Lecture 6, Thurs Feb 2: Mixed States

So far we've only talked about **pure states** (i.e., isolated quantum systems), but you can also have quantum superposition layered together with regular, old probabilistic uncertainty. This becomes extremely important when we talk about states where we're only measuring one part.

Last time we discussed the **Bell Pair**, and how if Alice measures her qubit in any basis, the state of Bob's qubit collapses to whichever state she got for her qubit. Even so, there's a formalism that helps us see why Bob can't do anything to learn which basis Alice makes her measurement in, and more generally, why Alice can't transmit *any* information instantaneously--in keeping with special relativity. This is the formalism of…

Mixed States

Which in some sense, are just probability distributions over quantum superpositions.
We can define a mixed state as a distribution over quantum states, so $\{p_i, \Psi_i\} = p_1, |\Psi_1\rangle, \cdots, p_n, |\Psi_n\rangle$ meaning that with probability p_i, the superposition is $|\Psi_i\rangle$.

^

Note that these don't have to be orthogonal

Thus, we can think of a pure state as a degenerate case of a mixed state where all the probabilities are 0 or 1.

The tricky thing about mixed states is that *different probability distributions over pure states, can give rise to exactly the same mixed state*. (We'll see an example shortly.) But to make manifest why information doesn't travel faster than light, we need a representation for mixed states that's unique. That's why we use:

Density Matrices

represented as $\rho = \sum_i p_i |\Psi_i\rangle\langle\Psi_i|$

$|\Psi_i\rangle\langle\Psi_i|$ is the **outer product** of Ψ with itself.

It's the matrix you get by multiplying $\begin{bmatrix}\alpha_1 \\ \vdots \\ \\ \alpha_N\end{bmatrix}\begin{bmatrix}\alpha_1^* & \cdots & & \alpha_N^*\end{bmatrix} = \begin{bmatrix}|\alpha_1|^2 & & & \\ & \ddots & \alpha_i\alpha_j^* & \\ & \alpha_i^*\alpha_j & \ddots & \\ & & & |\alpha_N|^2\end{bmatrix}$

Note that $\alpha_i\alpha_j^* = (\alpha_i^*\alpha_j)^*$, which means that the matrix is its own conjugate transpose

$\rho = \rho^\dagger$ That makes ρ a **Hermitian Matrix**.

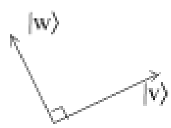
Some examples: $|0\rangle\langle 0| = \begin{bmatrix}1 & 0 \\ 0 & 0\end{bmatrix}$ $|1\rangle\langle 1| = \begin{bmatrix}0 & 0 \\ 0 & 1\end{bmatrix}$

Therefore an even mixture of them would be $\dfrac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2} = \begin{bmatrix}\frac{1}{2} & 0 \\ 0 & \frac{1}{2}\end{bmatrix} = \dfrac{I}{2}$

Similarly: $|+\rangle\langle+| = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$ $\qquad |-\rangle\langle-| = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$

And $\dfrac{|+\rangle\langle+| + |-\rangle\langle-|}{2} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} = \dfrac{I}{2}$

Note that an equal mixture of $|0\rangle$ and $|1\rangle$ is *different* from an equal superposition of $|0\rangle$ and $|1\rangle$ (a.k.a. $|+\rangle$), and so they have different density matrices. However, the mixture of $|0\rangle$ and $|1\rangle$ and the mixture of $|+\rangle$ and $|-\rangle$ have the same density matrix, which makes sense because Alice converting between the two bases in our Bell pair example should maintain Bob's density matrix representation of the state.



In fact, this is true of whichever basis Alice chooses: for any two orthogonal vectors $|v\rangle$ and $|w\rangle$, we have that $\dfrac{|v\rangle\langle v| + |w\rangle\langle w|}{2} = \dfrac{I}{2}$ (check this!)

Measuring ρ in the basis $|1\rangle, \ldots, |N\rangle$ gives us outcome $|i\rangle$ with probability:

$$\Pr(|i\rangle) = \rho_{ii} = \langle i|\rho|i\rangle$$

So, the diagonal entries of the density matrix directly represent probabilities.

You don't need to square them or anything because the Born Rule is already encoded in the density matrix (i.e. $(\alpha_1)(\alpha_1^*) = |\alpha_1|^2$)

$\begin{bmatrix} p_1 & & 0 \\ & \ddots & \\ 0 & & p_N \end{bmatrix}$ In particular, a density matrix that's diagonal is just a fancy way of writing a classical probability distribution.

