A_n B_n &= \langle \mathbf{A} \otimes \mathbf{B} \rangle \\
&= \langle \mathbf{P}_0^a \mathbf{P}_0^b \rangle + \langle \mathbf{P}_1^a \mathbf{P}_1^b \rangle - \langle \mathbf{P}_0^a \mathbf{P}_1^b \rangle - \langle \mathbf{P}_1^a \mathbf{P}_0^b \rangle \\
&= -1.
\end{aligned}$$

9

Similarly,

$$\begin{aligned}
\frac{1}{N} \sum_{n=1}^N A_n B'_n &= \langle \mathbf{P}_0^a \cos \phi \sigma_z^b \rangle - \langle \mathbf{P}_0^a \sin \phi \sigma_x^b \rangle - \langle \mathbf{P}_1^a \cos \phi \sigma_z^b \rangle + \langle \mathbf{P}_1^a \sin \phi \sigma_x^b \rangle \\
&= -\frac{1}{2} \cos \phi - \frac{1}{2} \cos \phi = -\cos \phi. \\
\frac{1}{N} \sum_{n=1}^N A'_n B_n &= \langle \mathbf{P}_0^b \cos \phi \sigma_z^a \rangle + \langle \mathbf{P}_0^b \sin \phi \sigma_x^a \rangle - \langle \mathbf{P}_1^b \cos \phi \sigma_z^a \rangle - \langle \mathbf{P}_1^b \sin \phi \sigma_x^a \rangle \\
&= -\cos \phi. \\
\frac{1}{N} \sum_{n=1}^N A'_n B'_n &= \langle \cos^2 \phi \sigma_z^a \sigma_z^b \rangle + \langle \cos \phi \sin \phi \sigma_z^a \sigma_x^b \rangle - \langle \cos \phi \sin \phi \sigma_x^a \sigma_z^b \rangle \\
&\quad - \langle \sin^2 \phi \sigma_x^a \sigma_x^b \rangle \\
&= \frac{\cos^2 \phi}{2} (-1 - 1) - \frac{\sin^2 \phi}{2} (-1 - 1) = \sin^2 \phi - \cos^2 \phi \\
&= -\cos 2\phi.
\end{aligned}$$

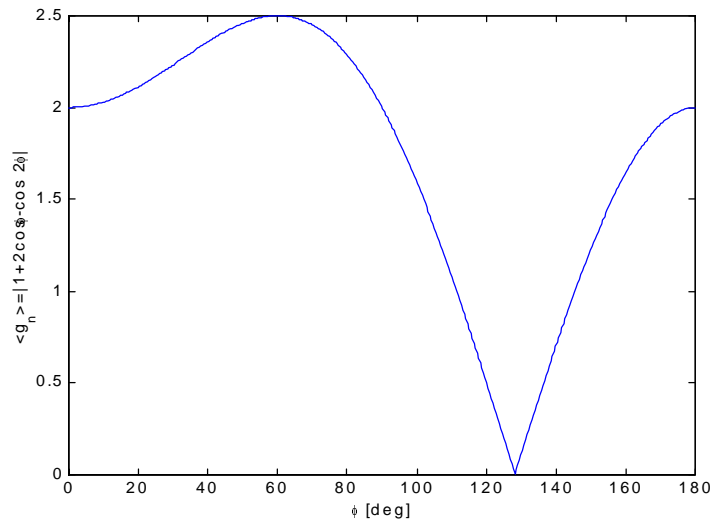
10

Finally, we can construct the overall quantity

$$\begin{aligned}
\frac{1}{N} \left| \sum_{n=1}^N g_n \right| &= |-1 - 2 \cos \phi + \cos 2\phi| \\
&= |1 + 2 \cos \phi - \cos 2\phi|.
\end{aligned}$$

11

Plotting this, we find that the Bell Inequality is violated ($\langle g_n \rangle > 2$) for $0 < \phi < 90^\circ$:



So what's going on here? From the graph we see that our Bell Inequality can be violated when the two possible measurements that Alice and Bob can perform correspond to projections on nonorthogonal bases. Hence what is being exploited here is the extra-strong "quantum correlation" between two particles that have been prepared in an entangled state such as

$$|\Psi_{ab}\rangle = \frac{1}{\sqrt{2}}(|0_a 1_b\rangle - |1_a 0_b\rangle). \quad 12$$

Recall from our last lecture that the quantum correlation between a and b survives changes of basis, whereas classical correlation

$$\rho_{ab} = \frac{1}{2}(|0_a 1_b\rangle\langle 0_a 1_b| + |1_a 0_b\rangle\langle 1_a 0_b|) \quad 13$$

does not.

From another point of view, the Bell Inequality we derived above tells us that there is no way to compose a mixed ensemble of factorizable quantum states $\{p_i, |\Psi_a^i\rangle \otimes |\Psi_b^i\rangle\}$ whose behavior simulates that of a *bona fide* entangled state.