While a pure state would look like $|\Psi\rangle\langle\Psi| = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$: that is, a matrix of rank one.

What if we want to measure a density matrix in a different basis?

Measuring ρ in the basis $\{|v\rangle, |w\rangle\}$ will give $Pr(|v\rangle) = \langle v|\rho|v\rangle$ and similarly for $|w\rangle$

You can think of a density matrix as encoding not just one but infinitely many probability distributions, because you can measure it in any basis.

The matrix $\frac{I}{2}$ that we've encountered above, as the even mixture of $|0\rangle$ and $|1\rangle$ (and also of $|+\rangle$ and $|-\rangle$) is called the **Maximally Mixed State**. This state is basically just the outcome of a classical coin flip, which gives it a special property:

Regardless of the basis we measure it in, both outcomes will be equally likely.

So for every basis $|v\rangle$,$|w\rangle$ we get the probabilities

$$\langle v|\tfrac{I}{2}|v\rangle = \tfrac{1}{2}\langle v|v\rangle = \tfrac{1}{2}$$

$$\langle w|\tfrac{I}{2}|w\rangle = \tfrac{1}{2}\langle w|w\rangle = \tfrac{1}{2}$$

This explains why Alice is unsuccessful in sending a message to Bob, by measuring her half of a Bell pair. Namely, because the maximally mixed state in any other basis is *still the maximally mixed state*.

So how do we handle unitary transformations with density matrices?

Since $\rho = \sum_i p_i |\Psi_i\rangle\langle\Psi_i|$, applying U to ρ means:

$$\sum_i p_i (U|\Psi_i\rangle)(U|\Psi_i\rangle)^\dagger = \sum_i p_i U|\Psi_i\rangle\langle\Psi_i|U^\dagger = U\rho U^\dagger$$

You can pull out the U's since it's the same one applied to each mixture.

It's worth noting that getting n^2 numbers in the density matrix isn't some formal artifact; we really do need all those extra parameters. What do the off-diagonal entries represent?

$$|+\rangle\langle+| = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

These are where all the 'quantumness' resides.
It's where the interference between $|0\rangle$ and $|1\rangle$ is represented.

The off-diagonal entries can vary depending on relative phase:

$|+\rangle$ has positive off-diagonal entries

$|-\rangle$ has negative off-diagonal entries

$$|i\rangle\langle i| = \begin{bmatrix} \frac{1}{2} & -\frac{i}{2} \\ \frac{i}{2} & \frac{1}{2} \end{bmatrix}$$

Later we'll see that as a quantum system interacts with the environment, the off-diagonal entries tend to get pushed down toward 0.

The density matrices in experimental quantum papers typically look like $\begin{bmatrix} \frac{1}{2} & \varepsilon \\ \varepsilon & \frac{1}{2} \end{bmatrix}$.
The bigger the off-diagonal values, the better the experiment, because it represents them seeing more of the quantum effect!

Which matrices can arise as density matrices?

We're effectively asking: What constraints does the form $\sum_i p_i |\Psi_i\rangle\langle\Psi_i|$ put on the matrix ρ?
Well, such ρ a must be:

- Square
- Hermitian
- $\sum_i p_{ii} = 1$ (which is to say: the **trace**, $Tr(\rho) = 1$)

Could $M = \begin{bmatrix} \frac{1}{2} & -10 \\ -10 & \frac{1}{2} \end{bmatrix}$ be a density matrix?

No. Measuring this in the $|+\rangle, |-\rangle$ basis would give $\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} M \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \frac{19}{2}$

Bad!

Remember that you can always transform ρ to $U\rho U^\dagger$, whose diagonal then has to be a probability distribution for all U. If we want that condition to hold, then in linear algebra terms, we need to add the restriction:

- All eigenvalues are non-negative (aka being **PSD: Positive Semidefinite**)

As a refresher: For the matrix ρ, the eigenvectors $|\Psi\rangle$ are the vectors that satisfy the equation:

$\rho|\Psi\rangle = \lambda|\Psi\rangle$ for some eigenvalue λ

If we had a negative eigenvector $|\Psi\rangle$,

the probability $\langle\Psi|\rho|\Psi\rangle = \lambda$ would be negative, which is nonsense.

Could we have missed a condition? Let's check.

We claim: any Hermitian PSD matrix with trace 1 can arise as a density matrix of a quantum state.

For such a ρ, we can represent it in the form $\sum_i \lambda_i |\Psi_i\rangle\langle\Psi_i|$ where the $|\Psi_i\rangle$ are the (unit) eigenvectors.

Then $\langle\Psi_i|\rho|\Psi_i\rangle = \lambda_i$, so the λ_i's sum to Tr(ρ)=1.

This process of obtaining eigenvalues and eigenvectors is called **eigendecomposition**.

We know the eigenvalues will be real because the matrix is Hermitian,

They're non-negative because the matrix is PSD.

One quantity you can always compute for density matrices is:

Rank

$rank(\rho)$ = the number of non-zero eigenvalues λ_i (counting with multiplicity)

A density matrix of rank n might look like $\begin{bmatrix} p_1 & & 0 \\ & \ddots & \\ 0 & & p_n \end{bmatrix}$

While a density matrix of rank 1 represents a pure state.

We know from linear algebra that the rank of an $n \times n$ matrix is always at most n. Physically, this means that every n-dimensional mixed state can be written as a mixture of at most n pure states.

In general, rank tells you the number of pure states that you have to mix to reach a given mixed state.

Now, consider the 2-qubit pure state $\frac{|00\rangle + |01\rangle + |10\rangle}{\sqrt{3}}$.

We'll give the first qubit to Alice and the second to Bob.
How does Bob calculate his density matrix?

By picking some orthogonal basis for Alice's side.

You can rewrite the state as $\sqrt{\frac{2}{3}}|0\rangle|+\rangle + \sqrt{\frac{1}{3}}|1\rangle|0\rangle$, which lets you calculate Bob's density matrix as:

$$\frac{2}{3}|+\rangle\langle+| + \frac{1}{3}|0\rangle\langle 0|$$
$$= \frac{2}{3}\begin{bmatrix}\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2}\end{bmatrix} + \frac{1}{3}\begin{bmatrix}1 & 0 \\ 0 & 0\end{bmatrix}$$
$$= \begin{bmatrix}\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3}\end{bmatrix}$$

In general, if you have a bipartite pure state, it'll look like $\sum_{i,j=1}^{N} \alpha_{ij}|i\rangle|j\rangle = |\Psi\rangle$

And you can get Bob's local density matrix

$$(\rho_{\text{Bob}})_{j,j'} = \sum_i \alpha_{ij}\alpha^*_{ij'}$$

The process of going from a pure state of a composite system, to the mixed state of part of the system, is called **tracing out**.

The Key Points:

1) A density matrix encodes all and only what is physically observable
   - Two quantum states will lead to different probabilities *iff* they have different density matrices
2) No-Communication Theorem
   - If Alice and Bob share an entangled state, nothing Alice chooses to do will have any effect on Bob's density matrix.

In other words, there's no observable effect on Bob's end. Which is the fundamental reason that quantum mechanics *is* compatible with the limitations of relativity.

We've already seen particular examples of both statements. But both of them hold in full generality, and you'll prove that in your homework!

OK, just to get you started a bit: recall that the No Communication Theorem says that, if Alice and Bob share an entangled state

$$|\Psi\rangle = \sum_{i,j=1}^{N} \alpha_{ij}|i\rangle_{\text{Alice}}|j\rangle_{\text{Bob}}$$

there's nothing that Alice can do to her subsystem that affects Bob's density matrix.

You already have the tools to prove this: just calculate Bob's density matrix, then apply a unitary transformation to Alice's side, then see if Bob's density matrix changes. Or have Alice measure her qubit, and see if *that* changes Bob's density matrix changes.

Note that if we condition on the *outcome* of Alice's measurement, then we do need to update Bob's density matrix to reflect the new knowledge: if Alice sees i then Bob sees j, etc. But that's not terribly surprising, since the same would also be true even with classical correlation! In particular, this doesn't provide a mechanism for faster-than-light communication.

To review, we've seen three different types of states in play, each more general than the last:

- Basis States, or Classical States
  - exist in a computational basis $|i\rangle$
- Pure States
  - superpositions of basis states $|\Psi\rangle = \sum \alpha_i |i\rangle$
- Mixed States
  - classical probability distributions over pure states $\rho = \sum \rho_i |\Psi_i\rangle\langle\Psi_i|$

Which represents the actual physical reality: pure or mixed states?

It's complicated. Sometimes we use density matrices to represent our probabilistic ignorance of a pure state. But when we look at part of an entangled state, a mixed state is the most complete representation possible that only talks about the part that we're looking at. We'll generally just focus on what these representations are useful for.